LITTLE KNOWN FACTS ABOUT EDUCATION: AN EMPIRICAL ANALYSIS

Richard John Murphy

Thesis submitted for a Ph.D. in Economics.

Department of Economics

University College London

May 2014

Declaration

I, Richard John Murphy confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Richard John Murphy

31st May 2014

Statement of conjoint work

Three out of my four chapters that form part of this thesis involve conjoint work, as specified below:

Chapter II "Top of The Class: The Importance of Ordinal Rank" is conjoint work with Felix Weinhardt. Overall, my contribution amounts to two thirds of the total paper.

Chapter III "Paying Out and Crowding Out? The Globalisation of Higher Education" is conjoint work with Steve Machin. Overall, my contribution amounts to two thirds of the total paper.

Chapter V "Ill Communication: Technology, Distraction & Workplace Productivity" is conjoint work with Louis-Philippe Beland. Overall, my contribution amounts to two thirds of the total paper.

Acknowledgements

Throughout my PhD I have received support, advice and encouragement from a great many people, so undoubtedly some names will be mistakenly omitted, and so apologies in advance if you are not mentioned. First and foremost I would like to thank my supervisor Steve Machin, who has provided me with invaluable guidance and insights, especially regarding graph formatting in Stata. Without his initial encouragement to pursue a Ph.D I would still be a research assistant. Not to say being a research assistant is bad. During my time at the Centre for Economic Performance as a research assistant, I learned much from my colleagues and I would like to thank all the staff of CEP. Especially Felix Weinhardt for making research a truly enjoyable experience. Gill Wyness for answering countless questions and correcting my English. And Amy Challen for her challenging questions and her re-writing or my ESRC studentship application for which I am eternally thankful. Others who have also helped with my writing and general life advice, to whom I'm grateful, are Kim Keller, Lee Elliot Major, and Nigel Rogers. Despite spending little time in the department, my countless discussions with my fellow UCL students have been a constant source of inspiration. I would like to thank David Card for letting me visit UC Berkeley and then for letting me come back – it was a life changing experience.

Last but not least, I would like to thank my family and friends for all their support, and persistence in asking when I will finally complete. I would particularly like to thank my parents for everything they have done over the years, and especially for badgering my primary school head teacher into putting me in for the 11+ exam, without which I would not be here now.

Funding acknowledgements

I gratefully acknowledge the funding that made my Ph.D. possible. I was funded by the Economics Department at UCL through the Gorman Grant, and by the ESRC through their scholarship. I would also like to thank the CEP for allowing me to work on my research whilst being employed at the centre.

Abstract

The thesis consists of four chapters utilising applied micro-econometric techniques to develop a deeper understanding of the education sector. I apply traditional economic concepts such as productivity, immigration, insurance and technological innovation to the field of education economics.

Chapter one considers the consequences of academic rank in primary school on later test scores. Using administrative data tracking the student population in England, I estimate the impact of rank on later attainment through the variation in the test-score distributions across schools. The positive impact of rank on attainment helps to explain some puzzles in the education literature, such as the lack of impact of selective schools.

The second chapter involves immigration and investigates how the influx of overseas students has affected enrolment of domestic students at UK universities. Using administrative data, I employ methods used in the labour literature to model crowd-out. I find no evidence of crowd-out of domestic students, and some evidence of crowd in amongst postgraduate students.

Chapter three establishes the threat of accusations as new source of demand for trade union membership amongst teachers. I model union membership as legal insurance, where demand is determined by the threat of accusations. I measure threat primarily through the incidence of media stories concerning teachers in the local area. Combining these data with union membership data from Labour Force Surveys, I find that unionisation rates increase with media coverage of allegations.

The final chapter is an estimation of the impact of restricting technology in the workplace on productivity. This is applied to the education setting using the autonomous decisions by schools to ban mobile phones. Obtaining histories of phone policies through surveys and combining this with administrative data on individual pupil level attainment, I use a difference in difference analysis to estimate the impact on student performance.

Contents

Index of Tables, Figures and Appendices

Chapter 2. 18

Chapter 2. Appendices

CHAPTER 1. INTRODUCTION

In this thesis I apply traditional economic concepts to various aspects of the human capital formation process and education systems more generally. As we continue to move into the new knowledge economy, human capital is becoming ever more important. Therefore questions such as, how do children learn best? what influences subject choice? and which institutional features can be improved? become increasingly significant. However, the topic of education is not new to the field of Economics. Smith (1776), Becker (1964), and Mincer (1974) all made seminal contributions to how the complex topic of education can be modelled through use of economics. These combined with methodological advances (Heckman 1979; Todd and Wolpin, 2003) have led to significant gains in our understanding of how we learn. Policy makers increasingly turn to economists when designing and evaluating policies. The four chapters of this thesis all address topics that should be of interest to policy makers, but equally find answers that may surprise.

Two of the chapters concern students at school and examine factors that impact on their educational outcomes. Chapter 2 explores how individuals' academic rank amongst their peers in primary school affects their academic achievement in secondary school. The widely held traditional opinion is that having higher attaining peers is always better for a child's attainment. However we show that an individual who has a higher rank, and hence has lower performing peers, goes on to do better in secondary school.

It is a natural instinct for humans to make comparisons (Festinger, 1954), yet little is known about the long run consequences of ordinal rank amongst peers on productivity. This chapter makes use of administrative data on the entire student population in England in their transition from primary to secondary school. We study the effect of rank conditional on prior relative test scores as well as age varying school and student effects. We use the natural variation in test score distributions across schools and across subjects within pupils to separately identify the impact of rank from other cardinal measures of attainment. A student with a given above average test score would be ranked higher in school cohorts with more compressed distributions. The main finding is that being highly ranked during primary school has large and robust effects on secondary school achievement. The effect of being top of the class is more important for boys than girls.

The paper goes on to examine potential channels that could account for these results, including learning about own ability. The chapter concludes that the most likely explanation is that the development of non-cognitive skills, such as confidence is improved when surrounded by peers who perform tasks worse than one's self. Supporting evidence is provided by combining the administrative data with survey data on twelve thousand students which directly measures task specific confidence. This result speaks to why integration programs such as bussing (Angrist and Lang 2004) or moving students to less deprived neighbourhoods (Kling at al. 2007) have typically failed to find improvements of low attaining students. Moreover it may also help to explain why students who just pass the entrance exams for selective schools fail to gain more than those that don't (Cullen et. al., 2006, Clark, 2010). The implication to policy is to highlight the importance of non-cognitive skills in the formation of human capital.

The other chapter that examines student achievement directly is Chapter V. Information technology is commonly viewed as increasing productivity (Kruger, 1993; Malamud, 2011). However, since Solow's (1987) infamous statement "You can see the computer age everywhere but in the productivity statistics", there has been a parallel literature that argues there have not been large improvements in productivity (DiNardo and Pischke, 1997) and offers four potential reasons for that (Brynjolfsson, 1991). This chapter adds to the literature by providing further evidence for another reason, that given the multifunctional nature of modern communication technology it has become a lot easier to be distracted. When this technology is easily to hand it effectively lowers the transition cost of shifting from work to procrastination to near zero.

This type of technology is common to many workplaces today, which leads to the question of how much productivity is lost through this type of distraction. The chapter estimates the impact of the removal of mobile phones on student productivity. Using an educational context has the advantage that individual productivity measures are readily observable in the form of externally marked test scores. We exploit schools' autonomy in mobile phone policies, meaning that schools introduced phone bans at different times. We use this variation in the implementation date to estimate the impact of the bans on student achievement through a difference in difference analysis. We collected the information on school policies through a survey of four large UK cities and combined it with long run administrative data on student performance. Our results indicate that after phones have been banned from schools, student grades on high stakes national examinations significantly improve, and that these gains are driven by the previously low achieving students. The chapter shows that access to technology is not always beneficial and that banning mobile phones from school premises can be a low cost policy to reduce educational inequality.

The other two chapters of the thesis take a step back and consider the settings of the educational system itself. Chapter III investigates how the rapid influx of overseas students to UK higher education has affected the number of domestic students. As the demand for highly skilled labour is growing globally, higher education and who has access to it become increasingly important topics. In the UK overseas students pay considerably higher tuition fees than domestic students, but despite this their numbers have tripled since 1995. A critical policy question is therefore, whether these students take the places of natives or whether the additional income they generate acts to subsidise domestic students.

We apply displacement models from the labour immigration literature to this education setting (Peri and Sparber, 2011). In order to establish causal estimates we apply two different methods. Firstly, we use the historical share of students from a sending country attending a university department combined with current national changes in the stock of students from this country as a shift-share instrument. Secondly, we use an exogenous change in the Chinese visa regulations and exchange rate in combination with strong revealed subject preferences as a predictor of overseas student growth across and within universities over time.

Using administrative data on the entire UK Higher Education population over the 1994/5 to 2011/12 academic years, we find no evidence that the big rise in international students enrolling in UK universities has crowded out domestic students. This is the case at undergraduate level as well as for taught and research postgraduates. Indeed, we find evidence of cross-subsidisation for postgraduates, especially on masters programmes. A possible reason for the stronger effect for postgraduates than for undergraduates are the government regulations limiting the growth of student numbers on undergraduate degrees, which do not apply as strictly to postgraduate degrees.

In the remaining chapter, I explore a curious development in the teacher labour market, namely that teachers' unionisation rate has been increasing during a period of general decline in union membership. There is a large body of research establishing that teachers are the most important factor within schools in determining student achievement (Rockoff, 2004; Rivkin et al. 2005). Therefore understanding how the teacher labour market operates is crucial, especially given the ongoing debate regarding union power and student outcomes (Hoxby, 1996; Lovenheim 2009). This chapter recognises the threat of accusation as new source of demand for union representation and analyses how the threat of accusation has increased union density in specific labour markets. The fact that society has become increasingly litigious, could have many repercussions on labour markets, especially for those sectors where employees have unsupervised interactions with vulnerable groups. A rational response to such changes would be an increase in demand for insurance against these risks.

I model union membership as a form of private legal insurance, where the decision to join is partly determined by the perceived threat of having an allegation made against the agent. This is examined by estimating the demand for union membership amongst UK teachers, which has been increasing over the last twenty years. I use media coverage of allegations relating to local teachers as a shock to the perceived threat. I find that unionisation rates increase with media coverage of relevant litigation at the regional and national levels. Ten relevant news stories in a region increase the probability of union membership by 5 percentage points. The size of the effect is dependent on the similarity of the teacher to the one mentioned in the story.

This chapter provides a reason for why the demand for union membership in the teaching and related sectors has increased. Moreover, I provide a further answer to the puzzle of why individuals choose to join a union even if they could free ride and receive the higher pay and working conditions derived through union action without having to pay the union dues. Unions offering a private, excludable service can maintain demand for membership, as long as demand for that service remains. The implication for policy is that there may be an increasing unmet demand for union membership in previously under unionised service sectors. Furthermore this means that governments have the power to temper the demand for union membership through introducing regulations that protect individuals from allegation. Suggestive evidence of this can be seen in the fall in union density post 2005 governmental reforms on newspaper reporting; the union density continued to fall despite the worsening in economic conditions which is traditionally thought of as a key driver of union demand (Blanchflower et al. 1990).

References

Angrist, J and Lang, K. (2004). “Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program," *American Economic Review*, 94 (5), pp.1613-1634

Blanchflower, D. Crouchley. R. Estrin, S. and Oswald, A., (1990). Unemployment And The Demand For Unions, *Working Papers 372*, London School of Economics - Centre for Labour Economics.

Becker, G. S. (1964). Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. Chicago, University of Chicago Press.

Brynjolfsson, E. (1991). Information technology and the re-organization of work: theory and evidence. *Ph.D. Dissertation*, MIT Sloan School of Management, 1991

Clark, D. (2010) “Selective Schools and Academic Achievement” *The B.E. Journal of Economic Analysis and Policy*, 10(1), pp.1935-1682.

Cullen, J.B., Jacob, B.A. and Levitt S., (2006). “The effect of school choice on participants: evidence from randomised lotteries”, *Econometrica*, 74(5), pp.1191-1230

DiNardo, J. E. and Pischke, J.S. (1997). The returns to computer use revisited: Have pencils changed the wage structure too? *The Quarterly Journal of Economics* 112(1), 291–303.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, pp. 117–140.

Heckman, J. (1979). Sample selection bias as a specification error, *Econometrica*, 47, pp. 153-61.

Hoxby, C.M. (1996). How Teachers' Unions Affect Education Production *The Quarterly Journal of Economics*, Vol. 111, No. 3 (Aug., 1996), pp. 671-718

Kling, J., Liebman, J. and Katz, L. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, 25(1), pp. 83-119.

Krueger, A. B. (1993). How computers have changed the wage structure: evidence from microdata, 1984–1989. *The Quarterly Journal of Economics,* 108(1), 33–60.

Lovenheim M. F. (2009). The Effect of Teachers’ Unions on Education Production: Evidence from Union Election Certifications in Three Midwestern States. *Journal of Labor Economics*, Vol. 27, No. 4 (October 2009), pp. 525-587

Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research

Malamud, O. and Pop-Eleches C. (2011). Home computer use and the development of human capital. *The Quarterly Journal of Economics* 126(2), 987–1027.

Peri, G. and Sparber C. (2011) Assessing Inherent Model Bias: An Application To Native Displacement In Response To Immigration, Journal of Urban Economics, 69, 82–91.

Todd P.E. and Wolpin K.I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement, *The Economic Journal*, 113 F3-F33.

Smith, A. (1776). An Inquiry into the Nature and Causes of the Wealth of Nations. London: Methuen & Co., Ltd. . 1904, Library of Economics and Liberty [Online] available from http://www.econlib.org/library/Smith/smWN.html; accessed 30 May 2014; Internet.

Solow, R. (1987), We'd better watch out. *New York Times* (July 12) Book Review, 36.

CHAPTER 2.

TOP OF THE CLASS:

THE IMPORTANCE OF ORDINAL RANK

I would like to thank my co-author Felix Weinhardt for his contribution. Moreover we thank Johannes Abeler, Esteban Aucejo, Thomas Breda, Andrew Clark, Susan Dynarski, Ben Faber, Eric Hanushek, Brian Jacob, Pat Kline, Steve Machin, Magne Mogstad, Imran Rasul, Olmo Silva, Kenneth Wolpin, Gill Wyness, and participants of the CEP Labour Market Workshop, the Sussex University, Queen Mary University and Royal Holloway-University departmental seminars, the CMPO seminar group, the RES Annual Conference panel, IWAEE, the Trondheim Educational Governance Conference, the SOLE conference, CEP Annual Conference, the UCL PhD Seminar, the BeNA Berlin Seminar, IFS seminar and the CEE Education Group for valuable feedback and comments.

1 Introduction

It is human nature to make comparisons against one's peers. Individuals make comparisons in terms of characteristics, traits and abilities in tasks (Festinger, 1954). However, individuals also often use cognitive shortcuts to simplify decision-making (Tversky and Kahneman, 1974). One such shortcut would be to use simple ordinal rank instead of detailed cardinal information. Rather than working out where one stands in relation to the group mean, one might say 'I am taller than Gill but shorter than Sarah'. In this simplified way of conceptualising the world, when we are making decisions one would be placing weight on ordinal rank as well as relative or absolute information. Indeed, it has recently been shown that ordinal rank, in addition to relative position, is used when individuals make comparisons with others (Brown et al., 2008; Kuziemko et al., 2011; Card et al., 2012). If people are ranking themselves amongst their peers, then ordinal in additional to cardinal information has the potential to affect investment decisions. These could in in turn determine later productivity, through various mechanisms such as learning about ability or the development of confidence.

This paper examines, in the context of education, the additional impact of ordinal rank on subsequent productivity. Students in England take externally marked national exams at the end of primary school at age 11. We use this to calculate their rank amongst their primary school peers in three subjects. These students then start attending secondary schools with a new set of peers and are tested again in the same subjects three years later. We use this setting to estimate the effect of age-11 rank on age-14 test scores, in a new peer environment conditional on prior (age-11) relative test scores, in our main specification. To do so, we use administrative data on the entire English public school population as they move from a primary into secondary education.[1]

The rank parameter is identified from the variation in test score distributions across and within primary schools cohorts, so that the same score relative to a school mean can have different ranks. Our estimates show that being highly ranked amongst your peers in a subject has large and robust effects on later performance in that same subject. Moreover, the impact of rank is significant across the entire rank distribution. These estimates use the school-by-subject-by-cohort variation in rank for a given test score and therefore allow for gains from individuals being ranked highly in one subject to impact on results on other subjects. We also provide more demanding within-student specifications, which absorb the average growth rate of a student between age 11 and 14,

[1] Public schools account for 93% of the total student population in England. Comparable data for the remaining 7% attending private schools are not available.

therefore removing subject-spillovers and so reflect student specialisation. In these specifications, the variation used for estimation is the within-student-across-subject differences in rank conditional on test scores and average prior peer quality. We argue that conditioning on these age-11 test scores, primary-subject-cohort effects and individual student effects, the rank of a student in a subject within primary school is effectively random.

Notably, primary school peers determine the rank measure but we estimate its effects on outcomes after the transition to secondary school. This makes our approach resilient to reflection problems (Manski, 1993) as the average student has 87% new peers in secondary school. We are therefore not relating individual and group outcomes from within the same peer group, as cautioned against by Angrist (2013). Moreover, our estimation sidesteps the standard issue of including a lagged dependent variable and individual effects simultaneously (Nickell, 1981), as the individual effects are recovered from test scores across subjects of a student at age 14, rather than from average test scores over time.

The effects of rank that we present are large in the context of the education literature, with a one standard deviation increase in rank improving age-14 test scores by 0.08 standard deviations. This is of comparable magnitude to being taught by a teacher one standard deviation above average (Aaronson, et al., 2007; Hanushek et al., 2005). As expected, the estimates relying on within-student variation in primary rank, conditional on ability, are smaller. Here, a one standard deviation increase in rank improves subsequent test scores in that subject by 0.055 within student standard deviations. This would mean being ranked at the 75th percentile of your primary school peers in a subject as opposed to the 25th percentile, improves age 14 test scores by 0.2 standard deviations in that same subject.

The paper goes on to examine the nature of these effects and finds that they exist throughout the rank distribution, implying that students accurately place themselves within their class, despite not being explicitly informed of their rank. This is likely to occur due to the repeated interactions among peers throughout the six years of primary school as well as seating arrangements that that reflect rank positions in many English primary schools[2]. Moreover, for nearly all rank positions boys are more affected, both positively and negatively, than girls. Boys at the top of the class in a subject gain four

[2] In English Primary schools it is common for students to be seated at tables of four and for them to be set by pupil ability. Students can be sat at the 'top table' or the 'bottom/naughty table'. This could assist students in establishing where they rank amongst all class members through a form of batch algorithm, e.g. 'I'm on top table, but I'm the worst, therefore I'm fourth best.'

times more than comparable girls. Low-income students also gain more from being top of the class but are less negatively affected by being ranked below the median.

Having presented this range of findings, the paper examines and tests threats to identification such as other forms of peer effects, measurement error and sorting to schools by parents. Using simulations we demonstrate that our findings are robust against non-linear peer effects, large measurement errors and are not just a statistical artefact. Using additional survey data, we further show that parental occupations predict subject-specific primary attainment but not rank. For example, children of accountants do better in Maths than in English. In contrast, parental occupational background has no relation to ordinal rank conditional on attainment. We interpret this as direct evidence for our main identification assumption that primary rank, conditional on class means and own attainment, is effectively random.

Finally the paper discusses a number of mechanisms that could account for these results: learning about own ability (Ertac, 2006); competitiveness; external (parental) investment by task; and environment favouring certain ranks; but provides evidence that the mechanism that best accommodates all the findings is through the development of non-cognitive skills such as confidence. Combining our administrative data with survey data containing direct measures of subject-specific confidence, we show that those who ranked higher in primary school have larger measures of later confidence, conditional on relative test scores and student effects. Mirroring our findings on attainment, we find that boys' confidence is more affected by their school rank than girls' confidence.

To build intuition for the effect of confidence, consider a child being the best in their neighbourhood at basketball. She will consider herself to be good at basketball, gaining confidence in her basketball abilities and resulting in her enjoying basketball more. This would then lead her to invest more time in playing basketball and so further develop her skills. Similarly, in the labour market, individuals rate their productivity in a task relative to their colleagues, and this in turn could determine in which field they specialize. More relevant to this paper, one might consider one's own school career. Upon starting school, we may not know which subjects we are good at. But, through ranking ourselves relative to our peers, we develop a sense that we are a 'math person' or a 'language person'. A 'math person' would be more confident in solving mathematical problems and enjoy math more, and therefore may invest more time into math homework, all of which could be reflected in their future math test scores.

We believe this paper has two main contributions. First, to the best of the authors' knowledge this is the first large-scale study to document the effects of ordinal rank in a task on later productivity. Critically, this study documents an additional effect of ordinal rank, after controlling for prior achievement and the relative distances between peers, i.e. cardinal measures of performance. Therefore, we believe rank could be considered a new factor in the education production function.[3] Besides implications on partial equilibrium considerations of parents regarding the choice of the best school for their children, this finding has more general implications relating to informational transparency and productivity. For instance managers or teachers could improve productivity by emphasising an individual's local rank position if that individual has a high rank. Alternatively, if an individual is in a high performing peer group and therefore may have a low local rank but ranks high globally, a manager should make the global rank more salient.

Secondly, we believe that the result that ordinal rank matters for later outcomes has the potential to add to the explanation of findings in the following education topics where placing individuals amongst high-performing peers has had mixed results: school integration (Angrist and Lang, 2004; Kling *et al.* 2007) selective schools (Cullen, Jacob and Levitt, 2006; Clark 2010); and affirmative action (Arcidiacono *et al.* 2012; Robles and Krishna, 2012). Moreover the finding that rank may exacerbate early differences in achievement due to individual investment decisions based on relative performance contributes to the literatures on ethnicity (Fryer and Levitt, 2006; Hanushek and Rivkin 2006; 2009), gender (Burgess et al, 2004; Machin & McNally, 2005) and relative age in cohort (i.e. Black *et al.*, 2011).

The remainder of the paper is laid out as follows. Section 2 reviews the literature on social comparisons. Section 3 sets out the empirical strategy and how the rank parameter is separately identified from relative achievement. This is followed by a brief description of the UK educational system, the administrative data, as well as the definition of rank used. Section 5 sets out the main results, nonlinearities and the heterogeneity by gender and parental income. Section 6 discusses and tests threats to identification such as peer effects, measurement error and endogenous sorting. Section 7 discusses potential mechanisms and provides additional survey evidence. Section 8 outlines other topics in education, which corroborate these findings. Finally, we conclude and discuss policy implications.

[3] There is a broad range of literature on the determinants of academic achievements, including natural ability (Watkins et al., 2007), family background (Hoxby, 2001), school inputs (Hanushek, 2006; Page et al., 2010), peer effects (Carrell et al., 2009; Lavy et al., 2012), and non-cognitive skills (Heckman et al., 2006); however, rank position has not yet been researched.

2 Related Literature

The importance of ordinal rank rather than relative position for individuals was first forwarded by Parducci (1965) with range frequency theory. This has the theoretical prediction that comparisons are based upon ordinal position of items within a comparison set. This prediction has been illustrated empirically recently by Brown *et al.* (2008) and Card *et al.* (2012), who show an individual's rank in addition to relative position in an income distribution is an important determinant of satisfaction. However, the economic literature on rank effects on productivity is sparse.[4]

A related study on rank and informational transparency finds that providing employees with their productivity rank within the firm increased output throughout the productivity distribution (Blanes i Vidal and Nossol 2011). This is explained by workers becoming concerned about their rank position, as the impact occurred after the feedback policy was announced but before the information was released.[5] Genakos and Pagliero (2012) find that in a tournament setting, where payoffs are based on relative performance and with continuous rank feedback, performance decreases as individuals are ranked higher.[6] In both of these papers, individuals are concerned about their relative positions amongst their immediate peers. The education setting of this study varies in two critical ways. Firstly, students are graded on their absolute performance according to national scales, rather than relative to their peers. Secondly, we are estimating the effect of rank amongst previous peers on contemporaneous test scores, and not the effects of rank within the same peer group. Moreover, whilst both of these papers use rank measurements, neither additionally controls for relative distances, and are therefore not separating rank effects from any cardinal relative effects.

The importance of ordinal rank in addition to relative position has been empirically illustrated by Brown et al. (2008) and Card et al. (2012). Most related to our paper, Clark et al. (2010) compares directly the importance of ordinal rather than relative position on discretionary work effort. They find that an employee's income rank was a

[4] The discussion of social comparisons is often framed in the form of peer effects (Falk and Ichino, 2006;i.e. Mas and Moretti, 2009, Carrell et al., 2009; Lavy et al., 2012) or the introduction of relative achievement feedback mechanisms (Eriksson et al. 2009; Azmat and Iriberri, 2010). These studies tend to find positive effects of peer quality on contemporaneous productivity, and that relative performance feedback increases productivity when there are piece rate incentives.

[5] Kosfeld and Neckerman (2011) examine the use of rankings as a non-monetary incentive and find increases in productivity. Specific to education, Jalava, Joensen and Pellas (2013) find that rank based grading increases test performance.

[6] Brown (2011) shows in a tournament setting that when an individual of known outstanding ability (high prior high rank is known) is placed into a group those ranked immediately below them, have a large fall in productivity compared to low ranked participants.

stronger determinant of stated work effort compared with the average reference group income and so conclude that comparisons are ordinal rather than cardinal. This is similar to our paper as we also in effect estimate effects of rank and relative position, but different because we observe rank effects in a real effort setting rather than in stated amounts.

3 Empirical strategy

3.1 The measurement of rank net of ability and cardinal factors

In order to identify the effect of primary school rank on later outcomes we require variation in rank for a given ability. Moreover, to separately identify the effect of ordinal rank from relative position requires variation in rank for a given distance from mean peer achievement. This comes about through the variation in test score distributions across school cohorts, within and across primary schools, so that students with the same test scores and same distance to peer means can have different ranks. Furthermore, as test score distributions vary across subjects within a school and cohort, a single student with the same score in all three subjects, as well as the same relative distances to peer mean achievements in these subjects, could have different subject ranks. To see this, consider the case illustrated by Figure 1, which shows unimodal and bimodal distributions of hypothetical English test scores for two cohorts of students in a primary school. The school has a very similar intake of students year-on-year, with the same number of students, and moreover the same mean, minimum and maximum student test scores despite having different distributions. As both cohorts have the same mean test scores, students who achieved the same absolute test scores across cohorts (Y), would also have the same relative score compared to the mean of their peers. However, the cohorts have different test score distributions, in the first students are more clusterd around the mean score and in the second test scores are more dispersed and has a bimodal distribution. Due to these different distributions, a student who scored Y in the unimodal cohort is ranked second, whilst the one in the bimodal cohort is ranked fifth.[7]

Note that the rank effect that is identified conditional on the student fixed effect differs in interpretation from school-subject-cohort effects that we just illustrated because it uses the variation in test score distributions *across subjects* within a cohort.

[7] This is similar to Brown et al. (2008) who rely on the variation in the earnings distributions of a subset of workers across firms to separately identify the effect of relative earnings and ranks in earnings on employee satisfaction.

The variation used here is analogous to Figure 1 but comparing differences in distributions across subjects rather than within subjects and over time.

3.2 A rank-augmented education production function

This section uses the standard education production function approach to derive a rank-augmented value added specification that can be used to identify the effect of primary school rank, measured as outlined in section 3.1, on subsequent outcomes.

To begin, we consider a basic contemporaneous education production function, using the framework as set out in Todd and Wolpin (2003), for student *i* studying subject *s* in primary school *j*, cohort *c* and in time period $t = [1,2]$:

$$Y_{ijsct} = X_i'\beta + v_{ijsct} \quad (1)$$

$$v_{ijsct} = \mu_{jsc} + \tau_i + \varepsilon_{ijsct}$$

where *Y* denotes national academic percentile rank in subject *s* at time *t* and X is a vector of observable non-time varying characteristics of the student. Here β represents the permanent impact of these non-time varying observable characteristics on academic achievement. In this specification there are two time periods, in period one students attend primary school and in the second period students attend secondary schools. The error term v_{ijsct} has three components; μ_{jsc} represents the permanent unobserved effects of being taught subject *j* in primary school *s* in cohort *c*. This could reflect the effect of a teacher being particularly good at teaching maths in one year but not English, or that a student's peers were good in English but not in science; τ_i represents permanent unobserved student characteristics, this would include any stable parental inputs or natural ability of the child; ε_{ijsct} is the idiosyncratic time specific error which includes secondary school effects. Under this restrictive specification only ε_{ijsct} could cause the national academic rank of a student to change between primary and secondary school, as all other factors are permanent and have the same impact over time.

This is a restrictive assumption, as the impact of observable and unobservable characteristics are likely to change as the student ages. One could imagine that neighbourhood effects may grow in importance as the child grows older, and that the effects of primary school are more important when the child is young and attending that school. Therefore we extend the model by allow for time-varying effects of these characteristics:

$$Y_{ijsct} = \beta_{Rank\,t} R_{ijsc} + X_i'\beta + X_i'\beta_t + v_{ijsc} \quad (2)$$

$$v_{ijsc} = \mu_{jsc} + \mu_{jsct} + \tau_i + \tau_{it} + \varepsilon_{ijsc}$$

where β_t allows for the effect of student characteristics to vary over time. We have also introduced the parameter of interest $\beta_{Rank\ t}$, which is the effect of having rank R_{ijsc}, in subject s in cohort c and in primary school j on student achievement in that subject in the subsequent period *t*. We are interested in longer-run effects of rank positions students had during early education stages. We therefore assume that there is no effect of rank in the first period *t=1* as there is no prior rank $\beta_{Rank\ 1} = 0$. We will hence be estimating $\beta_{Rank\ 2}$, the effect of primary school rank on period 2 outcomes. To simplify the notation the time subscript will be dropped, as only one rank parameter is estimated, β_{Rank}.

This specification also allows for the unobservables to have time varying effects. Again τ_i represents unobserved individual effects that capture all time constant effects of a student over time and μ_{jsc} represents the permanent effects of being taught in a specific school-subject-cohort. Now additionally we have τ_{it} and μ_{jsct} allowing for these error components to vary over time so that students can have individual-specific growth rates as they grow older, or that primary school teachers can affect the efficiency of their students to learn a certain subject in the future.

Given this structure we now state explicitly the conditional impendence assumption that needs to be satisfied for estimating an unbiased rank parameter. Conditional on student characteristics, time varying and permanent primary school-subject-cohort level and individual effects, we assume there would be no expected differences in students' outcomes except those driven by rank.

$$Y_{Ri} \perp R_i \mid X_i, \mu_{jsc}, \mu_{jsct}, \tau_i, \tau_{it} \text{ for all R} \tag{3}$$

To achieve this we require measures of all these factors that may be correlated with rank and final outcomes. Conditioning on prior test scores will absorb all non-time varying effects as they will effect period-1 test scores to the same extent as period-2 test scores. Any input, observable or unobservable, that would affect academic attainment is captured in these test scores.[8] Therefore we can express period two outcomes, age 14 test scores, as a function of rank, prior test scores, student characteristics and unobservable effects.

$$\begin{aligned} Y_{ijksc2} = {} & \beta_{Rank} R_{ijsc} + f\left(Y_{ijsc1}\left(X_i{}'\beta, X_i{}'\beta_1, \tau_i, \mu_{jsc}, \tau_{i1}, \mu_{jsc1}\right)\right) + \\ & + X_i{}'\beta_2 + \mu_{jsc2} + \tau_{i2} + \epsilon_{ijksc} \end{aligned} \tag{4}$$

Using lagged test scores means that the remaining factors are those that affect the learning in period 2, between ages 11 and 14 ($X_i{}'\beta_t, \mu_{jsc2}, \tau_{i2}$). In our regressions, we

[8] Examples of these effects include students' innate ability, parental investment, teacher effects, peer effects and primary school infrastructure

will allow the functional form of this lagged dependent variable to take two forms, either a 3rd degree polynomial or a fully flexible measure, which allows for a different effect at each national test score percentile. As we can observe certain characteristics and primary school attended, β_2 and μ_{jsc2} can easily be estimated. The interpretation of μ_{jsc2} is that some primary schools are more effective at teaching for a later test than others, in a way that does not show up in the end-of-primary age-11 test scores.

The discussion of recovering τ_{i2}, the second period academic growth of individual i is below, but it is worth spending some time interpreting what the rank coefficient represents without its inclusion. Being ranked highly in primary school may have positive spillover effect in other subjects. Any estimation, which allows for individual growth rates during secondary school (second period), would absorb any spillover effects. Therefore, leaving τ_{i2} in the residual means that the rank parameter is the effect of rank of the subject in question and the correlation in rank from the other two subjects, as we have test scores for English, mathematics and science.

In the second period the student will be attending secondary school k which may affect later test scores by subject, π_{ksc}, which is a component of the error term ϵ_{ijksc}, where $\epsilon_{ijksc} = \pi_{ksc} + \varepsilon_{ijksc}$. As stated above conditional on time-varying student effects, prior subject test scores and the other stated factors, we do not expect that these components will be correlated with primary rank. This is primarily because general secondary school effects are absorbed by the time varying student effect but we will return to the issue of secondary school choice in subsection 3.3.1.

The first two specifications that we estimate that will recover the effect of rank due to overall changes in effort which allow for spill-overs between subjects, are the following:

$$Y_{ijksc2} = \beta_{Rank} R_{ijsc} + f\left(Y_{ijsc1}\right) + X_i{}'\beta_2 + \mu_{jsc2} + \epsilon_{ijksc} \qquad (5)$$

$$\text{Where } \epsilon_{ijksc} = \tau_{i2} + \pi_{ksc} + \varepsilon_{ijksc}$$

$$Y_{ijksc2} = \beta_{Rank} R_{ijsc} + f\left(Y_{ijsc1}\right) + X_i{}'\beta_2 + \mu_{jsc2} + \pi_{ksc} + \psi_{ijksc} \qquad (6)$$

$$\text{Where } \psi_{ijksc} = \tau_{i2} + \varepsilon_{ijksc}$$

Note that we can further augment these regressions by using the average student growth rate across subjects to recover individual growth effects, τ_{i2}. Note that despite using panel data, this is estimating the individual effect across subjects and not over time. Lavy (*et al.* 2012) also use a student-fixed effects strategy to estimate ability peer effects. Applied to this setting, when allowing for student effects, we effectively compare relative rankings within an individual, controlling for national subject-specific ability. The variation used to identify rank is correlation between the differential growth rates by subject within each student and prior subject ranking. Therefore any individual characteristic that is not realised in age-11 test scores but contributes towards age-14 test scores is accounted for, including secondary school attended, as long as the effects are not subject specific.

$$Y_{ijksc2} = \alpha_t + \beta_{Rank}R_{ijsc} + f\left(Y_{ijsc1}\right) + \tau_{i2} + \mu_{jsc2} + \epsilon_{ijksc} \qquad (7)$$

$$\text{Where } \epsilon_{ijksc} = \pi_{ksc} + \varepsilon_{ijksc}$$

In these specifications the rank parameter only represents the increase in test scores due to subject specific rank, as any general gains across all subjects would be absorbed by the student effect. This can be interpreted as the extent of specialisation in subject *s* due to primary school rank. It is for this reason, and the removal of other covarying factors, that we would expect the coefficient of the rank effect in specification (7) to be smaller than those found in (5) or (6).

Finally, to also investigate potential non-linearites in the effect of ordinal rank on later outcomes, i.e. are effects driven by students being top or bottom of the class, we replace the linear ranking parameter with indicator variables according to quantiles in rank plus additional indicator variables for those at the top and bottom of each school-subject-cohort (the rank measure is defined in section 4.2). We allow for non-linear effects according to vingtiles in rank, which can be applied to all the specifications presented.[9]

$$Y_{ijksc2} = \beta_{R=0}Bottom_{ijsk} + \sum_{n=1}^{20} I_n R_{ijsc}\,\beta_{n,Rank} + \beta_{R=1}Top_{ijsc}$$
$$+ f\left(Y_{ijsc1}\right) + \tau_i + \mu_{jsc} + \epsilon_{ijksc} \qquad (8)$$

In summary, if students react to ordinal information as well as cardinal informaiton, then we would expect the rank in adition to relative achivement to have a signficant effect on later achievement when estimating these equations. This is what is picked up by the β_{Rank} -coefficient. The following sections discuss potential threats to identification, the setting, and how rank is measured before we turn to the estimates.

3.3 Threats to identification

3.3.1 Secondary school selection

A concern may remain that students could select secondary schools based on their rank in a particular subject during primary school in addition to their age-11 test scores. If, for example, students who were top of their class in maths aspire to attend a

[9] Estimates are robust to using deciles in rank rather than vigntiles and can be obtained upon request.

secondary school that specialises in maths, our estimates could be confounded by secondary school quality. This might seem unlikely because we know that ability sorting for secondary schools in England is largely based on average rather than subject-specific abilities (Lavy *et al.* 2012).

Fortunately, our data allows us to address this concern directly by additionally controlling for secondary school attended. The resulting specification additionally allows for period-2 achievement to vary by secondary school k in subject s of cohort c, π_{ksc2}.[10] Intuitively, this is comparing students who are exposed to the same secondary school influences, thus identifying effects net of any potential subject-rank driven sorting into secondary education. However, secondary school attended can be argued to be an outcome, and therefore should not be conditioned upon. Specifications that include these effects are not our preferred model and should only be used as an indication of the extent that secondary school selection has effects on the estimates. As we will see, this modification does not affect our results.

3.3.2 Unobserved individual factors and parental background

Even with this flexible specification, one may still not be convinced that we are identifying the effect of rank on subsequent educational attainment. The rank of a student in primary school may be correlated with other unobserved individual factors that affect students' outcomes. An example of this could be unobserved individual or parental aspirations that correlate with primary rank and later value added, i.e. a competitive child or 'pushy parents'. Furthermore, using across-school variation might be problematic if schools transformed a student's ability into test scores non-monotonically.[11] We believe that the student fixed effects approach outlined by specification (7) addresses most of these concerns as all unobserved factors that affect test scores in all subjects in a similar way are now controlled for.

Notice that while these general factors such as 'pushy parents' that could induce correlations between primary rank and later outcomes are now controlled for, the remaining assumption required for identification is that such unobserved factors are not subject-specific. We return to this issue in Section 6.4 where we show that parental occupations predict primary-subject test scores yet are orthogonal to primary-subject rankings.

[10] We use the Stata command `reg2hdfe` for these estimations (Guimaraes and Portugal, 2010).
[11] If some schools are better at teaching low (high) ability students, then the ranking technology for ability may be different across schools.

3.3.3 Ability peer effects and measurement error

Notice that all of our estimation specifications include primary-subject-cohort effects, which is also necessary to account for potential measurement error in the age-11 test scores arising through unobserved classroom-level shocks. In particular, if there are unobserved primary-school factors, these will create noise in the test score but not in the rank, as the ranking itself is mean-independent. As a result, the ranking variable could start to pick up ability-related information that could not be fully controlled for using only the national percentile test score. Including primary-school effects clears this kind of measurement error from the primary rank variable. We will return to these issues in Section 6, where we also examine in detail how these rank effects interplay with ability peer effects, as well as potential threats placed by various kinds of measurement errors. We will conclude that whilst the most obvious candidates, i.e. classroom-level shocks and ability peer effects are controlled for directly in our setting, that higher order issues of measurement error and transitory non-linear peer effects do not invalidate our approach.

4 Institutional setting, data and descriptive statistics

4.1 The English School System

The compulsory education system of England is made up of four Key Stages (KS); at the end of each stage students are evaluated in national exams. Key Stage Two (KS2) is taught during primary school between the ages of 7 and 11. The median size of a primary school cohort and the average primary school class size is 27 students (DFE, 2011). Therefore, when referring to primary school rank, one could consider this as class rank.[12] At the end of the final year of primary school when the students are aged 11 (Year 6), they take KS2 tests in English, math and science. These tests are externally graded on a national scale of between 0-100. This makes it possible to make comparisons in student achievement over time and across schools.

Rather than receiving these raw scores, students are instead given one of five broad attainment levels. The lowest performing students are awarded Level 1, the top performing students are awarded Level 5. These levels are broad, which results in them being a coarse measure, with 85% of students achieve Level 4 or 5. These are non-high

[12] The maximum class size at Key Stage 1 is 30 students. A parallel set of results has been estimated using only cohorts of 30 and below, assuming these are single class cohorts. The results are qualitatively the same and are available from the authors upon request.

stakes exams for students and are mainly used by the government as a measure of school effectiveness.[13] This means that students do not know their underlying exact test score, which we can use to calculate their local ranks. Rather, students infer their rank position in class through repeated interaction, teacher feedback, and often through seating arrangements that reflect ability.

Students then transfer to secondary school, where they start working towards the third Key Stage (KS3). During this transition the average primary school sends students to six different secondary schools and the larger secondary schools receive students from 16 different primary schools. Importantly, admission into secondary schools is generally non-selective and does not depend on end-of-primary KS2 test scores. A subset of schools can select on ability (grammar schools) but these schools administer their own admission tests. The KS2 is thus a low-stakes test with respect to secondary school choice. In the new school, a typical student has 87% new peers upon arrival. This large re-mixing of peers is beneficial, as it allows us to estimate the impact of rank form a previous peer group on subsequent outcomes. Key Stage 3 takes place over three years and at the end of Year 9, all students take KS3 examinations in English, math and science at age fourteen. Again KS3 is not a high-stakes test and is externally marked.

Two years later, students take the national Key Stage 4 test at age 16 (KS4), which marks the end of compulsory education in England. The KS4 is graded from one to eight and students have some discretion in choosing the subjects they study and at what level. Since KS3 is graded on a fine scale [0-100], and students are tested in the same compulsory subjects only, we prefer this as the outcome measure for the purpose of our study. However, our results also hold using KS4 test scores[14].

4.2 Data Construction

4.2.1 Administrative data

The Department for Education (DfE) collects data on all students and all schools in state education in England. The Pupil Level Annual School Census (PLASC) collects student information such as gender, ethnicity, language skills, Special Educational Needs (SEN),

[13] The students also appear not to gain academically just from achieving a higher level. Regression discontinuity techniques show no gain for those students who just achieved a higher level. This setting is ideal for a regression discontinuity techniques as the score needed to reach a level changes by year and by subject, which would make it particularly hard to game.

[14] Results can be obtained from the authors upon request. They are not presented here due to issues relating to the comparability of these test scores across students as they can be entered into different exams, along with the coarseness of the measures and students choosing to study additional optional subjects.

or being Free School Meals Eligible (FSME). The number of students and student characteristics are used to determine school funding. The National Pupil Database (NPD) contains student attainment data throughout their Key Stage progression in each of the three compulsory subjects. Each student is given a unique identifier so that they can be linked to schools and followed over time, allowing the government to produce value added measures and publish school league tables. As the functions of both of these datasets are at the school level, no class level data is collected.

We have combined these data to create a dataset following the entire population of five cohorts of English school children. This begins at the age of 10/11 (Year 6) in the final year of Primary School when students take their Key Stage 2 examinations through to age 13/14 (Year 9) when they take Key Stage 3 tests. KS2 examinations were taken in the academic years 2000/2001 to 2005/2006 and so it follows that the KS3 examinations took place in 2003/2004 to 2007/8. From 2009 students were no longer externally assessed, instead teacher assessment was used to evaluate students at the end of Key Stage 3, hence this is the end point of our analysis.

We imposed a set of restrictions on the data to obtain a balanced panel of students. We use only students who can be tracked with valid KS2 and KS3 exam information and background characteristics, 83% of the population. Secondly, we exclude students who appear to be double counted (1,060) and whose school identifiers do not match (12,900), approximately 0.6% of the remaining sample. Finally, we remove all students who attended a primary school whose cohort size was smaller than 10, as these small schools are likely to be atypical in a number of dimensions; this represents 2.8% of students[15]. This leaves us with approximately 454,000 students per cohort, with a final sample of just under 2.3 million student observations, or 6.8 million student-subject observations.

As described in Section 4, the Key Stage test scores for both levels are percentalized by subject and cohort, so that each individual has six test scores between 0 and 100 (three KS2 and three KS3). This ensures that students of the same nationally relative ability have the same national percentile rank, as a given test score could represent a different ability in different years or subjects. Importantly, this allows for test score comparisons to be made across subjects and across time, this does not impinge on our

[15] Estimations using the whole sample are very similar, only varying at the second decimal point. Contact authors for further results.

estimation strategy, which relies only on heterogeneous test score distributions across schools to generate variation in local rank.[16]

We rank students in each subject according to their age 11 national test scores within their primary school by cohort. To have a comparable local rank measurement across schools of different cohort size we transform the rank position of individual i with the following normalization:

$$R_{ijsc} = \frac{n_{ijsc}-1}{N_{jsc}-1}, \qquad R_{ijsc} = [0,1] \tag{9}$$

where N_{jsc} is the cohort size of school j in cohort c of subject s. An individual's i ordinal rank position within this set is n_{ijsc}, which is increasing in test score. R_{ijsc} is the standardised rank of the student.[17] For example, a student who had the second best score from a cohort of twenty-one students (n_{ijsc}=20, N_{jsc}=21) will have R_{ijsc}=0.95. This rank measure will be approximately uniformly distributed, and bounded between 0 and 1, with the lowest rank student in each school cohort having R=0. In the case of draws of national percentile rank, each of the students is given the lower local rank.

Rank is dependent on students own test scores, but is determined by the scores of others in their school. Again consider the students who scored X and Y in cohorts with different test score distributions from Figure 1. The students who scored Y, being the same distance above the mean in both school cohorts would have a rank of R_{yA}= 0.9 in Cohort A (unimodal distribution) and R_{yB}= 0.6 in Cohort B (bimodal distribution). Similarly students who scored X would have a rank of R_{xA}= 0.1 in Cohort A and R_{xB}= 0.4 in Cohort B. It is the different distribution of peer test scores that allows for the separate identification of the rank effect from a relative ability effect. As there is information for three subjects for each student, a student can have a different rank for each subject within her primary school. This feature of the data allows us to include student fixed effects in some of our regressions.

Note that since the students do not receive their detailed test scores, they will not be able to derive this same rank score themselves, nor will they be given an official

[16] Estimations using standardised rather than percentalized test scores provide similar estimates to the first decimal place in linear specification. For non-linear specifications the effect of rank appears more cubic in nature. However, these estimations suffer from non-comparability as a given test score could represent a different ability in different years or subjects. Year/subject effects would not account for all these differences as there are likely to be distributional differences. Allowing for either functional form of test scores to be interacted by year and subject was extremely computationally intensive, given our already demanding specification. Basic results are available from the authors upon request.

[17] This is rank within school subject cohort, it cannot be done by class as no class level information is available. However, all estimations have been replicated on schools which have cohort sizes of under 30 (maximum class size) and have equivalent results. Obtainable upon request.

ranking. Instead, our measure of local rank is a proxy for the students' experiences over the past six years of interacting with their peers in the classroom. The existence of any effect is driven through student beliefs about their rank position within their class.

4.2.2 Survey data

Additional information about a subsample of students is obtained through a representative survey of 16,122 students from the first cohort. The Longitudinal Survey of Young People in England (LSYPE) is managed by the Department for Education and follows a cohort of young people, collecting detailed information on their parental background, academic achievements, out of school activities and attitudes.

We merge survey responses with our administrative data using a unique student identifier. This results in a dataset where we can track students from a primary school, determine their academic ranks and then observe their later measurements of confidence and attainment, allowing us to test if rank effects confidence conditional on attainment. This is the first research to merge LSYPE responses to the NPD for primary school information.

At age 14 the students are asked how good they consider themselves to be in English, maths and science, with 5 possible responses that we code in the following way; 2 'Very Good'; 1 'Fairly Good'; 0 'Don't Know'; -1 'Not Very Good'; -2 'Not Good At All'. We use this simple scale as a measure of subject specific self-concept. Whilst it is much more basic than surveys that focus on self-concept, it does capture the essence of the concept.

The matching between the NPD and LSYPE was perfect. However, the LSYPE also surveys students attending private schools that are not included in the national datasets; moreover, as students not accurately tracked over time have been removed, a further 3,731 survey responses could not match. Finally, 1,017 state school students did not fully complete these questions and so could not be used for the self-concept analysis. Our final dataset contains 11,898 student observations with self-concept measures. Even though the survey will not contain the attitude measures of every student in a school cohort, by matching the main data we will know where that student was ranked. This means we will be able to determine the correlation of rank on self-concept, conditional on age 11 test scores and school-subject-cohort fixed effects.

4.3 Descriptive statistics

4.3.1 Main sample

The data has the complete coverage of the state student population from age 10 to 14. We follow each student from their primary school through to secondary school, linking their rank in their school to their later outcomes. Table 1 shows summary statistics for all students that are used the analysis. Given that the test scores are represented in percentiles, all three subjects test scores at age 11 and 14 have a mean of 50 with a standard deviation of 28.

The within-student standard deviation across the three subjects English, math and science is 12.68 national percentile points at age 11 with similar variation in the age 14 tests. This is important as it shows that there is variation within student which is used in student effects regressions.

Information relating to the background characteristics of the students is shown in the lower panel of Table 1 half the student population is male, over four-fifths are white British and about 15 per cent are Free School Meal Eligible (FSME) a standard measure of low parental income.

Figure 2 shows the position of top- and bottom-ranked students, as defined by being in the top or bottom 5% within each school subject cohort, against their national percentile rank. We see the large variation across schools in the test scores that would make a student rank in the top or bottom. Whilst in the majority of schools students would need to score around the 90th percentile nationally to be a ‘top student’, in some schools a student need only be in the 50th percentile.

We use this variation of test scores across schools to identify the effect of rank separately from relative ability. This was previously illustrated for a theoretical case in Figure 1, which shows the rank of an individual is dependent on the distribution of test scores even when maximum, minimum and mean test scores are the same in both schools. In the top panel of Figure 3 we replicate this with actual student test score data from six primary schools. Each point represents a student’s age 11 English test score. Each row represents a school which has a student ranked in the 1st and 100th national percentiles, has a mean percentile of 54 and a student in the 93rd percentile in English. This is a very specific case, but in each the student at the 93rd percentile has a different rank. For the estimations, we use all subjects and the distributions of test scores across all primary schools whilst accounting for mean school-subject-cohort test scores. Therefore the lower panel of Figure 3 plots the rank of every student in each subject by de-meaned test score. The vertical thickness of the distribution of points indicates the support at throughout the rank distribution. For the mean students there is nearly full rank support.

That there are differences in test score distributions across schools will be the result of many factors. One example is that a school in a rural area where there is little school choice may have a wider test score distribution than a school in a city where there is more parental sorting. However, conditional on school-subject-cohort and student effects, we are confident that these factors will not bias our results.

4.3.2 Longitudinal Study of Young People in England

Appendix Table 3 shows descriptive statistics for the LSYPE sample which we use to estimate rank effects on a direct measure of self-concept. The LSYPE respondents are representative of the main sample, although mean age 11 test scores are slightly lower and the proportion of Free School Meal Eligible is higher than the national at 18.6% and 14.6% respectively (Appendix Table 3).

The LSYPE students are asked to rate themselves in each of the subjects from 'Not good at all' to 'Very Good,' which is summarized in Appendix Table 5. Our measure of self-concept is coarse, with only five categories to choose from and around 60% choosing 'fairly good'. We can see that students do think about their own ability, with less than 0.2% not having an opinion. As would be expected, those who considered themselves to be poor performers did tend to have lower average national KS2 percentile rank and lower rank within their school. However, there is also large variance in these ranks within these self-evaluated categories. For every subject, each self-assessment category with an opinion has at least one individual in the top 9% nationally, including those who considered themselves 'Not Good'. Similarly, each category has an individual in the lowest performing percentile nationally, even those who consider themselves very good.[18]

[18] In Appendix Table 5 we also show the performance of the top and the bottom 10% of students within each self-assessment category that are less affected by outliers. We continue to see very large variance within categories. Consider Science in Panel C: of those who consider themselves 'Very Good' the bottom 10% performers in this category are ranked at the 17 percentile point nationally, whereas the top 10% of performers in the category that rated themselves 'Not very good at all' ranked at 64^{th} percentile nationally.

5 Main Results

5.1 *Effect of Rank: comparing across school cohorts*

To begin the discussion of the results we present estimates of the impact of primary school rank on age 14 test scores. The estimates are reported in Panel A of Table 2, with the specifications becoming increasingly flexible moving across columns to the right. The first row shows estimates of the rank parameter using fully flexible set of controls for age 11 test scores, allowing each percentile score to have a different effect on later test scores. Due to computational constraints we are unable to run all specifications using this functional form and therefore the second row instead uses a third order polynomial of age 11 test scores. It appears that this is sufficient approximation to account for the effect of age 11 test scores. All estimates control for a set of student characteristics and have standard errors clustered at the secondary school level[19].

The first column is a basic specification, which only controls for age 11 test scores, student characteristics, along with cohort and subject fixed effects. This shows a comparatively large estimate: a student at the top of their cohort has an 11.6 larger national percentile rank gain in test scores compared to a student ranked at the bottom, *ceteris paribus*. However, this regression does not condition on school-subject-cohort effects and therefore the parameter cannot be interpreted as pure rank effect as it will also capture the effects of relative ability. Furthermore, it uses variation in average quality of students across schools, which might correlate to family background characteristics, later school quality, or other unobserved variables.

Indeed, this is what we find in column (2) which is significantly smaller and additionally allows for any primary school-subject-cohort effects (Specification 5). Using this specification, the effect of being ranked top compared to bottom *ceteris paribus* is associated with a gain 7.96 national percentile ranks (0.29 standard deviations) conditional on a cubic of age 11 test scores. This can be interpreted as the additional ordinal rank effect. Given the distribution of test scores across schools, very few students would be bottom ranked at one school and top at another school. A more useful metric is to describe the effect size in terms of standard deviations, a one standard deviation increase in rank is associated with increases in later test scores by 0.085 standard deviations or 2.36 national percentile points. Note that any determinant that has a permanent effect on student outcomes would be absorbed by prior test scores, this is

[19] Student characteristics are ethnicity, gender, ever Free School Meal Eligible (FSME) and Special Educational Needs (SEN)

the growth in national percentiles between the ages of 11 and 14 due to primary school rank.[20] In comparison with other student characteristics, females' growth rate is 1.01 national percentile points higher than males and free school meal eligible students on average lose 2.96 national percentile points (Appendix Table 6).[21]

We see that when additionally allowing for secondary school-subject-cohort effects (Specification 6) there is only a marginal impact on the estimates and are not significantly different from those in column 2. This is evidence that that there is negligible sorting into secondary schools by subject rank, conditional on student test scores. Given that secondary school attended can be argued to be an outcome, these effects will not be included in the within student analysis.

5.2 *Effect of Rank: within student analysis*

We now turn to estimates that use the within student variation to estimate the rank effect (Specification 7). Conditioning on student effects allows for individual growth rates, which absorb any student level characteristic. Since students attend the same primary and secondary school for all subjects, any general school quality or school sorting is also accounted for. Subject specific primary school quality is absorbed by the primary school-subject-cohort effects. This uses the variation in the relative growth rates across subjects within student according to differing rank in primary school.

Besides removing potential biases, the inclusion of student effects changes the interpretation of the rank parameter. The student effect will also absorb any spillover effects gained through high ranks in other subjects and is only identifying the relative gains in that subject. Accordingly the within student estimate is considerably smaller. The effect from moving to the bottom to top of class *ceteris paribus* increases national percentile rank by 4.56 percentiles, as we see in Panel A, column (4) of Table 2. To make a comparison in terms of standard deviations this effect is scaled by the within student standard deviation of national percentile rank (11.32). Therefore, conditional on student and school-subject-cohort effects, the maximum effect of rank is 0.40 standard deviations. This is a very large effect, but a change from last to best rank *within student* represents an extreme treatment. It is more conceivable for a student to move 0.5 rank

[20] Using teacher assessment data on student ability to rank students near start of primary school, at age 7. We find that students who are consistently top of their primary school do additionally better in age 14 test. Results available upon request, not presented as main result due to coarseness and reliability of age 7 test scores.

[21] Including the rank parameter in this specification reduces the Mean Square Error by 0.31. This is more than the reduction from allowing for a gender growth term (0.25) or an ethnicity growth term (0.28).

points, e.g. being at the 25th percentile in one subject and 75th at another. Our estimates imply that this student would improve their test scores in that subject by 0.20 standard deviations. In terms of effect size, given that a standard deviation of the rank within student is 0.138 for any one-standard deviation increase in rank, test scores increase by about 0.056 standard deviations.[22]

Again, if there were any general gains through achieving a high rank in one subject, this would be absorbed in the within student estimates, and thus could be interpreted as the between subject substitutions of effort allocation, or a lower bound of the effect of being highly ranked. The difference between the within school estimates (7.96) and the within student estimates (4.56) can be interpreted as an upper bound of the gains from spilovers between subjects. A more detailed interpretation of the differences in effect size are provided in Section 7.5 once a mechanisms has been established, and Appendix 3 which describes a basic model for this mechanism.

5.3 Non-linear Effects

The specifications thus far assumed the effect of rank is linear, however, it is conceivable that the effects of rank change throughout the rank distribution (Brown, 2011). To address this we allow for non-linear effects of rank by replacing the rank parameter with a series of 20 indicator variables according to the vingtiles in rank, plus top and bottom of class dummies, as outlined in specification (8).

The equivalent estimates from specification (5) and (7), i.e. without and with student fixed effects, are presented in Figure 4. The effect of rank appears to be almost linear throughout the rank distribution, with small flicks in the tails. Reassuringly, the placebo ranks (to be discussed in Section 6.3) are also insignificant when allowing for non-linear effects. In comparison, all rank coefficients are significantly different from the reference group of the median-ranked students (10th vingtile). This indicates that the effect of rank exists throughout; even those students ranked just above the median perform better three years later than those at the median. Given that students are not informed of rank, our interpretation of this is that students are good at ranking themselves within the classroom. This ranking developed through the constant exposure to peers over the length of primary school, which continually reinforces the effect on self-concept such that by the end of primary school they have strong beliefs about where they rank.

[22] For students with similar ranks across subjects the choice of specialization could be less clear. Indeed, in a sample of the bottom quartile of students in terms of rank differences, the estimated rank effect is 25% smaller than those from the top quartile. Detailed results available on request.

Finally, the fact that the rank effect exists throughout the distribution is in line with the idea that self-concept forms according to rank position.

5.4 Heterogeneity by gender and parental income

We now turn to how the effects of rank vary by student characteristics using the student fixed effects specification (7) with non-linear rank effects and interacting the rank variable with the dichotomous characteristic of interest.[23] The student characteristics are Male: Female and, FSME: Non-FSME. The baseline group coefficients and the interaction plus baseline coefficients are plotted to show the effect of rank on test scores for both groups, illustrating how the different groups react to primary school rank.[24]

The first panel in Figure 5 shows the how rank relates to the gains in later test scores by gender. Males are more affected by rank throughout 95% of the rank distribution, this is shown by the steeper gradient of the rank effect. Males gain four times more from being at the top of the class, but also lose out marginally more from being in the bottom half. As this is within a student variation in later test scores, the coefficient could be interpreted as a specialising term, implying that prior rank has a stronger specialising effect on males than females.

The second panel in Figure 5 shows that Free School Meal Eligible (FSME) students are less negatively affected by rank and more positively affected than Non-FSME students. FSME students with a high rank gain more than Non-FSME students, especially those ranked top in class, who gain almost twice as much. FSME students who are below the median have limited negative effects on later test scores. This could be interpreted as these students already having a low self-concept for other reasons and therefore the negative effects of low rank have less of an effect. Moreover the shallower gradient for Non-FSME students could also be interpreted that they are less affected by class rank as these students may have their academic self-concept being more be affected by factors outside of school.

[23] Interacting student characteristics rather than estimating the effects separately, ensures that students who attend the same school have the same relative. Use of interactions is preferred over separate regressions as the school-subject-cohort effects will be shared across groups and so relative test scores according to that school's mean will be the same for both.

[24] The student characteristics themselves are not included in the estimations, as they are absorbed by the student effects. These characteristics interacted by rank, however, are not absorbed by student effects, because there is variation within the student due to having different ranks in each subject.

6 Robustness

Some non-trivial empirical challenges arise when estimating the effect of rank conditional on test score because we do not independently observe both a students' rank and a student's ability. Instead, we rely on externally marked and nationally standardised tests at the end of primary school to derive a student's local rank during primary education and also use this measure to control for a student's subject-specific achievement. This may cause problems relating to the influence of peers, parents and measurement error on test scores.

6.1 Peer Effects

Firstly, given that we are discussing an atypical peer effect, it is important to address the issues associated with such.[25] Any primary school peer effects that have a permanent effect on test scores do not bias the estimates as they are captured in the end-of-primary school test scores. Furthermore, we can account for contemporaneous secondary peer quality with the inclusion of secondary school-subject-cohort effects.[26]

However, if peer effects have a transitory effect on test scores i.e. only current peers matter, any estimation of the effect of primary rank on age 14 test scores whilst controlling for primary test scores could be biased. This is to the extent that both the conditioning variable and rank will be correlated with primary peer effects. The intuition for this is as follows: in the presence of transitory peer effects, a student with lower quality peers would attain a lower primary school results than otherwise and also have a higher rank than otherwise. Thus, when controlling for primary test scores in the estimations, those who previously had low quality peers would appear to gain more as they now have a new peer group, who on average would be better. Since rank is negatively correlated with peer quality in primary test scores, it would appear that those with high rank make the most gains. Therefore having a measure of ability confounded by transitory peer effects would lead to an upward biased rank coefficient.

This is shown to be the case in Appendix 1, where we create a data generating process in which we specify that subsequent test scores are not effected by rank. Instead test scores are only a function of ability and individual linear or non-linear peer effects. To be cautious we allow for these peer effects to be 20 times larger than those found in

[25] The standard reflection problem is not a first order issue in this situation, as students are surrounded by 87% new peers when they transfer to secondary school, and the rank effect is generated by primary school peers.
[26] This has almost zero effects on our coefficients partly because of the large re-mixing of students during the primary-to-secondary transition, and the sorting to school not being rank dependent.

Lavy et al. (2012). We simulate these data 1000 times and estimate the rank parameter with different sets of controls. This shows that not controlling for the primary school peer group generates biased results, but that this bias is negligible when allowing for mean school-subject-cohort effects, even with these large non-linear peer effects. These simulations and further discussion can be found in Appendix 1 and Appendix Table 1.

6.2 Measurement Error

In addition to peer effects, individual test scores may be imperfect measures of inputs up until that point in time. Given that both rank and test scores will be affected by the same measurement error, but to different extents due the heterogeneous test score distributions, calculating the size of the bias is intractable. To gauge the extent of measurement error we again simulate the data assuming 20% of the variation in test scores is random noise, 70% student ability and 10% school effects, these proportions reflect that 80% of the variance of test scores is within schools and 20% across schools (Appendix 2). This shows that normally distributed individual-specific measurement error would work against finding any effects.

The intuition for this is the following: a particular student having a large positive measurement error would result in both an inflated end-of-primary score and a higher local rank measure. Both of these would work against finding positive effect of rank on later outcomes, as we explicitly control for prior attainment. This student’s later test scores would hence be benchmarked against other students' with the same end of primary result but higher actual ability. Since the student only got a high local rank because of the measurement error, this would downward bias any positive rank effect estimate.

6.3 Is rank just picking up ability?

Our estimates of primary school subject-specific rank are relatively large, given that we are conditioning prior test scores and individual growth. As rank is highly correlated with student ability and test scores, there could be a concern that measurement error in the test scores for ability may be recovered in the rank measurement, if rank is measured with less error than test scores.

Note that this is different from the measurement error concern discussed above. To address the specific measurement error problem of rank having less measurement error than test scores and thus containing residual ability information, we perform placebo tests. This involves generating a placebo-rank measure that uses underlying ability, but would not reflect the social comparison experiences of students. To achieve this we re-assigned randomly all students into primary schools by cohort and re-calculated the ranks that they would have had in these schools with their original age-11 test scores but with peers that they never actually interacted with. These placebo-ranks are highly correlated with age-11 test scores. If they were found to be significant determinants of later achievement, this would indicate that rank is picking up ability not captured in end of primary school outcomes. We re-estimate all the specifications fifty times using new placebo-ranks each time and present the mean results in Panel B of Table 2, and the non-linear effects in Figure 4. We find no effects of these placebo ranks on later test scores. From these simulation results we conclude that our findings are unlikely to be mechanically driven by measurement error in test scores.

6.4 *Are student effects enough? Primary school sorting and parental occupation*

The causal interpretation that we give to estimates relies on the conditional independence assumption. That a student's rank needs to be orthogonal to other subject-varying determinants of a student's later achievement. Given the student effects, the variation need not be orthogonal to general determinants of the student's achievement, but would need to vary within a student across subjects. A prime example of this could be the occupational background of the parents. Children of scientists may have a higher learning curve in science throughout their academic career for reasons of parental investment or inherited ability. Similarly children of journalists for English and children of accountants in maths. This will not bias our results as long as conditional on age 11 test scores parental occupation is orthogonal to primary school rank. Or more broadly, there would be a problem if conditional on other factors, rank was correlated to subject-varying determinates of future achievement. This might well be the case if parents strongly aspire for their child to rank top in that subject and also have a higher academic growth rate in that subject between the ages 11 and 14.

Typically parents are trying to get their child into the 'best school' possible in terms of average grades. This would work against any positive sorting by rank as higher average achievement would decrease the probability of their child having a high rank. This sorting on general achievement would be accounted for by the student fixed effect. However, if parents wanted to maximise their child's rank in a particular subject, this could bias the results. In order to do this they would need to know the ability of their child and all potential peers by subject. This is unlikely to be the case, particularly for such young children who have yet to enter formal education at age 4. Parents could possibly infer the likely distributions of peer ability if there is autocorrelation of the student achievement within a primary school. This means that if parents did know the ability of their child by subject, and the achievement distributions of primary schools they could potentially select a school on this basis.

We test for this by using the LSYPE sample which has information on parental occupation. All parental occupations are classified into English, math, science, or 'other' and then an indicator variable is created for each student-subject if they have a parent who works in that field [27]. This is taken as an indicator for the parents' subject preference. We then regress age-11 test scores on parental occupation, school-subject effects and student effects (Table 3, Panel A). This establishes that this measure of parental occupation is a significant predictor of student subject achievement even when allowing for individual effects. Then using rank as the dependent variable we test for a violation of the orthogonality condition in Panel B of Table 3. Here we see that whilst parental occupation does predict student achievement by subject, it does not predict rank conditional on test scores. This implies that parents are not selecting schools on the basis of rank for their child. We therefore do not reject that the orthogonality condition does not hold with respect to parental background. This does not rule out other co-varying factors that may bias the results but it provides us with confidence that this likely large factor does not.

7 Mechanisms

A number of different mechanisms could produce similar results; competitiveness; environmental favours certain ranks; external (parental) investment by task; students learn about their ability. In the following, we discuss how each coincides with the results presented so far.

7.1 Hypothesis 1: Competitiveness

If the goal of individuals was to be better than their peers, maximise rank, this could produce some of our results, but not the full pattern.

[27] Parental Standard Occupational Classification 2000 grouped in Science, Math, English and Other. **Science** (3.5%); 2.1 Science and technology, 2.2 Health Professionals, 2.3.2 Scientific researchers, 3.1 Science and Engineering Technicians. **Math** (3.1%); 2.4.2 Business And Statistical Professionals, 3.5.3 Business And Finance Associate Professionals. **English** (1.5%): 2.4.5.1 Librarians, 3.4.1 Artistic and Literary Occupations, 3.4.3 Media Associate Professionals. **Other**: Remaining responses.

To see this, consider two students of the same ability who attend the same secondary school but went to primary schools of different peer quality. The student attending the primary school of low quality peers could provide less effort in their end of primary school tests and still be ranked top. This student would then achieve lower end of primary school test scores than the student who faced competition in primary school. At secondary school when they have the same level of competition, and due to their same ability they will have the same expected age 14 test scores. In our estimation, controlling for prior test scores will make it appear that the student who faced lower competition and was ranked higher, had larger growth and thus generate the positive effect of rank.

However, if these mechanisms were driving the results, we would only expect to see these effects near the top of the rank distribution as it only applies to students who far exceed their peers and so get a lower than would be expected age-11 test scores. All those in the remainder of the distribution would be applying effort during primary school to gain a higher rank and so we should not see an effect. However given the result that the rank effect is approximately linear throughout it is unlikely that this type of competition mechanism is causing the effect.

It could still be the case that primary school subject rank is positively correlated with the degree of competitiveness of the student. Then those who are the most competitive increase their effort the most when entering secondary school and so have higher test score growth. Note that in the student effects specification any general competitiveness of an individual would be accounted for, this competiveness would need to vary by subject. As previously mentioned, any factor that varies by student across subjects conditional on prior test scores could confound –on in this case, explain- the results.

7.2 Hypothesis 2: The environment favours certain ranks

Another possible explanation for this finding is that the environment could favour the growth of certain ranks of agents. As an example, one can imagine primary school teachers teaching to the low ability students if faced with a heterogeneous class group[28]. If this were the case, teachers may design their classes with the needs of the lowest ranked students in mind. This means that these students would achieve higher age 11 test

[28] We have run estimations controlling for the within school-subject-cohort variance to take into account that high variance classes may be more difficult to teach. However, these cannot include school-subject-cohort or student effects, and thus the estimates should not be cleanly interpreted as ordinal rank affects. Therefore these specifications only allowed for general school effects or no school effects. The inclusion of a school-subject-cohort variance into these specifications does not significantly alter the rank parameter. Our findings can be presented upon request.

scores than they otherwise would have done and students further from the bottom lose out.

What would this mean for the rank effect estimates? Again consider two students of the same ability who attend the same secondary school but different primary schools, where one was top of year. The top student would get less attention during primary school and therefore get a lower grade than they otherwise would have done. At secondary school they have the same attention due to their same ability and get the same age 14 test scores. In our estimation, controlling for prior test scores will make it appear that the top student had higher growth and thus generate the positive effect of rank. Therefore, teachers teaching to the bottom student could also generate a positive rank effect. This would require primary school teachers only being effective with lowest ranked students and secondary school teachers teaching to each ability level equally. Note if primary teachers taught to the median student, those at both extremes would lose out. So instead of a linear effect, we would find a U-shaped curve with *both* students at the bottom and the top of the distribution gaining relatively more during secondary school.

If this was mainly due to the teacher focusing on those of low rank we would not necessarily expect to see large differences by gender, or free school meal status. We saw that males are more affected by rank than females, which would imply that males are more negatively affected by having the subject content not tailored to them e.g. top males under-achieve more during primary but catch up during secondary school. This is conceivable, however it runs counter to our estimate that males on average have lower growth in test scores between 11 and 14 (Appendix Table 6). Moreover, this does not also easily explain why free school meal students up to the middle of the class rankings are not negatively affected by the focus on the bottom, and those at the top of class are. Given these inconsistences, and that it relies on primary school teachers focusing solely on the lowest rank student and secondary school being tailored to student ability to generate similar effects, we doubt that his is the dominant reason for the effect.

7.3 Hypothesis 3: External (parental) investment by task

It may not be the students that are applying different effort by subject but that parents of the students are. Parents can assist the child at home with homework or other extra-curricular activities. If the parents know that their child is ranked highly in one subject, they might encourage the child to do more activities and be more specialised in this subject. Note that as we are controlling for student effects, this must be subject specific encouragement rather than general encouragement regarding schoolwork, and the additional investment must take place between ages 11 and 14. As we have already shown that conditional on test scores, parental occupation does not predict student rank, this hypothesis assumes parents react to achieved primary school rank rather than prior preferences.

However, we believe there are two further counter arguments for this mechanism. Firstly, whilst some parents may choose to specialise their child, others may want to improve their child's weakest subject. If parental investment focused on the weaker subject, this would reverse the rank effect for these students. To explain the positive rank effect, one would need to assume that the majority of parents wanted their child to specialise, which seems unrealistic for the ages eleven to fourteen. Secondly, parents are unlikely to be highly informed of their child's exact rank in class in the English context. Teacher feedback to parents will convey some information for the parents to act upon, such as the student being the best or worst in class, but may not be able to discern a difference from being near the middle of the cohort rankings. Our results however, show significantly different effects from the median for all vingtiles with school-subject-cohort effects[29].

7.4 Hypothesis 4: Students learn about their ability

Another possibility is that students use the information obtained by their local rank to learn about their subject-specific abilities, and as a result allocate effort accordingly. This is similar to the model proposed by Ertac (2006) where individuals do not know their own ability and therefore use their own absolute and relative performance to learn about it. Note that this mechanism does not change an individual's education production function, only their perception of it. We will argue below that this feature allows us to test the learning model, and fail to reject that the learning model has an effect. Thus we cannot provide evidence in favour of the learning model. .

Under the learning hypothesis students additionally use local rank information to make effort investment decisions across subjects, applying more effort according to where there is higher perceived ability. This would produce the same predictions by subject as a mechanism that changed the production function, however it has a different prediction for average grades. Students with larger differences between local and

[29] Information on the within student comparative advantage by subject would be easier for a teacher to communicate, and so parents could use this to specialize the student. However, these effects would then appear less significant in the school-subject-cohort effects specifications.

national ranks (in absolute terms) would have a more distorted information about their true abilities. These students would then have a higher misallocation of effort across subjects under the learning model, assuming diminishing returns to learning in each subject and that students want to allocate effort where they are most productive. Those with higher misallocation of effort would thus achieve lower overall grades, compared to students whose local ranks happen to closely align with national ranks. This is because this misallocation would lead to inefficient effort allocation across subjects and thus reduce average grades obtained. Whereas, if the rank effects were caused by actual changes in the education production function (and not just learning and changes in perceptions), even if local rank was different from national rank, this would not lead to a misallocation of effort in terms of maximising grades.

We do not have direct data on perceptions versus reality of costs, however we can test for misallocation of effort by examining how average grade achieved is correlated with misinformation. More precisely, we compute a measure of misinformation for students in each subject using their local rank $R_{ijsc} = [0,1]$ and national percentile rank $Y_{ijsc1} = [1,100]$ at age 11. Both are uniformly distributed and therefore we simply define misinformation Mis_{ijsc1} as the absolute difference between the two after rescaling percentile rank:

$$Mis_{ijsc1} = \left| R_{ijsc} - {Y_{ijsc1}}/{100} \right| , where\ Mis = [0,1] \tag{10}$$

This measure takes the value zero for students where their local rank happens to correspond exactly to the national rank. A large value, on the other hand, indicates large differences between local and national rank. Averaging this metric across subjects within student provides a mean indicator of misinformation for each student. To test directly if a student with a large amount of disinformation does significantly worse, we use a specification similar to (5) but with the by subject variation removed as we are examining the effect on average test scores. We estimate the following specification:

$$\bar{Y}_{ijc2} = \beta_{Rank}\bar{R}_{ijc} + f\left(\bar{Y}_{ijc1}\right) + X_i{'}\beta_2 + \varphi_{jc2} + \beta_{Mis}\overline{Mis}_{ijc1} + \epsilon_{ijksc} \tag{11}$$

$$\text{Where } \epsilon_{ijc} = \tau_{i2} + \varepsilon_{ijc}$$

where $\bar{Y}_{ijct}$ is average test scores across subjects in period *t*, $\bar{R}$ is average rank, φ_{jc} are primary school-cohort effects and $\overline{Mis}$ the additional misinformation variable. If the amount of misinformation caused them to misallocate effort over subjects we would expect $\beta_{Mis} < 0$, alternatively the null hypothesis local rank causes changes to the actual production function and $\beta_{Mis} = 0$.

$$H_1\text{: } Learning\ \beta_{Mis} < 0$$
$$H_0\text{: } Null\ \beta_{Mis} = 0$$

We obtain the following estimates using our full sample of 2,271,999 students. For benchmarking purposes, we first estimate a version of specification (11) without the additional misinformation variable (Table 4). The effect of average rank on average test score is estimated at 10.7 and highly statistically significant.30 Column (2) adds the coefficient for the effect of misinformation, which is estimated to be small and statistically insignificant whilst the rank parameter remains almost unchanged. Given this specification we fail to reject the null hypothesis that the amount of misinformation does not cause students to misallocated effort. We therefore conclude that the learning mechanism alone is unlikely to generate our results, though we fully acknowledge the limitations of this test, in particular that we cannot control for primary-cohort-subject or student fixed effects in this specification.

7.5 Hypothesis 5: Rank position develops self-concept

An alternative explanation is that when surrounded by people who perform a task worse (better) than oneself, one develops a positive (negative) self-concept in that area. The psychological-education literature uses the term self-concept, which is formed through our interactions with the environment and peers (O'Mara et al., 2006). Individuals can have positive or negative self-concept about different aspects of themselves.

Applied to our setting, we envisage that students with higher rank would develop positive academic self-concept. Self-concepts can be subject specific as well as for academic work generally, so that a student can consider themselves good at school but still bad at math (Marsh et al., 1988; Yeung et al., 2000). Valentine et al. (2004) found that students with a high self-concept would also develop positive non-cognitive skills such as confidence, resilience, and perseverance. There is also broad agreement in the psychological literature that academic self-concept is most malleable before age 11 (Tiedemann, 2000; Lefot et al., 2010; Rubie-Davis, 2011), which is when we measure rank. The importance of such non-cognitive skills for both academic attainment and non-academic attainment is now well established (Heckman and Rubinstein, 2001; Borghans et al., 2008; Lindqvist and Vestman, 2011).

[30] This is about three points lager compared to our previous estimates of Table 2 (column 2). Note, however, that this specification does not allow controlling for Primary-cohort-subject effects. Instead, only Primary-cohort effects can be included.

Therefore, the hypothesised mechanism is that an individual's relative rank in a task amongst peers affects self-concept. This in turn has an impact on non-cognitive skills like resilience, persistence and confidence which affects the costs of effort for that task or task-specific productivity directly. An exemplary basic behavioural model that works through the changing-cost channel is provided in Appendix 3[31]. Students want to maximise total grades for a given total cost of effort, and have differential abilities and costs of effort for each subject. Students who have a high rank in a subject during primary school develop a positive self-concept, and have a lower cost of effort in that subject in secondary school. This will shift the student's iso-cost line out along one axis for this subject and therefore they can reach higher isoquant and will optimally invest more effort in that subject (Appendix Figure 1, Panel B). If there were any general gains in confidence, that would reduce the cost of any academic effort and cause a parallel shift out of the iso-cost line and therefore more effort would be allocated to all subjects.

Applied to the results, the smaller estimates from the pupil fixed effects specifications, will have had any general confidence effect absorbed and so will only be picking up the effect of within student reallocation of effort across subjects. The school-subject-cohort effect specifications, allow for spillovers between subjects within a pupil and can so be interpreted as a culmination of the income and substitution effects of rank across subject and so are accordingly larger (Appendix Figure 1, Panel C).

To provide evidence for this mechanism we link the administrative data to the Longitudinal Survey of Yong People in England (LSYPE). We are able to match approximately twelve thousand students from the survey who answer questions on their self-concepts in each subject. This allows us to test directly if rank position within primary school has an effect on this measure of self-concept, conditional on attainment. The specifications are equivalent to (5) and (7) with the dependent variable now being student confidence. Since this survey was only run for one cohort, the school-subject-cohort effects are replaced by school-subject effects.

Panel A of Table 5 presents these results and demonstrates that conditional on age 11 test scores students with a higher primary school rank position are significantly more likely to say that they are good in that subject (column 1). Controlling for school-subject effects, the impact of moving from the bottom of class to the top is 0.196 points on a five point scale (-2, 2), or about twenty per cent of a standard deviation in our self-

[31] Self-concept may instead affect a student's ability in a task rather than cost of effort. This would lead to the same predicted changes in the effort ratios and empirical results. If we had time use data we would be able to differentiate between these causes, however given the data available, we are unable to determine if it is costs or abilities that are affected.

concept measure (see column 2).[32] This suggests that students develop a clear sense of their strengths and weaknesses depending on their local rank position, conditional on relative test scores.

While we would prefer to have a measure of self-concept directly at age 11 at the end of primary school, these measures are only available to us just prior to the age 14 tests. Therefore, in Panel B we additionally control for contemporaneous attainment at age 14, which is an outcome. To cautiously interpret these estimates, students with 'the same' age 11 and 14 results have higher self-concept if they have had a higher local rank in that subject in primary school.

Note column (2), the specifications allowing for primary-subject effects cannot reject the null hypothesis that rank has no effect on self-concept. A reason for this is that there are few students per primary school in this survey (4.5 students conditional on at least one student being in the survey); as the survey was conducted at secondary schools. The small number of students per school severely limits the degrees of freedom in each school-subject group, the lack of variation is exacerbated due to the coarseness of the self-concept variable. This is exacerbated further when additionally conditioning on individual student effects column 3. To obtain a clearer view of the effect of rank on contemporaneous self-concept we estimate how rank based on age-14 test scores within a secondary school subject affects subject confidence conditional on secondary-subject effects and individual effects. The advantage of this is that there are on average 20 students for each school that has students in the survey.[33] These results can be found in Panel C, where we see that conditional on school-subject effects, moving from bottom to top of class improves confidence by 0.43 on the 5 point scale. Allowing for individuals to have different levels of confidence and only using the variation between subjects reduces the parameter to 0.38 but remains significant at 1% (column 3).

Furthermore, we examine the heterogeneity of these effects by estimating the effect of age 14 rank on confidence separately by gender, conditional on student and school-subject effects (lower part of Panel C). We find that the effect on male confidence is five times larger than the effect on females ($\hat{\beta}_{rank\,male}$ =0.61, $\hat{\beta}_{rank\,female}$ =0.12), which mirrors the results we find for the effect of rank on later test scores. Unfortunately due to

[32] The standard deviation of the self-concept measure is 0.99.

[33] The reason why we do not look at the effect of KS3 rank on later outcomes is due to the tracking by subject in secondary school, which will be related to rank. This is not an issue with primary school rank, because even if there were tracking in primary schools, when moving to secondary school, students with the same test scores (but different primary ranks) would be assigned to the same track.

the smaller sample size of the LSYPE, we are unable to produce the effects non-linearly or by FSME status.

The magnitudes of the secondary school ranks effects on secondary confidence are large, but we may expect the contemporaneous effect of primary rank on confidence at age 11 to be even larger, as self-concept is thought to be more malleable at this age (Tiedemann, 2000; Lefot et al., 2010; Rubie-Davis 2011). Moreover, we find indicative evidence that later confidence is affected by previous primary school rank.

Overall given the effects of rank on direct measures of self-concept and the heterogeneity of effects found in the main results we are confident in our conclusion that self-concept forms according to rank position and that this affects later investment decisions.

8 Corroborating research

The finding that higher peer quality could have negative effects on later outcomes may seem controversial, but there are a number of topics in education that have findings which corroborate this hypothesis.

Research on selective schools and school integration have shown mixed results from students attending selective or predominantly non-minority schools (Angrist and Lang, 2004; Clark 2010; Cullen, et al., 2006; Kling *et al.,* 2007). Many of these papers use a regression discontinuity design to compare the outcomes of the students that just passed the entrance exam to those that just failed. The general puzzle is that many papers find no benefit from attending these selective schools. However, our findings would speak to why the potential benefits of prestigious schools may be attenuated through the development of negative self-concepts amongst these marginal/bussed students, who necessarily would also be the low ranked students. This is consistent with Cullen, Jacob and Levitt (2006), who find that those whose peers improve the most gain the least: 'lottery winners have substantially lower class ranks throughout high school as a result of attending schools with higher achieving peers and are more likely to drop out'. Similar effects are found in the Higher Education literature with respect to affirmative action policies (Arcodiacono *et al.,* 2012; Robles and Krishna, 2012).

The early formation of self-concept and specialisation could also partly explain why some achievement gaps increase over the education cycle. Widening overall education gaps have been documented by race (Fryer and Levitt, 2006; Hanushek and Rivkin 2006; 2009), small differences in early overall attainment could negatively affect general academic self-concept, which would lead to decreased investment in education and exacerbate any initial differences. In the case of gender a gap occurs by subject, where males are overly represented in mathematics and science by the age of 18, despite girls outperforming boys at early ages in these subjects (Burgess et al, 2004; Machin & McNally, 2005). Even with girls performing better in all subjects, if boys do comparatively less badly in mathematics and are more affected by rank for investment decisions, then they would chose to invest more in those subjects. Finally the literature on age-effects in education shows that older children do better compared to their younger peers (i.e. Black *et al.*, 2011). The development of positive self-concepts of the older children at an early age due to initial differences is a potential mechanism for the continuation of these effects as the students grow older.

9 Conclusions

Individuals continuously make social comparisons, which can affect our beliefs and investment decisions. If individuals make these comparisons using ordinal as well as cardinal information, then an individual's rank amongst their peers could impact on their investments and later productivity.

This paper examined how, conditional on relative achievement, rank amongst peers affects subsequent performance. Applied to an education setting we establish a new result, that rank position within primary school has significant effects on secondary school achievement. Moreover, a higher rank also improves students' confidence, an important non-cognitive skill. These rank effects are in addition to any effect caused by a student's relative distance from the class mean.

The approximately linear impact of rank implies that students are very good at determining their rank amongst their peers. Furthermore, there is significant heterogeneity in the effect of rank with males being influenced considerably more. We find male confidence in a subject is five times more affected by their rank amongst their peers compared to females. Accordingly, male students specialise more according to their primary school rank than females. To the extent that boys gain four times more in later test scores from being top of the class compared to comparable female students. Contrastingly, students with low parental income background are not negatively affected by low rank positions during primary education. Together, we take this as evidence that an individual's rank amongst their peers during primary school affects their self-concepts over many dimensions which in turn are likely to impact on the development of task specific non-cognitive skill and subsequent investment decisions.

We cannot fully exclude other mechanisms, such as learning about ability, to generate parts of these results. However, we have shown that differences between local and national ranks have no negative impact on average performance. This speaks in favour of mechanisms that change the actual grade production function either through shifting task-specific productivity or cost, and against learning models where only student perceptions are affected. Given the impact of rank on a direct measure of confidence, we thus believe that rank is most likely to affect later results through non-cognitive skills.

It is worthwhile to think about policy implications of this finding. With specific regards to education, these findings lead to a natural question for a parent deciding on where to send their child (in partial equilibrium). Should my child attend a ‘prestigious school’ or a ‘worse school’ where she will have a higher rank? Rank is just one of many factors in the education production function, and therefore choosing solely on the basis of rank is unlikely to be correct. The authors are currently not aware of any study that identifies the effectiveness of schools in terms of standard deviations[34]; therefore, we use estimates of the impact of teachers as an indicative measure for effects of school quality for this benchmarking exercise. A teacher who is one standard deviation better than average improves student test scores by 0.1 to 0.2 standard deviations (Aaronson, et al. 2007; Rivkin et al. 2005). Comparatively we find that a student with one standard deviation higher rank in primary school will score 0.08 standard deviations better at age 14.[35] Forthcoming work will look at the longer run impacts of primary school rank, as well as changes in school ranks from moving schools.

We believe these findings have general implications for productivity and informational transparency. To improve productivity it would be optimal for managers or teachers to highlight an individual’s local rank position if that individual had a high local rank. If an individual is in a high-performing peer group and therefore may have a low local rank but a high global rank a manager should make the global rank more salient. For individuals who have low global and local ranks, managers should focus on absolute attainment, or focus on other tasks where the individual has higher ranks.

[34] Evaluations of school effectiveness using admission lotteries (i.e. Hoxby et al. 2009, Angrist et al. 2010, Dobbie and Fryer 2011, Abdulkadiroglu et al. 2011) are comparing effectiveness between types rather than the whole distribution of effectiveness.

[35] Note that these are still not directly comparable because the effect of the teacher is annual and quickly fades out, whereas the rank treatment lasts the duration of primary school (5 years) and the effect is found three years later.

Finally these findings have general implications regarding the formation of non-cognitive skills and productivity. Given the heterogeneous effects of rank it would be possible to organise groups by individuals characteristics and abilities to maximise output. However this would be very cumbersome and administratively intensive. Therefore the key implication is that non-cognitive skills such as confidence, perseverance and resilience have large effects on productivity. Rank can be thought of as just one treatment that impacts on these behaviours, however there are many other interventions that could have positive effects on all individuals within a group and not just those above the median.

References

Aaronson, D., Barrow, L. and Sander, W. (2007). Teachers and student achievement in the Chicago public high schools, *Journal of Labor Economics*, 25(1), pp. 95–135.

Abdulkadiroglu, A., Angrist, J., Dynarski, S., Kane, T., and Pathak, P. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilot, The *Quarterly Journal of Economics*, 126(2), pp. 699-748.

Angrist, J. (2013). The Perils of Peer Effects, NBER Working Paper No. 19774.

Angrist, J and Lang, K. (2004). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program, *American Economic Review*, 94(5), pp. 1613-1634.

Angrist, J., Dynarski, S., Kane, T., Pathak, P., and Walters, C. (2010). Inputs and impacts in charter schools: KIPP Lynn, *American Economic Review*, 100(2), pp. 239-243.

Arciciacono, P., Aucejo, E. and Spenner, K. (2012). What Happens After Enrolement? An Analysis of the Time Path of Racial Differences in GPA and Major Choice, *IZA Journal of Labor Economics*, 1(5).

Azmat, G, and Iriberri N. (2010) The importance of relative performance feedback information: Evidence from a natural experiment using high school students, *Journal of Public Economics*, 94(7–8), pp. 435-452.

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2011). Too Young to Leave the Nest? The Effects of School Starting Age, *The Review of Economics and Statistics*, 93(2), pp. 455–467.

Borghans, L., Duckworth A., Heckman J., Weel ter Bas (2008). The Economics and Psychology of Personality Traits, *Journal of Human Resources*, 43(4), pp.872-1059.

Brown G., Gardner, J., Oswald, A. and Qian, J. (2008). Does Wage Rank Affect Employee's Well-being?, *Industrial Relations*, 47(3), pp. 355-389.

Brown J., (2011). Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars, *Journal of Political Economy*, 119(5), pp. 982-1013.

Burgess S., McConnell B., Propper C., Wilson D., Girls Rock, Boys Roll: An Analysis of the Age 14-16 Gender Gap in English Schools, *Scottish Journal of Political Economy*, Scottish Economic Society, 51(2), pp. 209-229.

Card, D., Mas, M., Moretti, E., and Saez, E. (2012). Inequality at Work: The Effect of Peer Salaries on Job Satisfaction, *American Economic Review*, 102(6), pp.2981-3003.

Carrell, S.E., Fullerton R.L., and West, J.E. (2009). Does Your Cohort Matter? Estimating Peer Effects in College Achievement, *Journal of Labour Economics*, 27(3), pp.439-464.

Clark, A., and Oswald, A. (1996). Satisfaction and Comparison Income, *Journal of Public Economics*, 61(September), pp. 359-381.

Clark, A., Masclet, D. and Villeval M.C., (2010). Effort and comparison income: experiment and survey evidence, *Industrial and Labour Relations Review,* 63(3) pp. 407-426.

Clark, D. (2010) Selective Schools and Academic Achievement *The B.E. Journal of Economic Analysis and Policy*, 10(1), pp. 1935-1682.

Cullen, J.B., Jacob, B.A. and Levitt S., (2006). The effect of school choice on participants: evidence from randomised lotteries, *Econometrica*, 74(5), pp.1191-1230.

DFE (2011). Class Size and education in England evidence report, Research Report DFE-RR169.

Dobbie, W. and Fryer, R. (2011). Are high quality schools enough to close the achievement gap? Evidence from the Harlem Children's Zone, *American Economic Journal: Applied Economics*, 3(3), pp. 158-187.

Eriksson, T., Poulsen, A., Villeval, M., (2009). Feedback and incentives: experimental evidence, *Labour Economics*, 16, pp. 679–688.

Falk, A., Ichino, A., (2006) Clean evidence on peer pressure *Journal of Labor Economics* 24 (1), pp. 39–57

Festinger, L. (1954). A theory of social comparison processes *Human Relations*, 7, pp. 117–140.

Fryer, G. and Levitt, S. (2006). The Black-White Test Score Gap Through Third Garde, *American Law and Economics Review*, 8(2), pp. 249-281.

Guimaraes, P. and Portugal, P. (2010) A Simple Feasible Alternative Procedure to Estimate Models with High-Dimensional Fixed Effects, *Stata Journal*, 10(4), pp. 628-649.

Hanushek, E. (2006). School Resources, *Handbook of the Economics of Education*, 2(Chapter 14).

Hanushek, E., Kain, J. and Rivkin, S. (2005). Teachers, Schools, and Academic Achievement, *Econometrica*, 73(2), pp. 417–458.

Hanushek, E. and Rivkin, S. (2006). School Quality and the Black-White Achievement Gap, NBER Working Paper No. 12651.

Hanushek, E. and Rivkin, S (2009). Harming the Best: How Schools Affect the Black-White Achievement Gap, *Journal of Policy Analysis and Management*, 28(3), pp. 336-393.

Heckman J., Stixrud, J. and Urzua, S. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior, *Journal of Labor Economics*, 24(3), pp. 411-482.

Hoxby, C. (2001). If Families Matter Most, Where Do Schools Come In? in T. Moe, ed. A Primer on American Schools. Stanford: Hoover Institution Press.

Hoxby, C., Murarka, S., and Kang, J. (2009). How New York City's charter schools affect achievement., Cambridge, MA: New York City Charter Schools Evaluation Project.

Kling, J., Liebman, J. and Katz, L. (2007). Experimental Analysis of Neighborhood Effects, *Econometrica*, 25(1), pp. 83-119.

Kosfeld, M. and Neckermann, S. (2011). Getting More Work for Nothing? Symbolic Awards and Worker Performance, *American Economic Journals: Microeconomics*, 3(3), pp. 86-99.

Lavy, V., Silva, O. and Weinhardt, F. (2012). The Good, The Bad and The Average: Evidence on Ability Peer Effects in Schools, *Journal of Labor Economics*, 20(2), pp.367-414.

Lindqvist, E., and Vestman, R., 2011. The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment, *American Economic Journal: Applied Economics*, 3(1), pp. 101-28.

Kuziemko, I, Buell,R., Reich,T., and Norton,M. (2011). Last-place Aversion: Evidence and Redistributive Implications, NBER Working Paper No. 17234.

Machin, S., and McNally, S. (2005). Gender and student achievement in English schools, *Oxford review of economic policy,* 21(3). pp. 357-372.

Manski, C. (1993). Identification of Endogenous Social Effects: The Reflection Problem, The Review of Economic Studies, 60(3), pp. 531-542.

Marsh, H., Byrne, B. and Shavelson, R. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement, *Journal of Educational Psychology*, 80, pp.366–380.

Mas, A. and Moretti, E., (2009). Peers at work, *American Economic Review*, 99(1), pp. 112–145.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects, *Econometrica*, 49(6), pp. 1417-1426.

Page, M., Carrell, S., and West, J. (2010). Sex and Science: How Professor Gender Affects the Gender Gap, *Quarterly Journal of Economics,* 125(3), pp.1101-1144.

Parducci, A (1965). Category Judgment: A Range-frequency Theory, *Psychological Review*, (72), pp. 407-418.

Robles, F. V.C. and Krishna, K. (2012). Affirmative Action in Higher Education in India: Targeting, Catch Up, and Mismatch, NBER Working paper No.17727.

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school, *Journal of Educational Psychology*, 92(1), pp 144-151.

Todd, P.E. and Wolpin K.I., (2003). On The Specification and Estimation of The Production Function for Cognitive Achievement, *The Economic Journal*, 113(1), pp. 3-33.

Tversky A. and Kahneman, D., (1974). Judgment under Uncertainty: Heuristics and Biases, *Science*, 185(4157), pp. 1124-1131.

Valentine, J., DuBois, D. and Cooper, H. (2004). The relations between self-beliefs and academic achievement: A systematic review, *Educational Psychologist*, (39), pp.111–133.

Watkins, M., Lei, P. and Canivez, L. (2007). Psychometric intelligence and achievement: A cross-lagged panel analysis, *Intelligence*, 35, pp.59-68.

Yeung, A. and Lee, F. (1999). Self-concept of high school students in China: Confirmatory factor analysis of longitudinal data, *Educational and Psychological Measurement*, *59,* pp.431–450.

Table 1: Descriptive statistics of Student Characteristics

	Mean	S.D.	Min	Max
Panel A: Student Characteristics				
Age 11 test scores				
KS2 English	50.29	28.03	1	100
KS2 Math	50.52	28.19	1	100
KS2 Science	50.01	28.03	1	100
Within Student KS2 S.D.	12.68	7.70	0	57.16
Age 14 test scores				
KS3 English	51.23	28.18	1	100
KS3 Math	52.89	27.55	1	100
KS3 Science	52.91	27.53	1	100
Within Student KS3 S.D.	11.32	7.19	0	56.59
Panel B: Rank Characteristics				
Rank English	0.488	0.296	0	1
Rank Math	0.491	0.296	0	1
Rank Science	0.485	0.295	0	1
Within Student Rank S.D.	0.138	0.087	0	0.58
Panel C: Background Characteristics				
SEN	0.175	0.380	0	1
FSME	0.146	0.353	0	1
Male	0.499	0.500	0	1
Ethnicity				
White British	0.837	0.370	0	1
Other White	0.019	0.135	0	1
Asian	0.058	0.234	0	1
Black	0.030	0.171	0	1
Chinese	0.003	0.053	0	1
Mixed	0.017	0.128	0	1
Other	0.011	0.104	0	1
Unknown	0.026	0.158	0	1

Notes: 6,815,997 observations over 5 cohorts. Cohort 1 takes Key Stage 2 (KS2) examinations in 2001 and Key Stage 3 (KS3) examinations in 2004. Cohort 5 takes KS2 in 2005 and KS3 in 2008. Test scores are percentalized tests scores by cohort-subject. All test scores come from national exams which are externally marked. The analysis stops in 2008 as after this point Key Stage 3 exams became internally assessed.

Table 2: Effect of Primary School Rank (Age 11) on Age 14 Test Scores

	Raw (1)	Primary (2)	Primary-Secondary (3)	Primary-Student (4)
Panel A: The effect of primary rank				
Primary Rank	11.551**	7.662**		
Flexible Age 11 Test Scores	*0.293*	*0.145*		
Primary Rank	11.001**	7.960**	7.901**	4.562**
Cubic Age 11 Test Scores	*0.298*	*0.145*	*0.146*	*0.132*
Panel B: The effect of placebo rank				
Placebo Rank	0.0055	0.015		
Flexible Age 11 Test Scores	*0.100*	*0.011*		
Placebo Rank	0.0045	0.013	0.016	-0.008
Cubic Age 11 Test Scores	*0.100*	*0.116*	*0.119*	*0.137*
Student characteristics	✓	✓	✓	Abs
Age 11 Test Scores	✓	✓	✓	✓
Primary-cohort-subject Effects		✓	✓	✓
Secondary Effects			Abs	Abs
Secondary-cohort-subject Effects			✓	
Student Effects				✓

Notes: Results obtained from twelve separate regressions based on 2,271,999 student observations and 6,815,997 student-subject observations. The dependent variable is by cohort by subject percentalized KS3 test scores. All specifications control for Key Stage 2 results, student characteristics, cohort effects and subject effects. Student characteristics are ethnicity, gender, free school meal (FSME) and special educational needs (SEN). Coefficients in columns (2) and (3) are estimated using Stata command reg2hdfe allowing two high dimensional fixed effects to be absorbed. Standard errors in italics and clustered at 3,800 secondary schools. Abs indicates that the effect is absorbed by another estimated effect. ** 1% sig.

Table 3: Balancing by Parental Occupation

	Primary (1)	Primary-Student (2)
Panel A: Effects on age-11 tests		
Parental Occupation	7.722**	1.706*
	0.840	*0.783*
Panel B: Effects on Ordinal Rank		
Parental Occupation	-0.004	0.000
	0.005	*0.034*
Primary-subject Effects	✓	✓
Student Effects		✓

Notes: Results obtained from regressions based on 31,050 subject-student observations for which parental occupations could be identified from the LSYPE. Detailed occupational coding available from the authors on request. Panel A has KS2 as dependent variable, in Panel B KS2 with polynomials up to cubic are included as controls. All regressions control for student characteristics and subject effects. Regressions in column (2) estimated using Stata command reg2hdfe. ** 1%,* 5% significant.

Table 4: Effect of Average Primary School Rank (Age 11) and Misinformation on Average Age 14 Test Scores

	Raw (1)	Primary (2)
Primary Rank	10.710**	10.694**
	0.223	*0.223*
Misinformation	-	-0.361
	-	*0.233*
Student characteristics	✓	✓
Age 11 Test Scores (cubic)	✓	✓
Primary-cohort-subject Effects	✓	✓

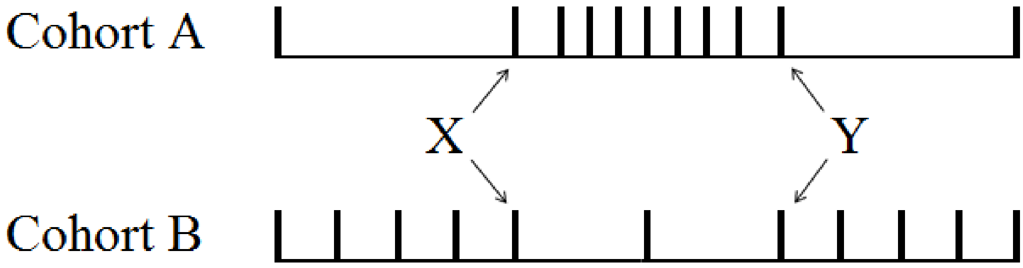
Notes: Results obtained from two separate regressions based on 2,271,999 student observations averaged over subjects where column (2) includes an additional explanatory variable on misinformation. The dependent variable is by cohort by subject percentalized average KS3 test scores. The misinformation measurement is the average absolute difference between local rank and national percentile rank for each student. Student characteristics are ethnicity, gender, free school meal (FSME) and special educational needs (SEN). Coefficients are estimated using Stata command reg2hdfe allowing two high dimensional fixed effects to be absorbed. Standard errors in italics and clustered at 3,800 secondary schools. ** 1% sig.

Table 5: Effect of Rank on Self-Concept

	(1)	(2)	(3)
Panel A: Self-Concept on Age 11 Test Scores			
Primary Rank	0.563**	0.196*	0.056
	0.038	*0.117*	*0.18*
Panel B: Self-Concept on Age 11 & 14 Test Scores			
Primary Rank	0.436**	0.109	0.014
	0.039	*0.115*	*0.079*
Panel C: Self-Concept on Age 14 Test scores			
Secondary Rank	0.897**	0.427**	0.382**
	0.048	*0.099*	*0.155*
Secondary Rank – Male Students	0.754***	0.530***	0.606***
	0.059	*0.126*	*0.206*
Secondary Rank – Female Students	1.067***	0.317*	0.115
	0.071	*0.166*	*0.233*
School-by-subject effects		✓	✓
Student Effects			✓

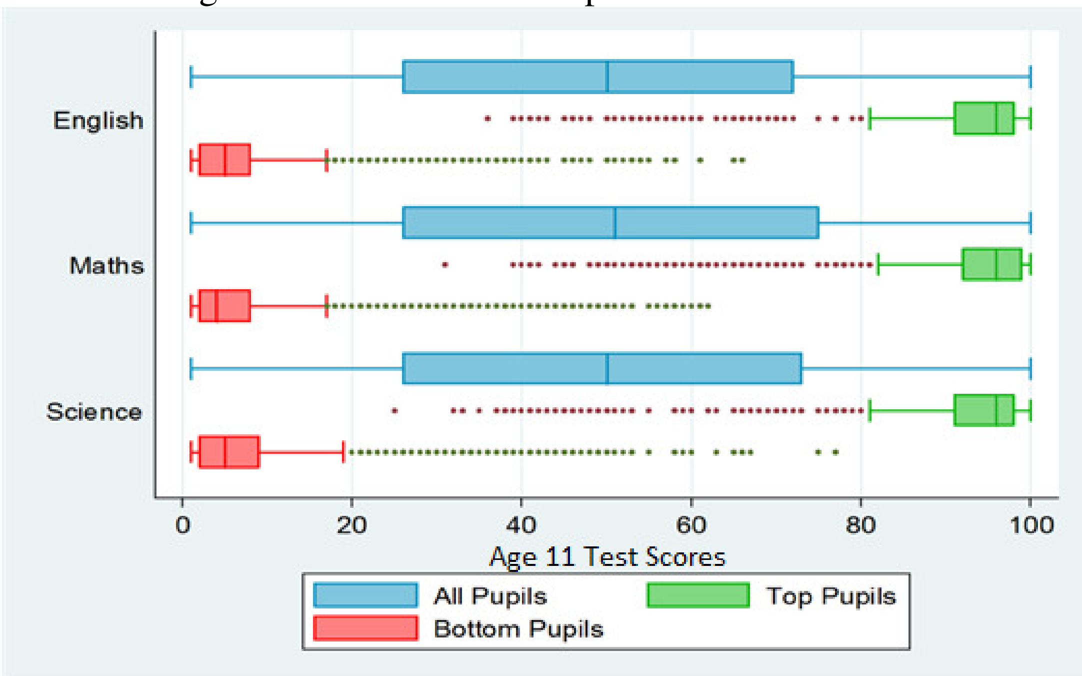
Notes: Results obtained from fifteen separate regressions based on 11,558 student observations and 34,674 student-subject observations from the LSYPE sample (17,415 female, 17,259 male). For descriptives, see Appendix Table 3. The dependent variable is a course measure of self-concept by subject. All specifications in columns 1 and 2 control for observable student characteristics, these are absorbed by the student effect in column 3. Student characteristics are ethnicity, gender, free school meal (FSME) and special educational needs (SEN).Panels A and B condition on age 11 test scores (cubic) and primary school by subject effects. Panels B and C condition on age 14 test scores (cubic) and secondary school by subject effects. Cohort effects are not included because the LSYPE data is only available for one cohort. Standard errors in parenthesis and clustered at 796 secondary schools ** 1% sig. * 10% sig.

Figure 1: Rank Dependent on Distribution Given Absolute and Relative Score



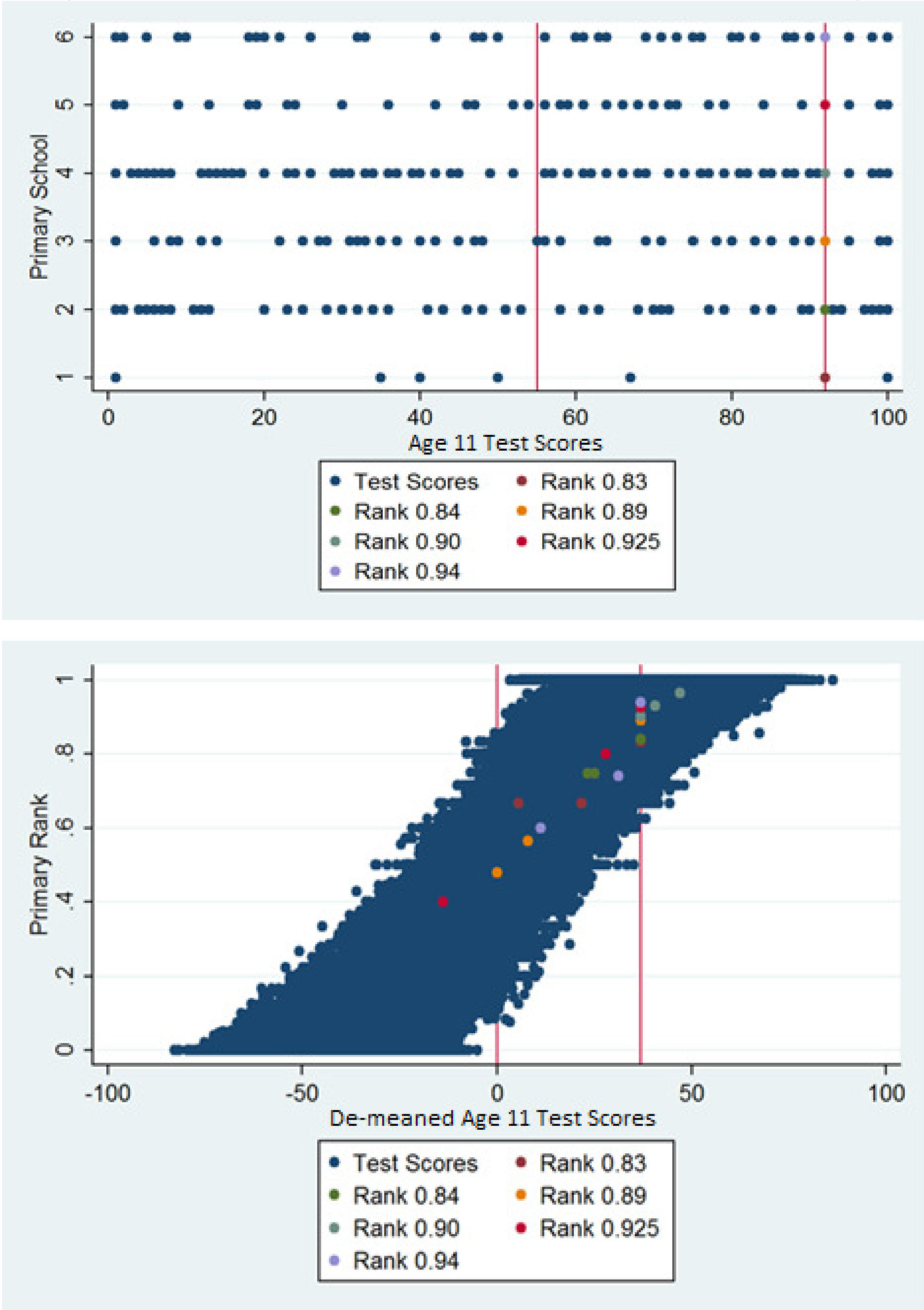
Notes: This figure illustrates that students with the same test score relative to the group mean can have different ranks depending on the distribution of test scores. Two cohorts of eleven students are represented, with each mark representing a student's test score. Test scores are increasing from left to right. Each cohort has the same minimum, maximum and mean test scores. Cohort A has a unimodal distribution and Cohort B has a bimodal test score distribution. A student with a test score of X in Cohort A would have a lower rank than the same test score in Cohort B. Similarly a test score of Y would be ranked differently in Cohorts A and B. Given the definition of rank given in Section 5.2, the rank measurements for score X are R_{xA}= 0.1 and R_{xB}= 0.4 and for Y are R_{yA}= 0.9, R_{yB}= 0.6. This is based on the illustration from Brown et al. (2008).

Figure 2: Distribution of Top and Bottom Students



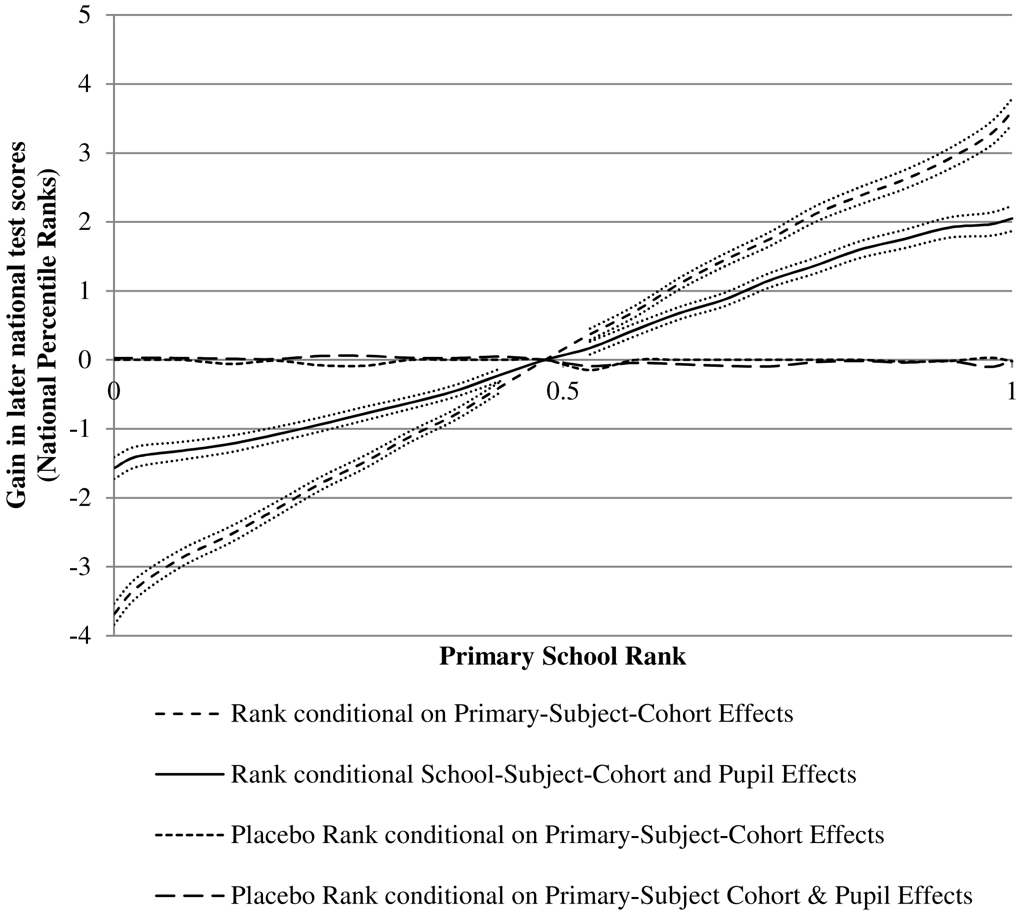
Notes: Box plots of age 11 test scores. Top Students are defined as students ranked in the top 5% of their school-subject-cohort ($Top\ Pupil = R_{ijs}[0.95, 1]$). Bottom Students are defined as students ranked in the bottom 5% of their school-subject-cohort ($Bottom\ Pupil = R_{ijs}[0, 0.05]$). The ends on the whiskers mark the most extreme value within 1.5 inter-quartile-ranges of the nearest quartile. Note that individual test scores have been randomly altered enough to ensure anonymity of individuals and schools. They are for illustrative purposes only, and in no way affects the interpretation of these figures.

Figure 3: Rank Distributions in Schools and Across Subjects



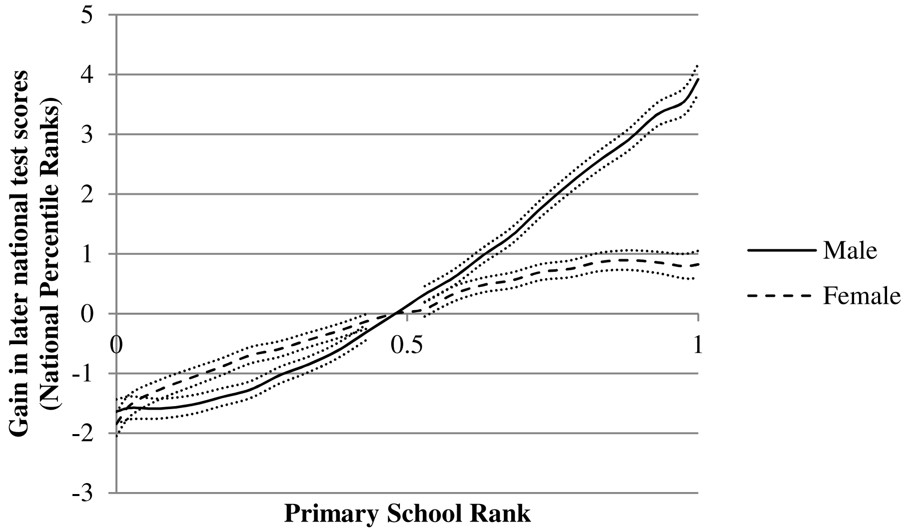
Notes: In the upper panel each point represents a student's Key Stage 2 test score. The six schools that are represented have the same mean (54), minimum (0) and maximum (100) tests scores in English, and also have a student with a test score of 93. Each student with the test score of 93 has a different rank. The lower panel shows all students in our data. The Y-axis is the primary rank of students and the X-axis shows the de-meaned test scores by primary school-subject-cohort. The colored points represent the three different test scores and ranks of students from Figure 5 with a test score of 93 in English. Note that the number of students per school as well as individual test scores have been randomly altered enough to ensure anonymity of individuals and schools. They are for illustrative purposes only and in no way affects the interpretation of these figures.

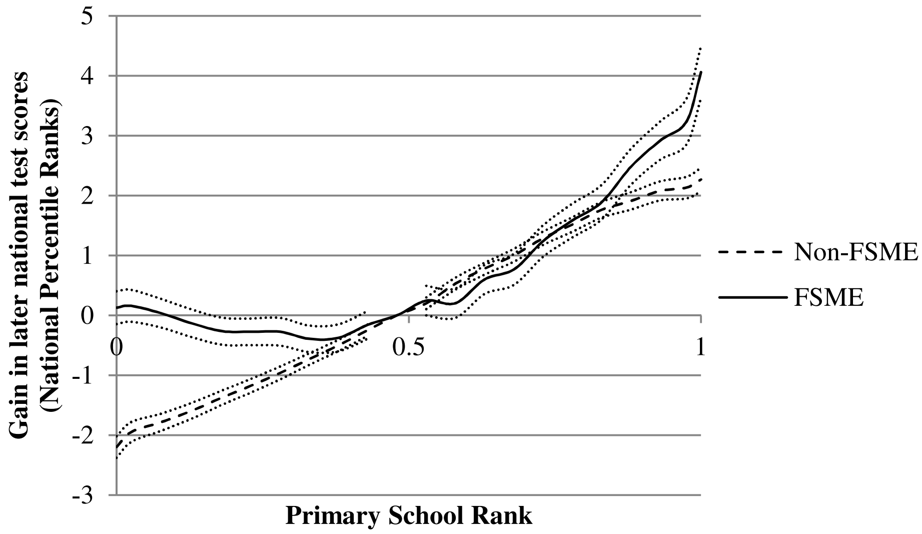
Figure 4: Effect of Placebo Primary School Rank on Secondary School Outcomes



Notes: Non-linear effect with dummies for the vingtiles of rank plus a dummy for being top or bottom of school-subject-cohort. All specifications have subject specific rank and test score across three subjects. Placebo rank generated from actual test scores but randomly allocated peers, using the actual distribution of primary school size. All standard errors clustered at the actual secondary school attended. Specification 1: Student characteristics and primary, subject and cohort effects. Specification 2: Primary-subject-cohort group effects and student effects. Dashed lines represent 95% confidence intervals.

Figure 5: Effect of Primary School rank on Secondary School outcomes by Student Characteristics





Notes: FSME stands for Free School Meal Eligible student. Effects obtained from estimating the effect of rank on Non-FSME (Female) students and the interaction term with FSME (Male) students. Non-linear effect with dummies for the vingtiles of rank plus a dummy for being top or bottom of school-subject-cohort. All estimates use subject specific rank and test score across three subjects and condition on Primary-subject-cohort group effects and student effects. Dashed lines represent 95% confidence intervals.

Appendix 1: Peer Effects

There are concerns that with the existence of peer effects, peer quality jointly determines both a student's rank position, as well as their age 11 results. This mechanical relationship could potentially bias our estimation. This is because in the presence of peer effects a student with lower quality peers would attain a lower age 11 test scores than otherwise and also have a higher rank than otherwise. Thus, when controlling for prior test scores in the age 14 estimations, when students have a new peer group, those who previously had low quality peers in primary school would appear to gain more. Since rank is negatively correlated with peer quality in primary, it would appear that those with high rank make the most gains. Therefore, having a measure of ability confounded by peer effects would lead to an upward-biased rank coefficient.

This situation could be present in our data. We propose a resolution through the inclusion of subject-by-cohort-by-primary school controls. These effects will absorb any average peer effects within a classroom. However, they will not absorb any peer effects that are individual specific. This is because all students will have a different set of peers (because they cannot be a peer to themselves). Therefore, including class level controls will only remove the average class peer effect. The remaining bias will be dependent on the difference between the average peer effect and the individual peer effect and its correlation with rank. We are confident that the remaining effect of peers on the rank parameter will be negligible, given that the difference between average and individual peer effect decreases as class size increases. The bias will be further attenuated because the correlation between the difference and rank will be less than one, and both effects are small.

We test this by running simulations of a data generating process where test scores are not affected by rank and are only a function of ability and school/peer effects; we then estimate the rank parameter given this data. We allow for the data-generating process to have linear mean-peer effects, as well as non-linear peer effects (Lavy *et al.* 2012). We are conservative and assume extremely large peer effects, allowing both types of peer effects to account for 10% of the variance of a student's subject-specific outcome. Given that the square root of the explained variance is the correlation coefficient, this assumption implies that a one standard deviation increase in peer quality improves test scores by 0.31 standard deviations. In reality Lavy *et al.* (2012) find a 1sd increase in peers only increases test scores by 0.015standard deviations, a 20th of the size.

The data generating process is as follows:

- We create 2900 students to 101 primary schools and 18 secondary schools of varying school sizes[36].
- A range of factors are used to determine achievement. Each of these factors are assigned a weight, such that the sum of the weights equal 1. This means weights can be interpreted as the proportion of the explained variance.
- Students have a general ability α_i and a subject specific ability δ_{is} taken from normal distributions with mean 0 and standard deviation 1. Taken together they are given a weighting of 0.7 as the within school variance of student achievement in the raw data is 0.85. Or a weight of 0.6 where rank effects exist.
- All schools are heterogeneous in their impact on student outcomes, which are taken from normal distributions with mean 0 and standard deviation 1. School effects are given a weighting of 0.1 as the across school variance in student achievement in the raw data is 0.15.
- Linear mean peer effects are the mean subject and general ability of peers not including themselves. Non-linear peer effect is the negative of the total number of peers in the bottom 5% of students in the population in that subject. Peer effects are given a weight of 0.1 much larger than reality.
- We allow for measurement error in test scores to account for 10% of the variance.
- We generate individual's i test scores as a function of general ability α_i subject specific ability δ_{is}, primary peer subject effects ρ_{ijs} or secondary peer subject effects σ_{iks}, primary school effects μ_j or secondary school effects π_k , Age 11 and 14 measurement error ε_{ijs} or ϵ_{ijks} , and primary school Rank ω_{ijs}.
    - Age 11 test scores
      Y_{ijs1}=0.7*(α_i+ δ_{is}) +0.10*μ_j +0.1*ρ_{ijs} +0.1ε_{ijs}
    - Age 14 test scores where rank has no effect (Panel A):
      Y_{ijks2}=0.7*(α_i+ δ_{is}) +0.10* π_k+0.1*σ_{iks} +0.1ϵ_{ijks}

[36] Primary school sizes; 14 students, 16, 25 students (x4 schools), 26 students (x5), 27 students (x10), 28 students (x10), 29 students (x10), 30 students (x60). Secondary School sizes: 26 students, 89 students, 153 students, 160 students, 162 students, 170 students, 174 students, 178 students, 180 students (x9),

- Age 14 test scores where rank has an effect (Panel B):

  Y_{ijks2}=0.6*(α_i+ δ_{is}) +0.10* π_k+0.1*σ_{iks}+ 0.1ω_{ijs} +0.1ϵ_{ijks}

We simulate the data 1000 times and then estimate the rank parameter using the following specifications, with and without school-subject effects.

$$Y_{ijks2} = \beta_{rank}Rank_{ijs} + \beta_{y1}Y_{ijs1} + \varepsilon_{ijks}$$

$$Y_{ijks2} = \beta_{rank}Rank_{ijs} + \beta_{y1}Y_{ijs1} + \sigma_{ijs} + \epsilon_{ijks}$$

The results from these estimations can be found in appendix Table 1 & 2 below. When rank does not have an effect, we would expect $\beta_{rank} = 0$, and when it does, $\beta_{rank} = 0.1$. With these inflated peer effects sizes, we find that controlling for school-subject-cohort removes enough of the positive bias to make the effect of peers negligible (Appendix Table 1 & 2, column 3). If there are large non-linear peer effects, then this specification introduces a negative bias; therefore our results could be seen as upper bounds (Appendix Table 2, column 3).

Appendix Table A1: Simulation of Rank Estimation with Peer Effects

	Mean peer effects		*Non-linear Peer Effects*	
	(1)	(2)	(3)	(4)
Panel A: Rank has no effect β_{rank}=0.0				
Mean $\hat{\beta}_{rank}$	0.046	0.000	0.302	-0.041
Mean SE of $\hat{\beta}_{rank}$	0.014	0.018	0.015	0.019
SE of $\hat{\beta}_{rank}$	0.015	0.019	0.031	0.020
95% Lower Bound	0.015	-0.037	0.243	-0.079
95% Upper Bound	0.077	0.035	0.364	-0.003
Panel B: Rank has an effect β_{rank}=0.1				
Mean $\hat{\beta}_{rank}$	0.099	0.100	0.304	0.068
Mean SE of $\hat{\beta}_{rank}$	0.014	0.017	*0.014*	*0.018*
SE of $\hat{\beta}_{rank}$	0.015	0.018	0.027	0.018
95% Lower Bound	0.069	0.066	0.252	0.033
95% Upper Bound	0.129	0.133	0.358	0.104
KS2 and Rank	✓	✓	✓	✓
School-Subject-Effects		✓		✓

Notes: 1000 iterations, 95% confidence bounds are obtained from 25^{th} and 975^{th} estimate of ordered rank parameters..

Appendix 2: Measurement Error in Test Scores

Test scores are scores are an imprecise measure of ability. Could this measurement error be driving the results? Given that rank and test scores will both be affected by the same measurement error (but to different extents due to heterogeneous test score distributions across classes), calculating the size of the bias is intractable. To gauge the potential effect of measurement error, we simulate the data generating process. This allows us to have a true measure of ability and a student test score of which 20% of the variation is measurement error. Comparing the estimates of the rank parameter both with and without measurement error provides us an indication of the extent to which measurement error could be driving the results. Rank measurement is then derived from the noisy test score measure in both cases.

The data generating process is as follows:

- 2900 students to 101 primary schools and 18 secondary schools of varying sizes[37].
- A range of factors are used to determine achievement. Each of these factors are assigned a weight, such that the sum of the weights equal 1. This means weights can be interpreted as the proportion of the explained variance.
- Students have a general ability α_i and a subject specific ability δ_{is} taken from normal distributions with mean 0 and standard deviation 1. Taken together they are given a weighting of 0.7 as the within school variance of student achievement in the raw data is 0.85. Or a weight of 0.6 where rank effects exist.
- All schools are heterogeneous in their impact on student outcomes, which are taken from normal distributions with mean 0 and standard deviation 1. School effects are given a weighting of 0.1 as the across school variance in student achievement in the raw data is 0.15.
- We allow for measurement error in test scores to account for 20% of the variance, double the effect of any school subject effects.
- We generate individual's i test scores as a function of general ability α_i subject specific ability δ_{is}, primary school effects μ_j or secondary school

[37] Primary school sizes; 14 students, 16, 25 students (x4 schools), 26 students (x5), 27 students (x10), 28 students (x10), 29 students (x10), 30 students (x60). Secondary School sizes: 26 students, 89 students, 153 students, 160 students, 162 students, 170 students, 174 students, 178 students, 180 students (x9),

effects π_k , Age 11 and 14 measurement error ε_{ijs} or ϵ_{ijks} , and primary school Rank ω_{ijs}

- Age 11 test scores

  Y_{ijs1}=0.7*(α_i+ δ_{is}) +0.10*μ_j +0.2ε_{ijs}
- Age 14 test scores where rank has no effect (Panel A):

  Y_{ijks2}=0.7*(α_i+ δ_{is}) +0.10* π_k +0.2ϵ_{ijks}
- Age 14 test scores where rank has an effect (Panel B):

  Y_{ijks2}=0.6*(α_i+ δ_{is}) +0.10* π_k+ 0.1ω_{ijs} +0.2ϵ_{ijks}

We simulate the data 1000 times and then estimate the rank parameter using the following specifications, with and without school-subject effects, controlling either for true ability (α_i+ δ_{is}) or age 11 test scores.

$$Y_{ijks2} = \beta_{rank}Rank_{ijs} + \beta_{Ability}Ability_{ijs} + \varepsilon_{ijks}$$

$$Y_{ijks2} = \beta_{rank}Rank_{ijs} + \beta_{Ability}Abiliyt_{ijs} + \sigma_{ijs} + \epsilon_{ijks}$$

$$Y_{ijks2} = \beta_{rank}Rank_{ijs} + \beta_{Y1}Y_{ijs1} + \varepsilon_{ijks}$$

$$Y_{ijks2} = \beta_{rank}Rank_{ijs} + \beta_{Y1}Y_{ijs1} + \sigma_{ijs} + \epsilon_{ijks}$$

The results of these specifications can be found in appendix Table 3 below. The ability specification produces unbiased results. When there is measurement error in the test score there is a downward bias to the rank effect when rank has an effect (Appendix Table 3, Column 5, Panel B). We find that including school-subject-cohort and student fixed effects removes this downward bias.

Appendix Table A2: Simulation with measurement error

	Condition on true ability: No measurement error			*Condition on test scores: Large measurement error*		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Rank has no effect β_{rank}=0.0						
Mean $\hat{\beta}_{rank}$	-0.001	0.000	0.000	-0.010	0.015	-0.000
Mean SE of $\hat{\beta}_{rank}$	0.021	0.020	0.025	0.029	0.034	0.041
SE of $\hat{\beta}_{rank}$	0.037	0.021	0.025	0.030	0.036	0.042
95% Lower Bound	-0.074	-0.039	-0.047	-0.068	-0.054	-0.082
95% Upper Bound	0.076	0.041	0.050	0.048	0.086	0.079
Panel B: Rank has an effect β_{rank}=0.1						
Mean $\hat{\beta}_{rank}$	0.099	0.100	0.100	0.053	0.113	0.100
Mean SE of $\hat{\beta}_{rank}$	0.021	0.020	0.025	0.027	0.032	0.038
SE of $\hat{\beta}_{rank}$	0.037	0.021	0.025	0.029	0.033	0.039
95% Lower Bound	0.026	0.061	0.053	-0.002	0.048	0.024
95% Upper Bound	0.176	0.141	0.150	0.109	0.179	0.176
Ability and Rank	✓	✓	✓	✓	✓	✓
School-Subject-Effects		✓	✓		✓	✓
Student Effects			✓			✓

Notes: 1000 iterations, 95% confidence bounds are obtained from 25^{th} and 975^{th} estimate of ordered rank parameters..

Appendix 3: Model of Effort Allocation

To explicitly describe the mechanism, we put forward a basic behavioural model of how rank could affect later actions through self-concept. There are two stages, a learning stage followed by an action stage. In the learning stage, individuals of heterogeneous ability in different tasks are randomly allocated into groups. They perform tasks and compare their abilities relative to others in their group. This forms their task specific self-concept and a general self-concept. In the second stage, individuals are put into a new peer group in which they perform the same tasks. The self-concepts formed in the first stage affects their costs of effort for each task in the second stage.[38] Individuals now allocate effort to each task to maximise output for a given level of effort and ability. In this simplified model we assume that individuals do not include later rank directly in their objective function.

Without losing generality, we apply this to the education setting where students vary in ability across subjects and are randomly allocated to primary schools where they form a self-concept in each subject during the first stage. This is generated through students interacting with their peers, such as observing who answers questions, and teacher grading. For the purposes of the model, we assume that students exert no effort during primary school, with outcomes being a product of ability and school factors.

In the second stage, we model students as grade maximising agents for a given total cost of effort T_i and subject ability level A_{is}. The grade achieved Y, by a student i in subject s is a function of ability A_{is} and effort E_{is} according to a to a separable production function where there are decreasing returns to effort in each subject, $0<\kappa<1$. For simplicity of notation, assume that there are only two subjects, s $=\{e, m\}$. The productivity of effort is additionally effected by subject specific school factors μ_s. The total test score of individual i is the sum of this function over subjects, therefore, for student i in school μ_{is} the education production is:

$$Y_i = f(A_{ie}, E_{ie}) + f(A_{im}, E_{im}) = \mu_{ie} \cdot A_{ie} \cdot E_{ie}{}^{\kappa} + \mu_{im} \cdot A_{im} \cdot E_{im}{}^{\kappa} \quad (11)$$

This can be rearranged in terms of E_{ie} so that an isoquant (Q_o) can be drawn for a given total grades Y_i, subject abilities and school effects, all the combinations of subject effort (see Figure 1).

[38] Self-concept instead affect an agent's ability in a task rather than cost of effort. This would lead to the same predicted changes in the effort ratios and empirical results. Given the data available, we are unable to determine if it is costs or abilities that are affected. With information on time allocated on each task a positive relationship with rank would imply cost reductions, whereas no changes or decreases would imply gains in ability. We have chosen costs, as this is the more parsimonious and intuitive of the two.

$$E_{ie} = \left(\frac{Y_i - \mu_{im} \cdot A_{im} \cdot E_{im}^{\kappa}}{\mu_{ie} \cdot A_{ie}}\right)^{\frac{1}{\kappa}} \tag{12}$$

The self-concept of each subject that was generated in the first stage determines the student's cost of effort. Those with a positive self-concept will find the cost of effort lower for example when faced with a difficult mathematics question, a student who considers herself good at mathematics would spend longer looking for a solution, compared to another student who may give up. Therefore, the cost of subject effort c_s is a decreasing function of school subject rank R_s, $C_s = g(R_s)\ where\ g' < 0$. We assume costs of subject effort are linear in effort applied to that subject. We also allow for a general cost of effort C_{ig}, which varies across individuals according to general academic self-concept and is a decreasing function of ranks in all subjects, $C_g = d(R_m, R_e)\ where\ d'(R_s) < 0$ for s={m,e}. This general cost function reflects a student's general attitude towards education, and is linear in the sum of effort applied across all subjects, $E_{im} + E_{ie}$. The total cost of effort T that a student can apply is fixed, however the inclusion of a general cost of academic effort term, means that the total effort applied by a student is very flexible.

$$T_i \geq C_{im} \cdot E_{im} + C_{ie} \cdot E_{ie} + C_{ig} \cdot (E_{im} + E_{ie}) \tag{13}$$

This allows for an isocost line to be drawn using the cost of effort in each subject as the factor prices for a given total effort (see Figure 1, Panel A). There is additionally a non-binding time constraint, normalising the total time available to one, $E_{ie} + E_{im} < 1$. As standard, the solution is where the technical rate of substitution equals the relative factor prices i.e. where the isoquant and isocost lines are tangential.

Therefore student *i* wants to maximise total grades by solving:

$$\max_{E_e E_m} Y(E_e, E_m) = f(E_e) + f(E_m)$$

$$= \mu_e \cdot A_e \cdot {E_e}^{\kappa} + \mu_m \cdot A_m \cdot {E_m}^{\kappa} \tag{11}$$

$$s.t.\ T \geq C_e \cdot E_e + C_m \cdot E_m + C_g \cdot (E_m + E_e) \tag{13}$$

$$1 > E_e + E_m$$

$$l = Y - \lambda(T - C_e E_e - C_m E_m - C_g \cdot (E_m + E_e))$$

$$\frac{dl}{dE_e} = 0 \ \rightarrow \ \frac{\partial Y}{\partial E_e} = \lambda(C_e + C_g)$$

$$\frac{dl}{dE_m} = 0 \ \rightarrow \ \frac{\partial Y}{\partial E_m} = \lambda(C_m + C_g)$$

$$\frac{dl}{d\lambda} = 0 \ \rightarrow C_e E_{ie} + C_m E_m + C_g \cdot (E_m + E_{ie}) = T$$

$$\frac{dY}{dE_s} = \kappa \cdot \mu_s \cdot A_s \cdot E_s^{\kappa-1}$$

Therefore

$$\kappa \cdot \mu_s \cdot A_s \cdot E_s^{\kappa-1} = \lambda\,(C_s + C_g)$$

Where λ reflects the marginal grade per effort and λ>0

$$\frac{\kappa \cdot \mu_e \cdot A_e E_e^{\kappa-1}}{(C_e + C_g)} = \lambda = \frac{\kappa \cdot \mu_m \cdot A_m E_m^{\kappa-1}}{(C_m + C_g)}$$

This gives

$$\frac{(C_e+C_g)}{(C_m+C_g)} = \frac{\mu_e \cdot A_e \cdot E_e^{\kappa-1}}{\mu_m \cdot A_m \cdot E_m^{\kappa-1}} \tag{14}$$

It is also clear that given this specification effort exerted in a specific subject is dependent on the student's ability and cost of effort in that subject and general self-concept.

$$E_{is}^* = \left(\frac{\lambda(C_{is}+C_g)}{\kappa \cdot \mu_s \cdot A_{is}}\right)^{\frac{1}{\kappa-1}} \tag{15}$$

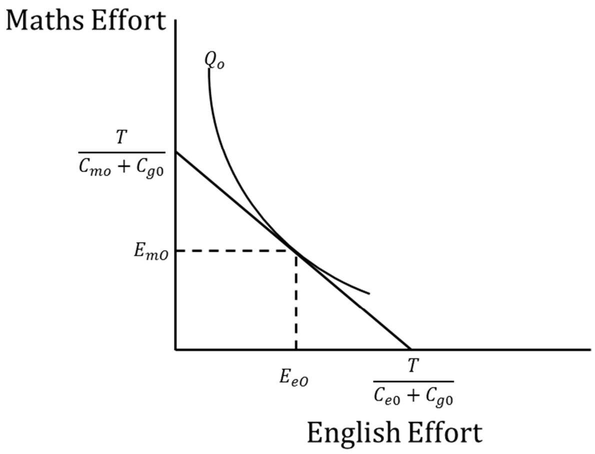
In the above λ reflects the marginal grade per effort where λ>0. As costs are decreasing in subject rank and $0 < \kappa < 1$ any increase in rank in subject s will increase the later effort allocated to that subject. A student with an improved English self-concept would now have a lower cost of learning English and therefore increase their English to math effort ratio. The reduced costs also induce an income effect as more effort can be allocated for the same total effort costs. The isocost line shifts outwards and a higher isoquant can be reached (Figure 1 Panel B). This student would now optimally chose to exert more effort in English ($E_1 > E_0$) and less effort in math ($M_1 < M_0$). As a result, the total grades that can be achieved for a given cost of effort and ability level is higher.

This has yet to take into account the reduction in general academic costs of effort C_g, due to a increase in general self-concept. An increase in general academic self-concept would reduce the cost for both subjects, and so there would only be an income effect. This shifts the isocost curve out, increasing the maximum possible English and Math effort that could be allocated (Figure 1 Panel C). Given this specification, the final effect on math effort is ambiguous, as it depends on the shape and position on the isoquants and the importance of general self-concept.

For the estimations that used the variation in rank within student, the individual effects absorb any individual general academic confidence gained by being ranked highly. These estimations are therefore equivalent to the case where C_g is fixed and we're just looking at the effect allocation across subjects. The specifications that do not include individual effects and instead use the within class variation do allow for spillover effects between subjects and so there can be general gains in confidence. This is the intuition for why the parameters recovered from the student effects estimations are smaller than those from the school cohort effect estimations. This two-subject example is for exposition only but easily extends to the setting where an individual is maximising total grades over three subjects.

Appendix Figure A1: Optimal Allocation of Effort

Panel A

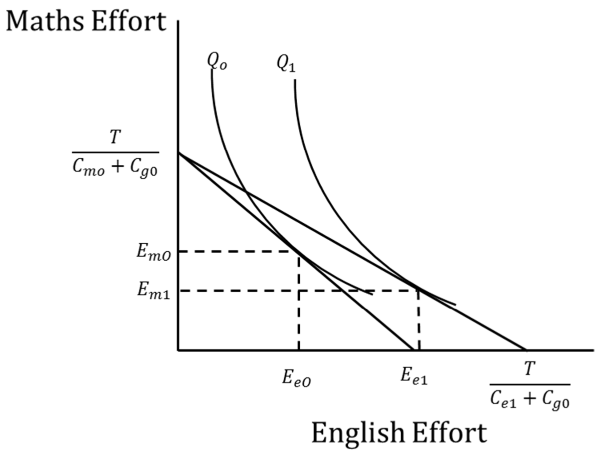


Isoquant Q_0:

$$E_{e0} = \left(\frac{Y_0-\mu_m \cdot A_m \cdot E_{m0}^{\kappa}}{\mu_e \cdot A_e}\right)^{\frac{1}{\kappa}}$$

Optimal English effort E_{e0} and math effort E_{m0}, given cost of English and math effort C_{eo} , C_{mo} . Marginal cost of effort equals marginal test score gain where isoquant and isocost curve are tangential.

Panel B

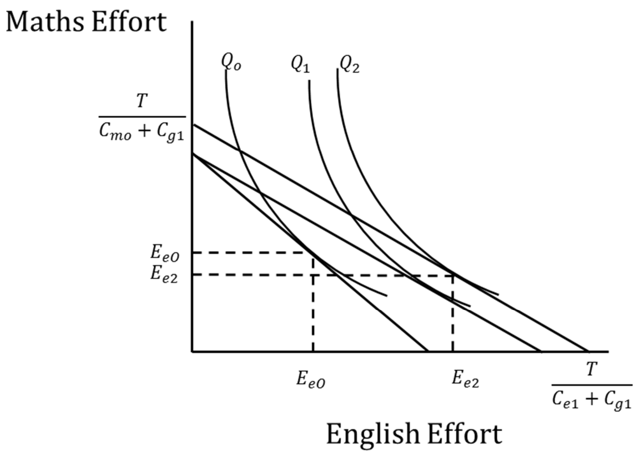


Isoquant Q_1:

$$E_{e1} = \left(\frac{Y_1-\mu_m \cdot A_m \cdot E_{m1}^{\kappa}}{\mu_e \cdot A_e}\right)^{\frac{1}{\kappa}}$$

A higher rank in English, improves English self-concept and reduces cost of effort in English $C_{e0}>C_{e1}$. Shifts isocost line out to new intercept on English axis. Increase English effort $E_{e0} < E_{e1}$ and decrease math effort $E_{mo} > E_{m1}$

Panel C



Isoquant Q_2:

$$E_{e2} = \left(\frac{Y_1-\mu_m \cdot A_m \cdot E_{m2}^{\kappa}}{\mu_e \cdot A_e}\right)^{\frac{1}{\kappa}}$$

A higher rank in English, also improves general self-concept and reduces cost of effort in both subjects $C_{s1}>C_{s2}$. Shifts isocost line out to new intercept on both axis. This increases effort applied in both subjects $E_{s1} < E_{s2}$.

Total effects: More effort applied to English $E_{e2} > E_{e0}$, ambiguous effect on math.

Table A3: Descriptive Statistics of LSYPE Sample

	Mean	S.D.	Min	Max
Panel A: Student Characteristics				
Age 11 test scores				
KS2 English	49.48	27.77	1	100
KS2 Math	50.11	28.37	1	100
KS2 Science	48.69	28.30	1	100
Within Student KS2 S.D.	12.71	7.69	0	47.44
Age 14 test scores				
KS3 English	50.67	28.00	1	100
KS3 Math	52.99	27.61	1	100
KS3 Science	52.21	27.71	1	100
Within Student KS3 S.D.	12.71	7.69	0	47.44
Panel B: Rank Characteristics				
Rank English	0.491	0.295	0	1
Rank Math	0.496	0.297	0	1
Rank Science	0.482	0.294	0	1
Within Student Rank S.D.	0.140	0.086	0	0.49
Panel C: Background Characteristics				
SEN	0.166	0.372	0	1
FSME	0.186	0.389	0	1
Male	0.498	0.500	0	1
Ethnicity				
White British	0.651	0.477	0	1
Other White	0.026	0.159	0	1
Asian	0.175	0.380	0	1
Black	0.081	0.273	0	1
Chinese	0.002	0.048	0	1
Mixed	0.002	0.046	0	1
Other	0.035	0.184	0	1
Unknown	0.028	0.164	0	1

Notes: 34,674 observations from the cohort who took KS2 in 2001 and KS3 in 2004. Test scores are percentalized tests scores by cohort-subject.

Table A4: Descriptive Statistics Top and Bottom Ranked Students

Panel A: Top

	National Average	Ranked in Top 5% Nationally (Age 11)	Ranked in Top 5% of Primary School (Age 11)	Self-concept Considered themselves: Very Good
Male	49.9%	49.3%	49.5%	53.5%
FSME	14.6%	4.8%	8.1%	18.5%
SEN	17.5%	2.2%	2.8%	11.2%
Minority	16.3%	13.8%	15.5%	41.1%
Obs.	6,815,997	353,464	365,176	8,192

Panel B:Bottom

	National Average	Ranked in Bottom 5% Nationally (Age 11)	Ranked in Bottom 5% of Primary School (Age 11)	Self-concept Considered themselves: Not Good
Male	49.9%	50.9%	51.5%	44.6%
FSME	14.6%	30.8%	23.7%	20.1%
SEN	17.5%	68.8%	61.4%	25.2%
Minority	16.3%	22.1%	17.9%	28.8%
Obs.	6,815,997	280,675	467,208	5,211

Notes: Data from 5 cohorts. Cohort 1 is age 11 in 2001 and age 14 in 2004, which is the only cohort we have self-concept measures for from the LSYPE dataset. Student characteristics are ethnicity, gender, free school meal (FSME) and special educational needs (SEN), minority is non-white.

Appendix Table A5: Descriptive Statistics of Self-Concept, National and Local Rank

		National Percentile Rank Age 11			Local School Rank*100 Age 11			
	Share	Mean	10th	90th	Mean	10th	90th	Obs.
How good do you think you are at…								
Panel A: …English?								
Not Good At All	1.1%	28	4	62	27	0	62	132
Not Very Good	13.5%	35	7	70	33	3	73	1563
Don't Know	0.1%	31	10	53	35	0	63	11
Fairly Good	62.5%	49	12	85	48	9	88	7222
Very Good	22.8%	62	21	95	63	20	96	2630
Panel B:…Math?								
Not Good At All	1.6%	25	3	56	22	0	56	188
Not Very Good	11.9%	31	5	62	29	2	64	1377
Don't Know	0.1%	53	12	90	56	10	93	15
Fairly Good	63.8%	47	12	85	47	9	86	7371
Very Good	22.6%	70	30	97	71	31	98	2607
Panel C:…Science?								
Not Good At All	2.1%	32	5	64	31	3	70	237
Not Very Good	14.8%	37	6	76	36	3	75	1714
Don't Know	0.2%	38	17	76	40	11	68	21
Fairly Good	57.4%	48	10	86	47	8	88	6631
Very Good	25.6%	59	17	94	60	18	95	2955

Notes: Results obtained from 11,558 student observations and 34,674 student-subject observations from LSYPE sample. Mean confidence is 0.91 with a standard deviation of 0.99.

Appendix Table A6: Age 14 Test Scores on Rank (showing controls)

	(Raw)	(1)	(2)	(3)
Panel A: The effect of primary rank				
Primary Rank	11.001**	7.960**	7.901**	4.562**
	0.298	*0.145*	*0.146*	*0.132*
Male	-0.912**	-1.007**	-0.833**	
	0.070	*0.045*	*0.021*	
Free School Meal Eligible	-6.451**	-2.962**	-2.651**	
	0.070	*0.030*	*0.027*	
Special Educational Needs	-5.148**	-4.401**	-4.308**	
	0.047	*0.033*	*0.032*	
Non-White British	1.201**	1.873**	1.526**	
	0.122	*0.525*	*0.045*	
Cohort Effects	✓	Abs	Abs	Abs
Subject Effects	✓	Abs	Abs	Abs
Cubic Key Stage 2 controls	✓	✓	✓	✓
Primary-cohort-subject Effects		✓	✓	✓
Secondary Effects			Abs	Abs
Secondary-cohort-subject Effects			✓	
Student Effects				✓

Notes: Results obtained from twelve separate regressions based on 2,271,999 student observations and 6,815,997 student-subject observations. The dependent variable is by cohort by subject percentalized KS3 test scores. All specifications control for Key Stage 2 results, student characteristics, cohort effects and subject effects. Science is the reference subjects, and the second cohort is the reference cohort. Student characteristics are ethnicity, gender, free school meal (FSME) and special educational needs (SEN). Coefficients in columns (2) and (3) are estimated using Stata command reg2hdfe allowing two high dimensional fixed effects to be absorbed. Standard errors in italics and clustered at 3,800 secondary schools. Abs indicates that the effect is absorbed by another estimated effect. ** 1% sig.

CHAPTER 3.

PAYING OUT AND CROWDING OUT? THE GLOBALISATION OF HIGHER EDUCATION

I would like to thank my co-author Steve Machin for his contribution. Moreover we acknowledge with thanks the support of the Sutton Trust for funding the initial stages of this research. We also thank the Higher Educational Statistical Agency (HESA) for the use of their data. HESA does not accept responsibility for any inferences or conclusions derived from the data by third parties. We would also like to thank participants at the RES, EALE, MECSC and PEGGED conferences for a number of helpful comments.

1 Introduction

International students bring considerable revenue to universities across the world, through their rising numbers and by paying higher tution fees than domestic students. A critical policy question is whether, because of their increasing numbers, these students take the places of natives or whether the additional income they generate acts to subsidise domestic students. This forms the subject matter of this study where we ask the question: do international students crowd out, or crowd in, domestic students to the UK higher education sector?

There have been very rapid and sizable increases in the number of international students globally, with 4.3 million students currently registered as studying abroad (OECD, 2013). Universities in the UK have been in a prime position to recruit these international students. They are generally considered to be of a high quality, with a number of universities placing very high in international world rankings.[1] Furthermore, given that English is the major *lingua franca* of business and academia, universities in English-speaking countries have a clear advantage in attracting international students. Thus, the UK ranks second, after Australia, in the percentage of enrolled university students who come from overseas (OECD, 2013).

Aggregate figures show that the total number of overseas students in UK universities has quadrupled since the 1994/1995 academic year, standing at 266 thousand full-time students by 2011/12. The postgraduate sector has seen the highest growth in overseas students in terms of proportions and absolute numbers. There are now over five times as many overseas taught postgraduates than there were in 1994/95, increasing from 28 thousand then to 140 thousand by 20011/12. Overseas typically pay higher tuition fees than domestic students and, as such, have become a major source of income for the UK HE sector, with estimates suggesting they currently contribute about 11.6 percent of the total income of the sector and 39 percent of all fee income from full time home and overseas students (HEIDI 2012), despite only making up 6% of students.

[1] For example, in Shanghai's Jiao Tong University's ranking two UK universities appear in the top 10 world rankings and five in the top 50 (and nine in the top 100).

One test of whether there is crowd in or crowd out comes from studying correlations between changes in the number of domestic students and changes in the number of international students within universities over time.[2] A negative association would correspond to crowd out (or displacement) and a positive one to crowd in (or subsidisation). This is not unlike the approach taken in the literature on immigration and the labour market where researchers look for possible displacement of native workers by immigrant flows (see, *inter alia*, Borjas, 1999, or Card and DiNardo, 2000).

However, as is the case with that literature, there are concerns related to endogenous sorting. In the case of HE it is common patterns of sorting to universities by domestic and overseas students which can render such estimates as biased. To address these concerns, in our analysis we therefore use two separate methods in attempts to identify a relationship between changes in domestic and foreign students that ensures the direction of causation flows from foreign to domestic student numbers.

The first of these has parallels with the labour economics literature on immigration where authors use the fact that immigrants from particular sending countries tend to settle in places where previous migrants from their country have settled (so called 'enclaves').[3] We adopt the same kind of exercise in terms of enrolment choices of international students. To do so, we use the historical share of students from a sending country attending a university department combined with current national changes in the stock of students from this country as a shift-share instrument to predict exogenous variations in the number of overseas students attending that university department.[4]

A second approach considers the fact that there have been very rapid increases in the number of students enrolling in UK universities from China, especially in the 2000s. We use a change in Chinese visa regulations in combination with strong revealed subject preferences and price sensitivity amongst Chinese students as a predictor of overseas student growth across and within universities over time.[5]

---

[2] This approach is adopted in some US work, for example, Borjas (2007).
[3] See Card (2005) or Card (2009).
[4] More precisely, in our empirical work below, we look at field of study and university as that is the level of analysis that our data permits. See the Data Appendix for more detail.
[5] The recent research on the effects of imports from China on the labour market and on firm productivity utilises big shifts in the Chinese share of imports to advanced countries in an analogous way (see Autor, Dorn and Hanson, 2013, and Bloom, Draca and Van Reenen, 2011).

We analyse administrative data for the entire UK population of Higher Education (HE) students over an eighteen year period (running from academic years 1994/95 through to 2011/12) which covers the time period when rapid internationalisation of the UK HE sector occurred. We document the scale of the increased supply of overseas students to the HE sector and then analyse its effects on the number of domestic students and home fee paying students using the methods outlined above.[6] The empirical analysis is carried out separately for undergraduates, taught postgraduates and research postgraduates.

Given the high policy relevance of these questions, it seems very surprising that there is hardly any research on the issue of whether there is any crowding out of natives by foreigners in university enrolments. This has not been empirically examined in the UK, and there are just a couple of research papers in the US. Borjas (2007) examines enrolment trends in US graduate programs from 1978 to 1998 and reports no average effect of foreign students on natives in graduate programs.[7] Hoxby (1998) examines if disadvantaged natives are affected by the presence of foreign students in higher education by exploiting a policy change in the fee structure in the Californian HE system. She also finds no significant effects, but does find indications that disadvantaged natives suffer a crowding out effect from immigrant students. Hoxby claims the likely mechanisms underpinning this work being through competition for affirmative action targets and financial aid.[8]

To preview our main findings, we find no evidence that the big rise in international students enrolling in UK universities has crowded out domestic students. This is the case at undergraduate level and for taught and research postgraduates. Indeed, we find

[6] Home fee students include all EU students who are those eligible for government subsidies and fee regulations and therefore face lower fees than overseas students.

[7] When extending his analysis to focus upon sub-groups, Borjas does find a significant negative effect for a subsample of native white males and demonstrates that this can neither be explained by demographics nor by a decline in demand for college places by males.

[8] In terms of using similar methodological approaches to our analysis (and those of Borjas and Hoxby), but in the very different context of compulsory schooling, Betts (1998) and Gould, Lavy and Paserman (2004) report that increased immigrant inflows have significant adverse effects on the educational outcomes of native students. This is, of course, a very different setting and in the context of HE (and opposite to compulsory schooling) in HE foreign students are more likely to come from higher socio-economic backgrounds. Other related schooling studies are Geay, McNally and Telhaj (2013) who look at non-native speakers in English schools and Ohinata and Van Ours (2013) who study immigrant children in Dutch schools.

evidence of subsidisation for postgraduates, especially on Master's programmes. Thus the answer to the question posed in the title of the paper seems to be one of crowd in rather than crowd out.

The rest of the paper is structured as follows. Section 2 offers a description of the UK higher education sector and how it has changed over time, placing a particualr references on the changing mix of domestic and foreign students. Section 3 describes the data we use and the research designs that we implement. Section 4 reports the results, while section 5 offers some conclusions.

2 The UK Higher Education Sector

2.2 Long Run Participation Trends

Figure 1 shows trends in higher education (HE) participation in the UK since the academic year 1981/82. In 1981/82 the participation rate in higher education for the appropriate age cohort (i.e. the flow of individuals participating at that time) was just over one in ten. It went up very rapidly during the 1990s expansion and has continued to rise, reaching 40 percent by the academic year 2011/12.

Thus, many more young people now attend higher education than in the past. Abstracting away from whether or not this is a good thing, and the issues to do with whether richer individuals (or those from higher social class backgrounds) benefitted more from this expansion,[9] the funding of HE changed radically over time period. The system moved from one of being broadly 'free' (i.e. non-fee paying with maintenance grants for students) to one where students pay fees and no longer receive maintenance grants, but have to take out loans to fund their education.[10]

As described in the introduction, there has also been a rapid expansion of the number of non-UK students attending UK universities. There are a number of reasons for this (which we detail below) as it has become evident that universities themselves have had become more reliant on generating more income from these international students.

[9] For discussion of the social mobility implications of HE expansion see Blanden and Machin (2004) or Lindley and Machin (2012).

[10] For more detail see Dearden, Fitzsimons, Goodman and Kaplan (2008) or Dearden, Fitzsimons and Wyness (2013).

Our focus in this paper is on the implications of this for UK students. Has the increased number of foreign students crowded out domestic students? Or has the increased income that universities receive from foreign students (typically charged at higher fees than domestic students[11]) enabled universities to take on home students and effectively crowded them in?

We study these questions using administrative data on the entire UK HE population over the 1994/95 to 2011/12 academic years. The start year is because of the creation of the so-called 'new' universities who we wish to include in our analysis and because consistently defined data from the Higher Education Statistics Agency (HESA) begins from then. Our analysis uses data covering the vast majority of students enrolled in universities over this time period, covering 161 universities in our full sample and 144 in the balanced panel who are observed in every year (more details on the data are given below and in the Data Appendix).

2.3 Trends in Student Numbers from 1994/5 to 2011/12

As already noted, the size of the UK HE sector has been rapidly growing. Table 1 shows summary statistics from our data, for the unbalanced and balanced panels of universities. As the numbers in Table 1 show, the total number of full-time students in all universities in the sample (the columns labelled 'full sample') increased from 1.06 million in the 1994/95 academic year to 1.65 million in the 2011/12 academic year. In 1994/95 there were 65 thousand international students enrolled – or around 6 percent. By 2011/12 this had risen to 246 thousand, or around 15 percent of all (full-time) students.

Sharp increases in the relative numbers of international students have occurred at both undergraduate level and at postgraduate level. Figure 2 shows trends over time (where the numbers of each are indexed at 1 in 1994/95) in the numbers of domestic and foreign undergraduates, taught postgraduates and research postgraduates. The relative

[11] Average international fees for an undergraduate (postgraduate) course were £9,360 (£9,520) in 2009/10 (Murphy, 2014). Comparatively for domestic undergraduates universities received £3,000 in fees and a mínimum of £3,947 in subsidies from the Government, dependent on subject and location.

increase is clear for all three, but is especially marked amongst full-time taught postgraduates.

2.4 Rules on University Admissions

It is worth noting that because of government funding of places for home undergraduates and taught postgraduates there is a constraint on growth during this time period due to government quotas. During this period universities educating these students received funding from the *Higher Education Funding Council for England* (HEFCE), who set a Student Control Number (SCN) dictating the maximum number of home fee students allowed to be enrolled by each university.[12] Universities who took on more home students than they were allocated were issued a fine per student. At the same time universities were allowed to under-recruit by up to 5 percent and still get funding for the full amount, which obviously acted to increase per-pupil funding. Universities could choose not to count all taught postgraduate students against its Student Control Number, foregoing any government funding for these students.[13]

At the same time, universities were allowed to bid for Additional Student Numbers (ASN) to increase its SCN and its teaching grant accordingly. The way in which HEFCE awarded these ASNs changed over time. Prior to 2000/01 they were allocated according to government plans for growth of student numbers by region and subject area or for specific projects. From 2000/01, this was amended so that institutions could submit proposals for ASN. For an application to be successful an institution needed to have filled its existing student places, show excellence and provide evidence that there is demand for additional places. One feature of this was for universities to use overseas students as a signal of this demand. This potentially allowed the number of overseas students to influence the number of domestic students, even with government quotas. From 2005 the government suspended competitive ASN bids and once again allocated additional places according to specific developments or goals.

[12] Government funding was dependent on the location of the university and the cost of the subject taught. There are four subject categories which are given cost weightings; Class D courses involves only lectures have weight 1; Class C courses have a fieldwork or studio element have weight 1.3; Class B courses are laboratory-based subjects and have a weight of 1.7; Class A courses are medicine and dentistry and have a weight of 4.

[13] To do this the fees charged for the course need to exceed the teaching subsidy per student (£3,951) plus an assumed fee (£3,591).

Funding for research postgraduate students was not subject to these caps. Funding for these students is allocated in proportion to the number of home research postgraduates in their first 3 years of full time study (6 years for part time) in departments rated 4+ in the previous Research Assessment Exercise (RAE), weighted by London residence and subject costs. This is unrestricted funding with no caps on the maximum amounts of students at a university, however the total amount of money is capped and split amongst institutions.

By contrast to these regulations on home students, overseas students receive no government subsidies, and therefore have no limits to the number enrolled by individual institutions. The only limiting factor is the number of overseas student visas approved by the UK immigration office and from the sending country.

Hence the very large growth in the number of overseas students documented in Figure 2 and Table 1. By the 2011/12 academic year, there were 115 percent more research postgraduates, 229 percent more undergraduates and 547 percent more taught postgraduates from overseas as compared to 1994/95.

In addition to domestic students and overseas students are non-UK European Union (EU) students. EU regulations mean that all students domiciled in the EU are required to be treated in the same way. This means that all UK and EU students pay the same fees and are referred to as home students. A further consequence of this is that EU students receive the same funding as an equivalent domestic students and it is for this reason that they come under the same quota system as domestic students.[14]

[14] Therefore non-UK EU students displace domestic students on a one-for-one basis. In our empirical work, we thus show results treating UK only and UK plus EU students as domestic students and home fee students. As one would expect, given that from a university funding perspective the students are equivalent, results prove to be rather similar. In 2011/2012 non-UK EU students comprised 4 percent of home fee undergraduates, percent of taught postgraduates and 21 percent of research postgraduates (see Table 1).

2.5 Origin Countries of Overseas Students

The composition of overseas students has remained relatively stable over this time period by broad region of origin.[15] Figure 3 shows Asia to be by far the largest source region of students, followed by Africa and then North America.

However, as Figure 3 also shows, there has been one large and notable change in the composition of overseas students, namely the influx of Chinese students. In 1995 there were 1,510 Chinese students amongst all UK universities studying full time at any level. This remained fairly stable until 1998, after which it began where it began to increase rapidly, almost doubling year on year between 2000 and 2003. By 2005 there were 39,820 Chinese students, corresponding to an enormous 1,900 percent increase over a seven year period. Again this remained stable until 2009 when the number began growing quickly, the Chinese now account for more 4 percent of all students and 26 percent of all overseas students.

The rapid expansion and subsequent levelling off in these numbers was caused by a change in the Chinese visa licencing for students. In 1999 the Chinese Government introduced new regulations[16] which allowed for the formations of licencing agencies making it considerably easier for self-funded Chinese students to study abroad. Although it had been possible to self-fund since 1981[17], international study was still characterised by around 5000 government funded students from the leading universities being sent to strategically productive placements abroad. The opening of these agencies dramatically increased the size of the self-funded sector, in 1998 there were 11,443 self-funded students, by 2002 there were 117,000 (Li and Zhang, 2010).

The self-funded nature of Chinese students has meant they are very concentrated in certain fields of study. Figures A1 and A2 in the Appendix clearly show these strong subject preferences. The subjects with the largest increase in numbers were Business & Management, Maths & Computing, Economics and Engineering, whilst there is very little growth in the remaining subjects. Furthermore given that Chinese students became

[15] Countries were grouped to the following broad regions using NSCC groupings; Africa, Asia, Europe (EU and Non EU), Middle East, North America and the Rest of the World.

[16] Regulations for the Administration of Intermediary Agencies for Self-Funded Study Abroad 1999 (PRC).

[17] State Council – Interim Provisions for Study Abroad with Self-funding.

predominately self-funded they also became more price responsive. Prior to 1998 the demand of student places by Chinese students was unlikely to be responsive to the British Pound: Yuan exchange rate. After then the exchange rate became potentially more important as a determinant of the number of Chinese students, a feature we exploit in our empirical strategy below.

3 Data and Research Designs

3.2 Data Description

The administrative data we use comes from the Higher Education Statistical Agency (HESA) and contains information on all full time students studying at higher education institutions between the academic years 1994/95 and 2011/12, comprising 18.6 million individuals in total. We have count data for universities broken down across the following groups: 165 subject areas, 267 domiciles of origin, 3 levels of study, 3 fee statuses and 2 genders over 18 years.[18]

We conduct separate analyses for undergraduates, taught postgraduates and research postgraduates. To eliminate the issue of universities opening and closing (which could result in spurious results) we use a balanced panel of universities. This is defined as including those with a positive student count for each of the levels (UG, PGT, and PGR) in all of the years. This brings our sample to 144 institutions.[19] Summary statistics are shown in Table 1 (in the columns labelled 'balanced panel'). It is evident that our sample contains the vast majority of students as compared to the full sample (where there were 149 universities in 1994/95 and 161 in 2011/12), suggesting any entrants or exit are small.

There are many university courses on offer at UK universities. We have data on 165 distinct fields of study categories. During the early years of our data, the number of overseas students was relatively small in some universities and so therefore to ensure that there were sufficient non-zero shares of students from countries, we aggregated groupings of related subjects. The 165 subjects are grouped into 5 subject areas; (1)

[18] So, as an example, in a given academic year we can calculate the number of male French students at Oxford paying home fees who are studying physics at undergraduate level.

[19] For the 41universities that merge during the time period, we consider them as one university throughout our sample period.

Medicine, dentistry and subjects allied to medicine; (2) Sciences; (3) Social Sciences, Law and Business, (4) English, Languages and History; and (5) Creative arts, design and education.

Table 1 shows student numbers in each of these five groups, and also in twelve smaller groups which we can look at for some of our analysis (the part on Chinese student inflows) when we focus on a university panel rather than a field of study by university panel.[20] More details on these, and other definitional aspects of the data, are given in the Data Description in the Appendix.

3.3 Modelling Approach

Our initial research set up borrows from the related literature (in terms of approach) that studies the impacts of immigrant inflows on native outcomes. The most well-known work in this area studies the impact of inflows on labour market outcomes, but there are also studies looking at other outcomes like crime, use of public services and education.[21]

When considering the impact of immigrant flows on native outcomes, various authors (like Borjas, 2006, and Card, 2007) have set up empirical specifications to net out problems to do with initial conditions and mechanical biases. Peri and Sparber (2011) summarise these arguments and claim that, in the context of spatial variations across cities, the best representation relates changes in native or immigrant outcomes (employment in their case) that are scaled by the lagged size of their spatial unit (the city). Our analogous outcomes are domestic and international student numbers at study field by university level, so we develop a baseline estimating equation of the following form for subject by university i in year t:

$$(D_{it} - D_{i,t-1})/S_{i,t-1} = \alpha_i + \beta\,(F_{it} - F_{i,t-1})/S_{i,t-1} + T_t + \varepsilon_{it} \qquad (1)$$

where i denotes field of study by university, t denotes year, D is the number of domestic students, F is the number of international students, so that S = D + F, and the equation includes a full set of field of study by university fixed effects (α_i) and time effects (T_t)

[20] Results using 12 subject areas instead of the 5 broad areas are available from the authors upon request. They are similar, but because of a higher preponderance of zeros the first stage results were not as strong.

[21] Examples of studies of crime and immigration are, *inter alia*, Bell, Fasani and Machin (2013), on public services and immigration see Wadsworth (2013) and on education and migraton see Dustmann and Glitz (2011).

and an error term (ε_{it}). As the model is specified in changes it accounts for underlying differences across universities, with α_i controlling for average growth rates by university field of study and T_t accounting for annual aggregate growth rates. Therefore, the identifying variation comes from deviations of growth rates from university field of study growth trends.

In (1) β is the key parameter of interest for whether or not there is a crowding in or out within a university. A positive β is suggestive of subsidisation, whilst a negative β implies displacement. This coefficient can be interpreted as the number of domestic students who respond to each additional overseas student (e.g. a coefficient of -1 implies one-for-one crowding out). These estimates are not affected by cell size, nor is there any artificial correlation between the dependent and explanatory variable.

Whilst equation (1) is quite stringent in that it specifies the relationship in terms of within field of study by university changes and includes a full set of fixed effects, it does not (unless these fixed effects factor out any possible bias) account for the potential endogeneity of overseas demand. This is important as universities that experience shocks, such as changes in university rankings or new teaching buildings, may affect the supply or demand of places for domestic and overseas students simultaneously. To address this issue of common unobserved shocks, we use instrumental variable techniques to generate an exogenous source of variation in the number of overseas students at university subject area level.

We adopt two approaches to do this. The first employs the shift share approach that has been commonly used in the immigration literature (e.g. by Card, 2009). This approach relies on prior immigrant settlement patterns as a source of identifying information. The idea rests on the notion that the current relative flow of immigrants to a city is related to historical population shares. The thought experiment is that a city with an historically high share of the immigrants from a particular source country, is more likely to experience growth when the national amount of immigrants from that source country increases, compared to a city with a low historical share. The key assumption is that the national inflow rates from each source country are exogenous to conditions of any city.

When thinking of university enrolments, the conceptual analogue is that individuals from a particular origin country are more likely to go to universities, and study subjects, where previous students attended. Anecdotally, this seems reasonable in that there are well known examples of students from particular countries studying the same kind of degrees in particular countries. Obviously it is an empirical question as to how strongly the instrument predicts.

More formally, the instrumental variable we use to predict the change in the share of foreign students for field of study by university time is the following:

$$\Delta P_{it} = \sum_{c=1}^{n} (F_{cit0}/F_{ct0})\Delta F_{ct} \tag{2}$$

where we use the initial distribution of foreign students from country c and allocate the flow of foreign students from that country between period t and t+1, according to that distribution in time 0 and the total change in students from country c. We do this for 1994/95 to 1998/99 as the initial time period and predict future annual flows (2001/02 to 2011/12) to each university subject area. This means there is year on year by university subject area variation generated from a combination of national inflow figures and the historical shares.

From this we generate two instrumental variables. The first, IV1, groups all Non-EU countries together as one category. In this case c is an indicator for originating from non-home fee paying country. The second set, IV2, uses the shares on a country basis, so c represents the country of origin. IV2 has the benefit that, because it uses the proportion of students from an individual country, it allows for specific country university subject relationships. For example, if the science department at University A has a higher than average proportion of the national total of students from Non-EU country Z, then when the total amount of students from country Z increases we expect a larger than average increase in the number of overseas science students at University A. Furthermore if another Non-EU country Y which has no students at university A had an increase in their student numbers, this would not affect the number of overseas students at university A. Contrary to this method IV1 would use the share of Non-EU students as a whole, and any increase in the number of students from country Y at the national level would be in part allocated to university A.

The disadvantage of this method is that because IV2 uses specific country university subject area pairs, the proportion may be more liable to change over time. Even though these shares were generated over a four year period between 1995-8, some countries had little to no representations in some cells that have subsequently may have become more prominent. So any future increases in their numbers would not be reflected in predicted changes at that university subject level. The benefit of IV1 is that it captures the international character of a university, and therefore does not rely on specific countries to be present at a university during the 1995-8 period. The key assumption remains that the national inflow rates from each source country are exogenous to conditions in a particular university department. This is likely to be the case given that we have 570 university subject cells, each only contributing a small amount to the total.

The shift share approach assumes that historical shares of students at university departments are informative about current flows. It turns out that this is typically true as our results that the location patterns for past students are strongly reflective of where later students chose to study. However, for reasons we have already discussed, this is not likely to hold for China, which is an important country during this period of expansion. Prior to 1998 the majority of Chinese students were granted a student visa if their choice of course was supported by the state. However, once self-sponsorship became easier, the numbers of Chinese students studying Business or Economics increased very rapidly indeed.

We thus focus on inflows of Chinese students in detail and implement an alternative IV strategy which uses the Chinese policy change as a source of exogenous variation in the change in number of Chinese students attending UK universities. After the simplification of the Chinese student visa application process the number of overseas students studying Business or Economics rapidly increased.

This increase is shown very clearly in Figure 4 which shows big increases, for these two growth subjects. Figure 4 also (like the earlier Figure 3) shows that the majority of the growth occurs between 1999/2000 and 2004/05. Then, after a lull, a second strong growth phase occurred from 2008/09 to 2011/12. Interestingly, this occurred after the Chinese Yuan huge 80 percent appreciation against the Great British Pound during 2008, as is shown in Figure A3 of the Appendix.

It is evident that, with the increase in self sponsored students the number of Chinese students attending UK universities became increasingly dependent on the exchange rate. Table 2 shows the price sensitivity of Chinese students before and after the visa reform. The results show that changes in the number of Chinese students were uncorrelated with the exchange rate prior to the reform (when they were predominantly funded by the Chinese Government), but were significantly correlated post reform. Moreover, it is the growth subjects are the most sensitive to exchange rate fluctuations.

Therefore to implement an instrumental variable strategy in the study of changing Chinese student numbers we generate an indicator variable for each growth subject post reform, plus a growth indicator interacted with the Yuan:Pound exchange rate. We argue that these sources of overseas student growth are exogenous to university departments in the UK and so we use these combinations as instruments for the change in the number of overseas students at a university department.

4 Results

4.2 University by Field of Study Panel

Table 3 shows estimates of equation (1) for undergraduates, taught postgraduates and research postgraduates. Two specifications are reported for each, one where the dependent variable is the proportional change in the number of home students (i.e. UK and EU students) and one where it is the proportional change in the number of domestic students (i.e. UK only). The first row of Table 3 shows ordinary least squares (OLS) estimates, and the second row shows the two stage least squares (2SLS) estimates. The first stages corresponding to the latter (which are the same for the home and domestic specifications) are reported below these, together with associated F-statistics for the instruments.

The first thing to note from the Table is that all the OLS coefficients are estimated to be positive, and statistically significant, implying no evidence of crowd out. However, for the reasons articulated above, we need to consider what happens when we allow for common shocks to affect both changes in domestic and foreign students via our instrumental variable strategy. The F-tests reported for the first stage show that the instruments are very good predictors of the change in foreign students (they are all above 10, and some strongly so). Thus transposing over the enclave idea that has been exploited in the immigration literature to the inflow of foreign students to specific fields of study and university seems to work well. The positive enclave effect is intuitively very plausible (i.e. that foreign students go to study the same subjects in the same universities as previous students from their home country did). Interestingly, the 2SLS estimates are, like the OLS estimates, all positive as well. Thus our evidence is much more in line with the notion of crowding in, where foreign students bring in additional income that can cross-subsidise domestic students, rather than crowding out.

The pattern across the three groups of students is informative in this regard as well. The 2SLS estimates are significant and positive for the postgraduate groups, but not for the undergraduates. Thus for undergraduates, there is on average no crowd out or crowd in, but for postgraduates where universities have freedom to increase the number of home students and where fees can be sizable for foreign students, there is evidence of crowding in. The coefficient above unity for the taught postgraduates is suggestive that Master's courses are the major place where this occurs.

Table 4 reports separate estimates for the 20 Russell Group and 124 non-Russell Group universities in our sample.[22] The evidence of crowd in seems to be more marked for the former group. This is not so surprising given their ability to recruit international students and that these top universities charge higher tuition fees and, in doing so, generate a significant income stream. Again there is little significant evidence of crowd in of undergraduate students, which is reflective of government restrictions on undergraduate numbers.

4.3 Increases in Chinese Students

We next move to the analysis of Chinese student inflows. To do so the estimating equation is structured as before, but the key independent variable of interest becomes changes in the number of Chinese students, C:

$$(D_{it} - D_{i,t-1})/S_{i,t-1} = \alpha_i + \beta\,(C_{it} - C_{i,t-1})/S_{i,t-1} + T_t + \varepsilon_{it} \qquad (3)$$

[22] The full set of results, structured in the same way as Table 3 for the Russell Group and non-Russell Group universities are reported in Tables A2 and A3 in the appendix.

Table 5 shows the estimates of equation (3). The Table is structured in a similar way to Table 3, with OLS estimates in the first row, 2SLS estimates in the second and the first stages from the latter below. As with the panel analysis above, the OLS estimates all show positive estimated coefficients. Also in line with those results, the first stages using the reform/exchange rate instruments are strong, with one exception the research postgraduates.

The 2SLS estimates from the undergraduate and taught postgraduate regressions are both positive, although the undergraduate estimate is imprecisely determined. The taught postgraduates estimate is of similar magnitude to the earlier results and offers strong evidence of crowd in. Thus it seems that increased enrolment of Chinese students on Master's courses has become an important factor in generating income streams for UK universities that have also enabled universities to take on more domestic students in these subjects.

The estimated 2SLS coefficient for research postgraduates is the only negative one we have been able to uncover in our entire empirical analysis. It is, however, close to zero and, owing to the weak first stage, is very poorly determined. This is likely due to the relatively small increases in the number of Chinese students undertaking research degrees. Again, this offers no evidence whatsoever for the hypothesis that foreign students have crowded out domestic students in UK universities. Thus both of our causal approaches reach the same conclusion.

5 Conclusions

In this paper we study the rapid inflows of international students to UK universities, asking the questions as to whether their increased enrolment and paying out of high fees has had any impact on the enrolment of domestic students. We frame this as a question of whether one can detect any evidence that their increased numbers have displaced domestic students or whether their increased numbers have gone hand-in-hand with increased numbers of domestic students.

To properly consider this question, it is important to set up a research design that allows for common shocks that could cause numbers of domestic and foreign students to covary with one another. We do this in two ways. First, in a manner similar to that adopted in the immigration literature, we use the historical share of students from a sending country attending a university department combined with current national changes in the stock of students from this country as a shift-share instrument. Secondly, we use an exogenous change in the Chinese visa regulations and exchange rate in combination with strong revealed subject preferences as a predictor of overseas student growth.

Using administrative data on the entire UK HE population over the 1994/5 to 2011/12 academic years, in both of these approaches we find no evidence that the big rise in international students enrolling in UK universities has crowded out domestic students. This is the case at undergraduate level and for taught and research postgraduates. Indeed, in some cases we find evidence of subsidisation for postgraduates, especially on Master's programmes where numbers of domestic and foreign students have covaried positively with one another as both have increased significantly through time. For undergraduates, there is little evidence of crowding in or crowding out. This is suggestive evidence that the government quotas on the number of domestic students has impeded growth. Despite this, the increased number of overseas students combined with their higher than average fees would still have resulted in higher funding per domestic student.

References

Autor, D., Dorn, D. and Hanson, G. (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States, *American Economic Review*, 103(6), pp.2121-68.

Bartlett A., Preston, B. and Webb, S. (1981). Overseas Students And British Commercial Interests, in P. Williams ed. *The Overseas Student Question – Studies for a Policy*, Heinemann Educational Books/Overseas Students Trust: London.

Bell, B., Fasani, F. and Machin S. (2013). Crime and Immigration: Evidence from Large Immigrant Waves, *Review of Economics and Statistics*, 95(4), pp. 1278-1290

Betts, J. 1998. Educational Crowding Out: Do Immigrants Affect The Educational Attainment Of American Minorities?, in D. Hamermesh and F. Bean eds., *Help or Hindrance? The Economic Implications of Immigration for African-Americans*, Russell Sage Foundation: New York.

Blaug, M. (1981). The Economic Costs And Benefits Of Overseas Students, in P. Williams ed. The Overseas Student Question – Studies for a Policy, Heinemann Educational Books/Overseas Students Trust: London

Blanden, J. and Machin S. (2004). Educational Inequality and the Expansion of UK Higher Education, *Scottish Journal of Political Economy*, 51, 230-49.

Bloom, N., M. Draca and J. Van Reenen (2011). Trade Induced Technical Change: The Impact of Chinese Imports on Innovation, Diffusion and Productivity, *National Bureau of Economic Research*, Working Paper 16717.

Borjas, G. (1999). The Economic Analysis of Immigration, Chapter 28 in O. Ashenfelter and D. Card eds., *Handbook of Labor Economics*, Vol. 3, Amsterdam: North Holland, pp.1761-1800.

Borjas, G. (2005). Do Foreign Students Crowd Out Natives From Graduate Programs?, in R. Ehrenberg and P. Stephan eds. Science and the University, University of Wisconsin Press: Madison.

Bowden, R. and Watson D. (1999). Why Did They Do It?: The Conservatives And Mass Higher Education, 1979-97, *Journal of Education Policy*, 14, pp. 243-256.

Card, D. (1990). The Impact Of The Mariel Boatlift On The Miami Labor Market, *Industrial and Labor Relations Review*, 43, pp.245-257.

Card, D. (2005). Is The New Immigration Really So Bad?, *Economic Journal*, 115, F300-F323.

Card, D. (2009). Immigration and Inequality, *American Economic Review*, 99, pp. 1-21.

Card, D. and DiNardo J. (2000). Do Immigrant Inflows Lead To Native Outflows?, *American Economic Review*, 90, pp. 360-367.

Dearden, L., Fitzsimons, E., Goodman A., and Kaplan G. (2008). Higher Education Funding Reforms in England: The Distributional Effects and the Shifting Balance of Costs, *Economic Journal*, 118, F100–F125.

Dearden, L., Fitzsimons E. and Wyness G. (2013). Money for Nothing: Estimating the Impact of Student aid on Participation in Higher Education, *Department of Quantitative Social Science* Working Paper No. 13, Institute of Education, London.

Dustmann, C. and Glitz A. (2013). Migration and Education, in E. Hanushek, S. Machin and L. Woessmann eds. *Handbook of the Economics of Education*, Volume 4, Elsevier: Amsterdam.

Geay, C. McNally S. and Telhaj S. 2013. Non-native Speakers of English in the Classroom: What are the Effects on Pupil Performance?, *Economic Journal*, 123, F281-F307.

Gould, E., Lavy V. and Paserman D. (2009). Does Immigration Affect The Long Term Educational Outcomes Of Natives? Quasi-Experimental Evidence, *Economic Journal*, 119, pp.1243-69.

HESA (2010). Statistics - Finances of UK HE institutions - Table E, Higher Education Statistics Agency: Cheltenham.

HEIDI (2012). Higher Education Information Database for Institutions, Higher Education Statistics Agency: Cheltenham.

Hoxby, C. (1998). Do Immigrants Crowd Disadvantaged American Natives Out of Higher Education?, in D. Hamermesh and F. Bean eds., *Help or Hindrance? The Economic Implications of Immigration for African-Americans*, Russell Sage Foundation: New York.

Li, M. and Zhang Y. (2010). Two-Way Flows Of Higher Education Students In Mainland China In A Global Market: Trends, Characteristics and Problems, Higher Education in The Asia-Pacific, 36, pp.309-327.

Lindley, J. and Machin S. 2012. The Quest For More and More Education: Implications for Social Mobility, Fiscal Studies, 33, pp.265-86.

Murphy, R. (2014). The Causes and Consequences of the Growth in International Students in Higher Education, in Gerard, M. and S. Übelmesser eds.. *The Mobility of Students and the Highly Skilled*, MIT Press.

OECD (2013). Education at a Glance 2013: OECD Indicators, OECD: Paris.

Ohinata, A. and Van Ours, J. (2013). How Immigrant Children Affect the Academic Achievement of Native Dutch Children, *Economic Journal*, 123, F308–F331.

Peri, G. and Sparber C. (2011). Assessing Inherent Model Bias: An Application To Native Displacement In Response To Immigration, *Journal of Urban Economics*, 69, pp.82–91.

Wadsworth, J. (2013). Mustn't Grumble. Immigration, Health and Health Service Use in the UK and Germany, *Fiscal Studies*, 34, pp.55-82.

Table 1: Descriptive Statistics of Student Attendance

	Full Sample			Balanced Sample		
	1994/5	2011/12	Percentage Change	1994/5	2011/12	Percentage Change
Undergraduate Students						
Home	891.2	1240.7	39.2	873.2	1214.4	39.1
Domestic	854.3	1165.7	36.5	836.9	1140.9	36.3
Overseas	37.6	123.5	228.8	37.1	121.7	228.4
Total	928.8	1372.8	47.8	910.2	1343.2	47.6
Taught Postgraduate						
Home	69.1	101.2	46.5	68.0	99.3	46.2
Domestic	59.1	77.7	31.3	58.1	76.1	30.9
Overseas	15.0	96.9	547.1	14.7	95.4	549.9
Total	84.1	199.9	137.8	82.6	196.1	137.4
Research Postgraduate						
Home	32.7	48.0	46.7	32.1	47.1	46.8
Domestic	27.9	37.3	33.7	27.4	36.6	33.6
Overseas	12.0	25.8	114.9	11.9	25.4	114.0
Total	44.7	74.0	65.5	44.0	72.7	65.3
Subject Areas						
Medical	172.4	387.5	124.8	169.9	384.4	126.2
Science	283.9	339.8	19.7	279.8	335.9	20.1
Social Science	272.9	465.5	70.6	266.3	460.1	72.8
Languages/History	140.8	206.1	46.3	137.6	201.5	46.5
Arts & Other	187.5	247.9	32.2	183.2	246.5	34.5
Medicine & Dentistry	96.8	215.9	123.0	94.7	214.7	126.7
Biology & Veterinary	75.6	171.1	126.3	75.2	169.5	125.4
Physical Sciences	66.1	73.6	11.3	65.1	73.1	12.3
Maths & Computing	76.5	105.4	37.8	74.5	104.7	40.5
Engineering	95.7	109.1	14.0	93.8	106.9	14.0
Architecture & Technology	46.5	51.3	10.3	46.4	50.5	8.8
Law & Social Studies	120.8	193	59.8	117.8	190.8	62.0
Economics	22.7	34.7	52.9	21.7	33.9	56.2
Business & Management	127.8	237.6	85.9	126.8	235.6	85.8
Language & Humanities	140.8	206.1	46.4	137.6	201.5	46.4
Education & Creative Arts	138.5	241.9	74.7	136.5	241.1	76.6
Other & Combined	48.7	6	-87.7	46.7	5.4	-88.4
Total Overseas Students	64.6	246.2	381.1	63.7	242.5	380.7
Total Students	1057.6	1646.7	155.7	1036.8	1612.0	155.5
Number of Universities	149	161		144	144	

Notes: Totals shown in 1000s. *Source*: HESA administrative data of full time students at UK HE institutions

Table 2: Changes in The Sensitivity of Chinese Students to Pound-Yuan Exchange Rate, Pre- and Post-Reform

	Undergraduates			**Taught Postgraduates**			**Research Postgraduates**		
	Pre-Reform	Post-Reform	Change	Pre-Reform	Post-Reform	Change	Pre-Reform	Post-Reform	Change
Business	0.118	0.496	0.378	-1.280	2.050	3.330	-0.827	-0.561	0.266
	0.159	*0.075*	*0.176*	*0.670*	*0.361*	*0.761*	*1.064*	*0.262*	*1.096*
	[300]	[1400]		[301]	[1600]		[281]	[1424]	
Economics	0.016	0.414	0.398	0.102	2.185	2.287	-0.949	-0.503	0.447
	0.150	*0.116*	*0.189*	*1.043*	*1.521*	*1.107*	*1.271*	*0.416*	*1.337*
	[183]	[854]		[160]	[713]		[194]	[836]	
All Subjects	-0.021	0.113	0.134	-0.389	1.000	1.390	0.028	0.014	-0.017
	0.030	*0.016*	*0.034*	*0.148*	*0.099*	*0.178*	*0.210*	*0.034*	*0.213*
	[3156]	[14728]		[2905]	[14898]		[3026]	[14291]	

Notes: This table presents the estimates from 18 regressions of the normalized change in Chinese students on the exchange rate, for undergraduates, taught postgraduates and research post graduates. A further nine regressions are in the Change columns to estimate the change in this relationship.. Robust standard errors in italics. Numbers of students in square brackets. Regressions are weighted by the appropriate mean of the student populations over the differenced years. Source: British Pound Sterling and Chinese Yuan exchange rate from the International Monetary Fund

Table 3: University by Field of Study (5) Panel Estimates

Estimates of

$$(D_{it} - D_{i,t-1})/S_{i,t-1} = \alpha_i + \beta\,(F_{it} - F_{i,t-1})/S_{i,t-1} + T_t + \varepsilon_{it}$$

	Undergraduates		**Taught Postgraduates**		**Research Postgraduates**	
	Home	Domestic	Home	Domestic	Home	Domestic
Ordinary Least Squares:						
Change in Foreign	0.772	0.702	0.581	0.510	0.938	0.679
Students	*0.297*	*0.288*	*0.286*	*0.257*	*0.252*	*0.157*
Two Stage Least Squares: Squares:						
Change in Foreign	0.060	0.053	1.442	1.268	0.949	0.763
Students	*0.364*	*0.341*	*0.889*	*0.802*	*0.139*	*0.144*
First Stage: IV1	0.182		1.052		1.465	
	0.159		*0.288*		*0.305*	
First Stage: IV2	0.496		-0.199		0.462	
	0.153		*0.170*		*0.167*	
F-Test	23.04		11.32		26.86	
Sample Size	7,444	7,444	6,945	6,945	6,514	6,514
Number of	144	144	144	144	144	144

Notes: All regressions include, year and subject-institution fixed effects, with robust standard errors in *italics*. Regressions are weighted by the appropriate mean of the student populations over the differenced years. 2SLS F statistic is based on the Kleinbergen-Paap Wald F statistic, allowing for non iid errors.

Table 4: University by Field of Study (5) Panel Estimates, Russell Group and Non-Russell Group Universities

Estimates of

$$(D_{it} - D_{i,t-1})/S_{i,t-1} = \alpha_i + \beta\,(F_{it} - F_{i,t-1})/S_{i,t-1} + T_t + \varepsilon_{it}$$

	Undergraduates		**Taught Postgraduates**		**Research Postgraduates**	
	Home	Domestic	Home	Domestic	Home	Domestic
Panel A. Russell Group						
Two Stage Least Squares:						
Change in Foreign	2.241	2.140	2.286	2.092	0.866	0.683
Students	*1.273*	*1.304*	*0.653*	*0.628*	*0.190*	*0.162*
F-Test for First Stage	4.99		7.44		21.69	
Sample Size	1167	1167	1179	1179	1172	1172
Number of Universities	20	20	20	20	20	20
Panel B. Non-Russell						
Two Stage Least Squares:						
Change in Foreign	-0.034	-0.030	0.397	0.335	0.996	0.804
Students	*0.396*	*0.371*	*0.155*	*0.138*	*0.177*	*0.145*
F-Test for First Stage	19.64		37.62		16.55	
Sample Size	6,277	6,277	5,342	5,342	5,766	5,766
Number of Universities	124	124	124	124	124	124

Notes: All regressions include, year and subject-institution fixed effects, with robust standard errors in brackets. Regressions are weighted by the appropriate mean of the student populations over the differenced years. 2SLS F statistic is based on the Kleinbergen-Paap Wald F statistic, allowing for non iid errors.

Table 5: Chinese Students, University Panel by Field of Study (11) Estimates

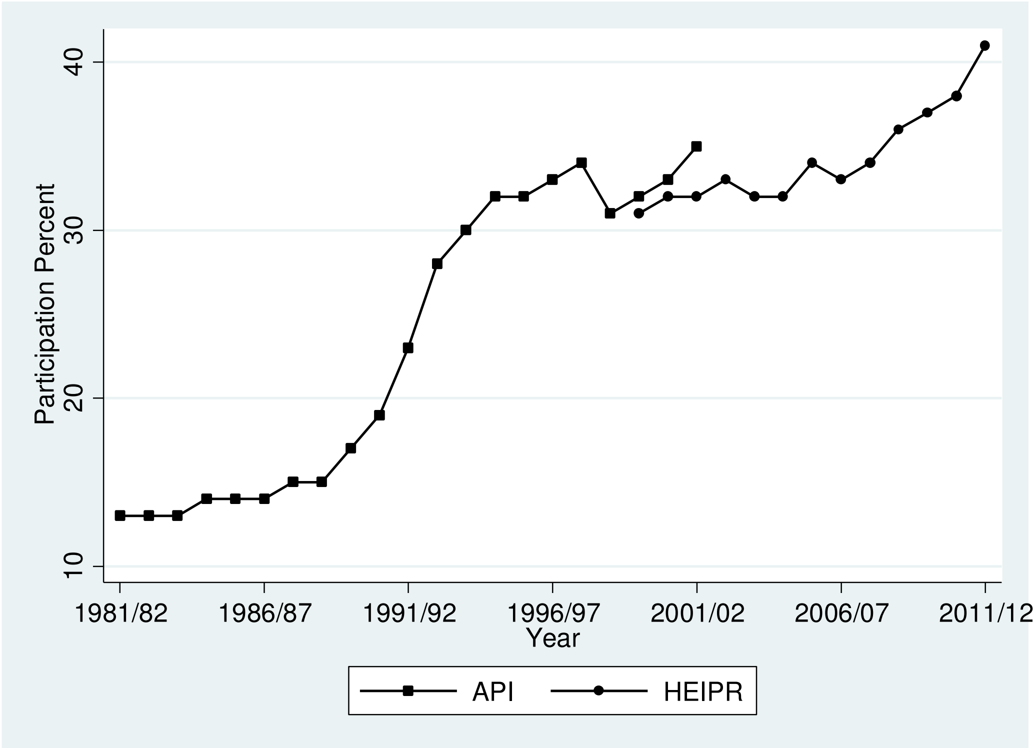
Estimates of

$$(D_{it} - D_{i,t-1})/S_{i,t-1} = \alpha_i + \beta\,(C_{it} - C_{i,t-1})/S_{i,t-1} + T_t + \varepsilon_{it}$$

	Undergraduates		**Taught Postgraduates**		**Research Postgraduates**	
	Home	Domestic	Home	Domestic	Home	Domestic
Ordinary Least Squares:						
Change in Chinese	0.487	0.540	0.388	0.577	1.111	1.587
Students	*0.138*	*0.144*	*0.081*	*0.104*	*0.279*	*0.342*
Two Stage Least Squares:						
Change in Chinese	0.815	0.584	1.538	1.705	-0.112	-0.374
	0.586	*0.601*	*0.828*	*0.930*	*0.587*	*0.958*
First Stage:						
Business X Reform	-0.029		-0.088		0.050	
	0.006		*0.028*		*0.021*	
Business X Reform	0.453		1.490		-0.602	
X Exchange Rate	*0.074*		*0.352*		*0.254*	
Economics X Reform	-0.023		-0.091		0.048	
	0.009		*0.082*		*0.031*	
Economics X Reform	0.367		1.746		-0.558	
X Exchange Rate	*0.111*		*1.055*		*0.397*	
F-Test	29.03		15.93		2.36	
Sample Size	17,663	17,663	17,386	17,386	17,026	17,026
Number of Universities	144	144	144	144	144	144

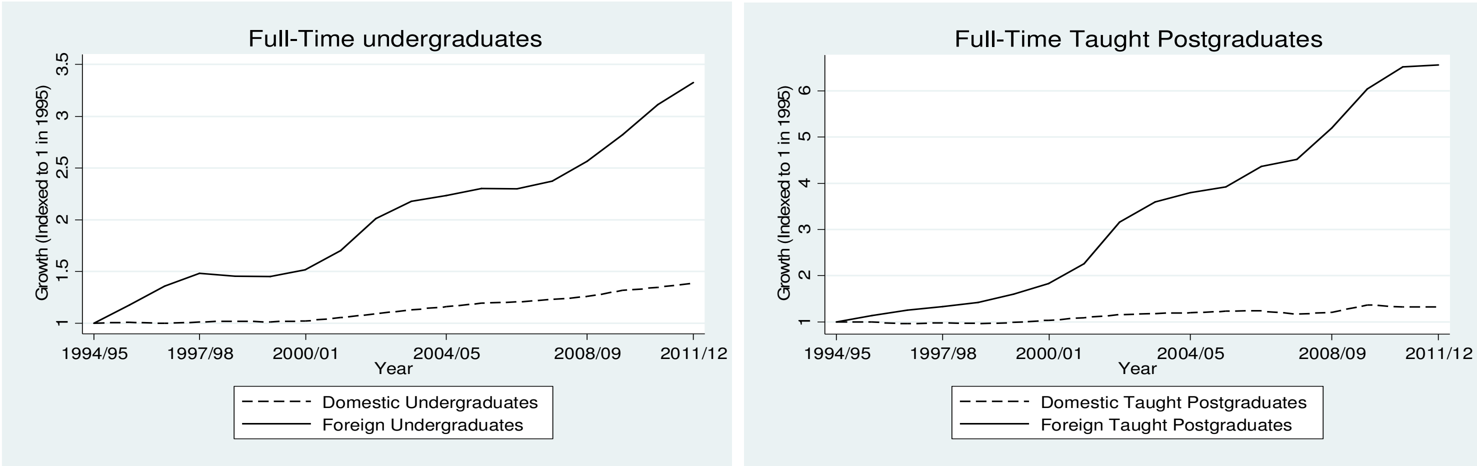
Notes: All regressions include, year and subject-institution fixed effects, with robust standard errors in brackets. Regressions are weighted by the appropriate mean of the student populations over the differenced years. 2SLS F statistic is based on the Kleinbergen-Paap Wald F statistic, allowing for non iid errors.

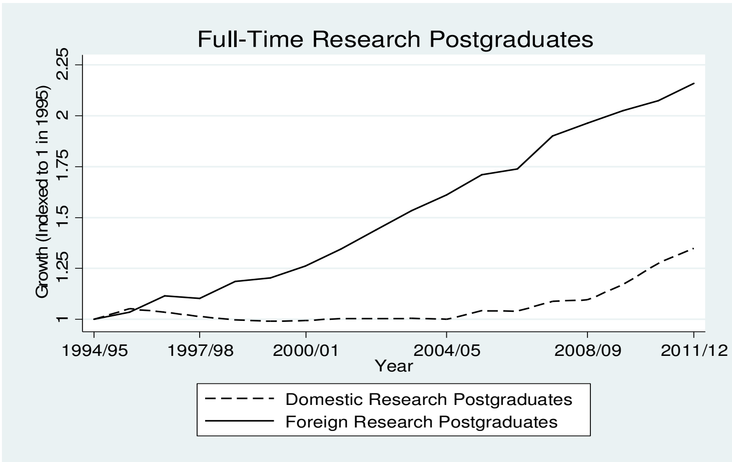
Figure 1 Trends in UK Higher Education Participation



Notes: The Age Participation Index (API) is the number of domiciled young people (aged less than 21) who are initial entrants to full time and sandwich undergraduate courses as a percentage of the 18 to 19 year old GB population. The API was discontinued in 2001 and replaced by the Higher Education Initial Participation Rate (HEIPR), which has a different definition as it covers entrants to HE from different age groups (for the one reported here covering ages 17 to 20).

Figure 2: Growth of Full-time Students by Level and Domicile in UK Higher Education

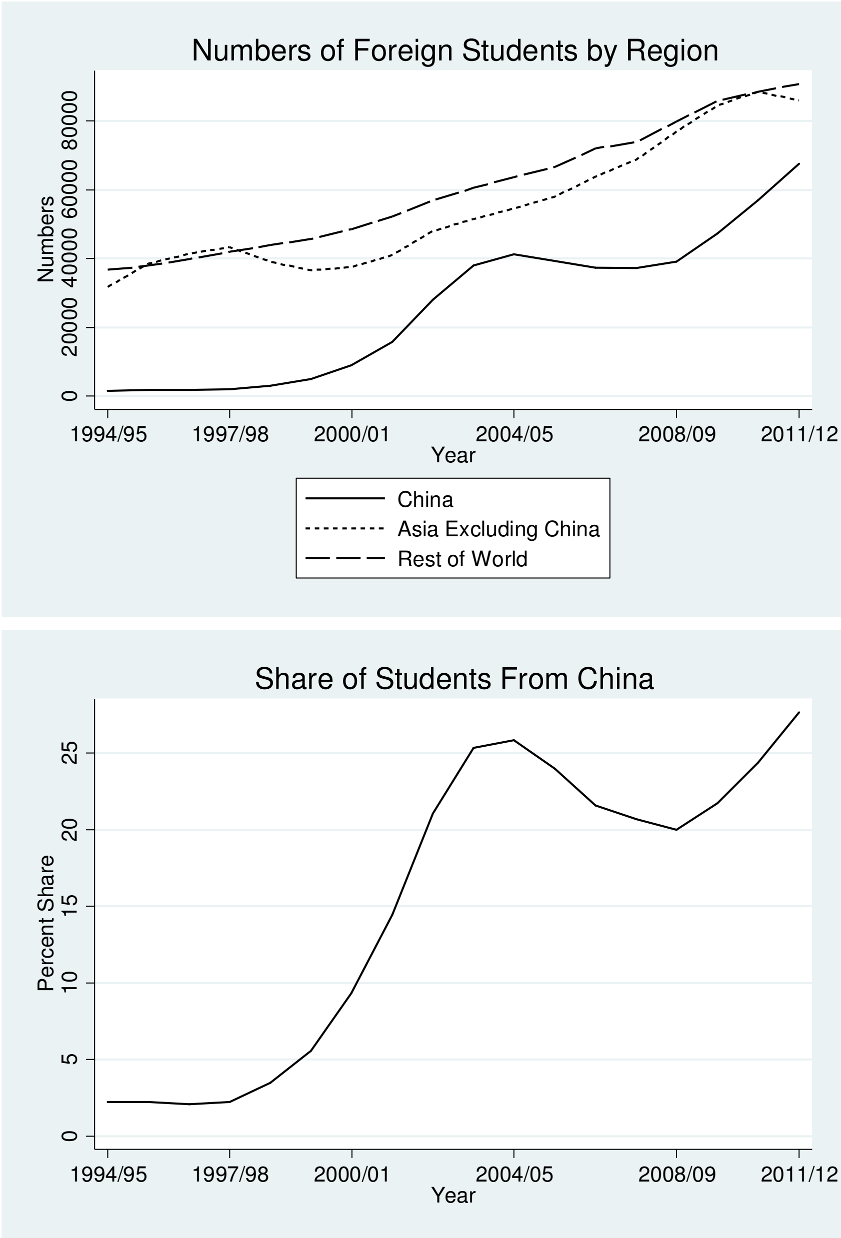




Notes: Domestic students are all students domiciled in the UK paying Home fee levels. Overseas students are all Non-UK students paying Non-Home fees. All totals rounded to nearest 100. Undergraduates; 1994/95- 854,300 domestic, 37,600 overseas; 2011/12 - 1,165,700 domestic, 123,500 overseas. Taught postgraduates: 1994/95- 59,100 domestic, 15,000 overseas. 2011/12- 77,700 domestic, 96,900 overseas. Research postgraduates: 1994/95- 27,900 domestic, 12,000 overseas. 2011/12- 37,300 domestic, 25,800 overseas

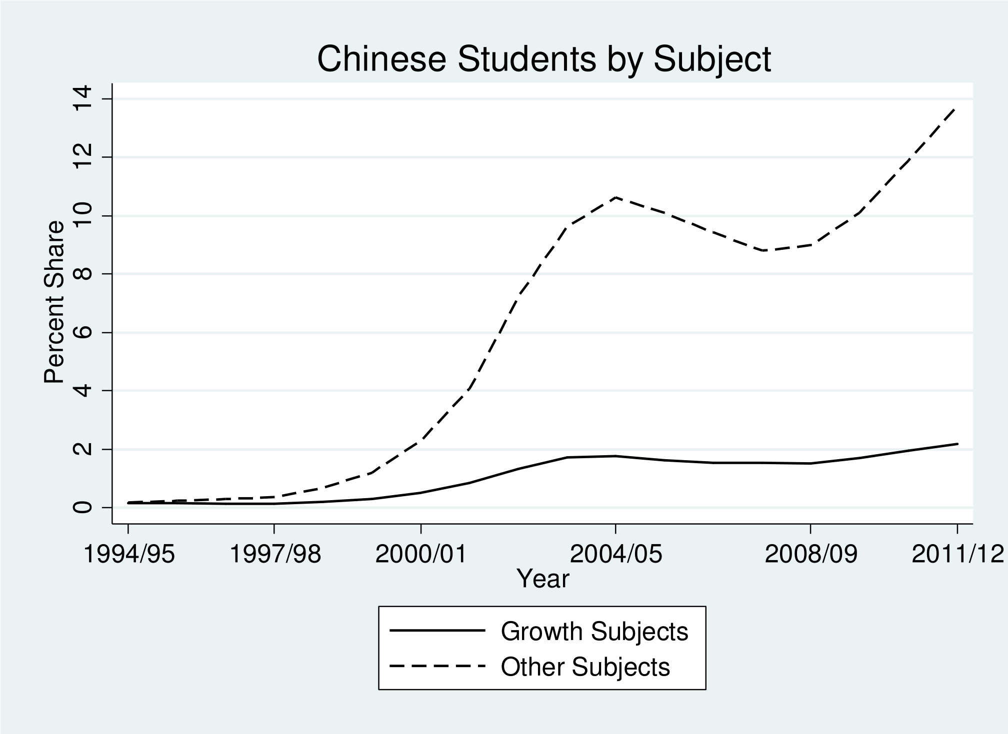
Source: HESA 1994/5- 2011/12 administrative population data.

Figure 3: Flows of International Students to UK Higher Education



Source: HESA 1994/5- 2011/12 administrative population data.

Figure 4: Flows of Chinese Students in Growth and Other Subjects



Source: HESA 1994/5- 2011/12 administrative population data.

Notes: Growth Subjects are 'Business and Management' and Economics

Appendix 1: Data Description

1 Basic Processing of HESA Data

We use HESA standard population measures from 1995 to 2012 (corresponding to the academic years 1994-1995 to 2011-2012), the maximum available amount of years with consistent data definitions. We use restricted to our analysis to the change in number of full time students in university subject areas. Full time students are defined as attending an institution for periods amounting to at least 24 weeks within the year of study and during those weeks studying at least 21 hours. Changes in student numbers are calculated on an annual basis and are standardised by the according total of the previous year. Weights are used in all calculations. Each observation is weighted by the mean student population of the lagged and previous year of the appropriate university subject areas.

2 Higher Education Institutions

There are 210 Higher Education Institutions in the UK over this time period. We have administrative data on 202 of these institutions. The missing universities are: Camborne School of Mines, Liverpool Hope University, Craigie College of Education, Duncan of Jordanstone College of Art, Manchester Business School, Welsh Agricultural College, University College Birmingham, and London Metropolitan University. They were not included as they either enrolled no students that met the student population definition or they had requested HESA not to release the data to researchers. Of the 202 institutions 41 merged with another university during the observation period and so had their totals retrospectively aggregated. This makes the unbalanced panel of 161 universities. Of these 17 are removed as they either open or close and therefore leaves us with a balanced panel of 144 of universities which are used in the final analysis. This consisted of the 144 universities that continually existed from 1995 to 2012.

3 Student Population

We use the HESA Student record which has counts of all students registered at a reporting higher education institution (HE institution) who follow courses that lead to the award of a qualification(s) or institutional credit, excluding those registered as studying wholly overseas e.g. overseas sandwich year students. If it is known at the beginning of the course that a student will spend a block of eight weeks or more in the UK as part of their programme then they are included on the Student record throughout, and not included in the Aggregate offshore record. Moreover Postdoctoral students are also not included in the HESA Student record.

From the HESA Student record the HESA standard HE population has is derived. It includes all higher education enrolments as at 1 December of the academic year, except: dormant students (those who have ceased studying but have not formally de-registered); incoming visiting and exchange students; students studying for the whole of their programme of study outside of the UK; students who left the institution prior to 1 December of the academic year, or who commenced a programme of study after this date; students on sabbatical; writing-up students

The population data is provided in the form of counts by specified student characteristics, e.g. total number of w level students from country x, paying fee level f studying subject y, at university z. These can be aggregated up to departmental, university or national levels for each of these categories. Brief definitions of these categories and how they are aggregated is given below.

4 Student Levels

The Level of Study refers to the qualification aim of the student. These are classified into four levels; *First Degree, Other Undergraduate, Postgraduate Taught, Postgraduate Research*.

First Degree and *Other Undergraduate* refer to Bachelor degrees (BSc, BA, etc.), first degrees with Qualified Teacher Status and equivalents including foundation degrees, diplomas in higher education (including those with eligibility to register to practice with a health or social care or veterinary statutory regulatory body), Higher

National Diploma (HND), and Higher National Certificate (HNC). These levels were combined together to form the undergraduate population measure.

Postgraduate Taught includes master's degrees (MSc, MA, etc.), postgraduate bachelor's degrees at level M and postgraduate diplomas or certificates not studied primarily through research, such as the Postgraduate Certificate in Education (PGCE). It will also include doctorate students not primarily taught through research. These form the second student level.

Postgraduate Research refers to all students studying towards a doctorate, master's degrees and postgraduate diplomas or certificates studied primarily through research.

5 Country of Origin

It is mandatory to collect the domicile of all students. These are mapped to countries using the National Statistics Country Classification 2006 grouping of countries (www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/national-statistics-country-classification/index.html) which provides 251 domiciles. These were reduced down to 75 countries, grouping all countries with less than 5000 students-years in the UK over the entire 18 year period into one category. This represented 6.3 percent of the total population or 9.8 percent of the Non-UK population. Where no data is supplied about the student's domicile, fee eligibility is used to assign to either UK region unknown or Non-European-Union unknown. These countries were basis to form additional regional totals; Domestic-from the UK; EU – students domiciled in the EU accounting for the growth in the EU; Non-EU – remaining countries.

6 Fee Level

Students are either eligible to pay Home Fees or Overseas fees. All students resident in the UK and the remainder of the EU are subsidised by the UK government and are eligible to pay Home Fees. Home Fee was originally set at zero (free) for undergraduate students, but was increased to £1000 per year in 1998/9, a maximum of £3000 in

2006/7.[23] The Home Fees for Taught and Research postgraduate students is unregulated but has remained comparatively low with median fees of £4000 in 2009/10 (Murphy, 2014). All Non-EU students are not subsidised and therefore pay the full market rate for a course. These fees are considerably more with the average Overseas fees for undergraduates were £9,360 and £9,520 for postgraduate students. The data provided information on the fee status of each student; 1) Eligible to pay home fees (87%); 2) Not eligible to pay home fees (10%); and 3) Eligibility to pay home fees not assessed (2%).

7 *Subject of Study*

In the UK system students are studying always towards a particular subject goal. All subjects are categorised into JACS subject 161 codes consisting of a letter followed by a single digit, where the initial letter identifies the subject group. There are 20 major subject groups;

1) Medicine & dentistry; 2) Subjects allied to medicine; 3) Biological sciences; 4) Veterinary & Agriculture science; 5) Physical sciences; 6) Mathematical & Computer science; 7) Engineering; 8) Mineral technology; 9) Architecture, building & planning; 10) Social, economic & political studies; 11) Law; 12) Business & administrative studies; 13) Mass communications & documentation; 14) English/Classics; 15) European Languages; 16) Modern Languages; 17) Historical & philosophical studies; 18) Creative arts & design; 19) Education; 20) Combined.

During the first estimation method using the historical shares of students from country x studying subject y at university z. During the mid-1990's there were fewer overseas students and therefore the subject areas were grouped into 5 major subject groups. For the second estimation method the growth over the whole period was use and therefore allowed to have more subject groupings, including separating the sub-group Economics from the major grouping of Social Science. The coding for these subject aggregations can be found in Data Appendix Table 1.

[23] After the end of the same the Home Fee tuition fee cap increased to £9000 in 2012/13.

8 Currency Exchange

The British Pound Sterling and Chinese Yuan exchange rate is obtained from the International Monetary Fund (http://www.imf.org/external/np/fin/data/param_rms_mth.aspx). This provided the daily exchange rates. The mean annual exchange rate was calculated on the academic year basis up until the September of that year. i.e. the mean exchange rate from September 1st 1994 to August 30th 1995 is used for the academic year 1995-1996. This reflects the exchange rate when potential students were deciding which country/university to attend.

9 Additional Data cleaning

HESA advised that the student totals for Cambridge in 2006 were incorrectly recorded. Correspondingly totals were interpolated by averaging preceding and proceeding years.

Table A1: Subject Coding

	JACS Subject Groups		5 Subject Groups		11 Subject Groups
1	Medicine & dentistry	1	Medicine, Dentistry & allied subjects	1	Medicine & Dentistry
2	Subjects allied to medicine	1	Medicine, Dentistry & allied subjects	1	Medicine & Dentistry
3	Biological sciences	1	Medicine, Dentistry & allied subjects	2	Biology & Veterinary Sciences
4	Veterinary & Agriculture science	1	Medicine, Dentistry & allied subjects	2	Biology & Veterinary Sciences
5	Physical sciences	2	Sciences and MECT	3	Physical Sciences
6	Mathematical & Computer science	2	Sciences and MECT	4	Maths & Computing
7	Engineering	2	Sciences and MECT	5	Engineering
8	Mineral technology	2	Sciences and MECT	6	Architecture & Technology
9	Architecture, building & planning	2	Sciences and MECT	6	Architecture & Technology
10	Social, economic & political studies	3	Social Sciences, Law & Business	7	Law & Social Studies
11	Law	3	Social Sciences, Law & Business	8	Economics
12	Business & administrative studies	3	Social Sciences, Law & Business	9	Business & Management
13	Mass communications & documentation	4	English, Language & History	10	Language & Humanities
14	English/Classics	4	English, Language & History	10	Language & Humanities
15	European Languages	4	English, Language & History	10	Language & Humanities
16	Modern Languages	4	English, Language & History	10	Language & Humanities
17	Historical & philosophical studies	4	English, Language & History	10	Language & Humanities
18	Creative arts & design	5	Creative Arts, Design, Education & Other	11	Education & Creative Arts
19	Education	5	Creative Arts, Design, Education & Other	11	Education & Creative Arts
20	Combined & Other	NA	Not Used	NA	Not Used

Table A.2: Russell Group Universities by Field of Study (5) Panel Estimates

Estimates of

$$(D_{it} - D_{i,t-1})/S_{i,t-1} = \alpha_i + \beta\,(F_{it} - F_{i,t-1})/S_{i,t-1} + T_t + \varepsilon_{it}$$

	Undergraduates		**Taught Postgraduates**		**Research Postgraduates**	
	Home	Domestic	Home	Domestic	Home	Domestic
Ordinary Least Squares:						
Change in Foreign	0.806	0.714	2.241	1.943	0.743	0.554
Students	*0.251*	*0.245*	*1.273*	*1.164*	*0.156*	*0.123*
Two Stage Least Squares:						
Change in Foreign	2.241	2.140	2.286	2.092	0.866	0.683
Students	*1.273*	*1.304*	*0.653*	*0.628*	*0.190*	*0.162*
First Stage: IV1	0.085		0.865		0.950	
	0.147		*0.496*		*0.279*	
First Stage: IV2	0.501		0.102		0.400	
	0.184		*0.444*		*0.156*	
F-Test	4.99		7.44		21.69	
Sample Size	1,167	1,167	1,179	1,179	1,172	1,172
Number of Universities	20	20	20	20	20	20

Notes: All regressions include, year and subject-institution fixed effects, with robust standard errors in italics. Regressions are weighted by the appropriate mean of the student populations over the differenced years. 2SLS F statistic is based on the Kleinbergen-Paap Wald F statistic, allowing for non iid errors.

Table A.3: Non-Russell Group Universities by Field of Study (5) Panel Estimates

Estimates of

$$(D_{it} - D_{i,t-1})/S_{i,t-1} = \alpha_i + \beta\,(F_{it} - F_{i,t-1})/S_{i,t-1} + T_t + \varepsilon_{it}$$

	Undergraduates		Taught Postgraduates		Research Postgraduates	
	Home	Domestic	Home	Domestic	Home	Domestic
Ordinary Least Squares:						
Change in Foreign	0.718	0.703	0.292	0.260	0.955	0.689
	0.321	*0.311*	*0.074*	*0.066*	*0.272*	*0.169*
Two Stage Least Squares:						
Change in Foreign	-0.034	-0.030	0.397	0.335	0.996	0.804
Students	*0.396*	*0.371*	*0.155*	*0.138*	*0.177*	*0.145*
First Stage: IV1	0.223		1.234		1.827	
	0.226		*0.241*		*0.476*	
First Stage: IV2	0.466		-0.265		0.449	
	0.165		*0.184*		*0.214*	
F-Test	19.64		37.62		16.55	
Sample Size	6,277	6,277	5,342	5,342	5,766	5,766
Number of Universities	124	124	124	124	124	124

Notes: All regressions include, year and subject-institution fixed effects, with robust standard errors in brackets. Regressions are weighted by the appropriate mean of the student populations over the differenced years. 2SLS F statistic is based on the Kleinbergen-Paap Wald F statistic, allowing for non iid errors.

Figure A1: Total Numbers of Chinese Students by Field of Study

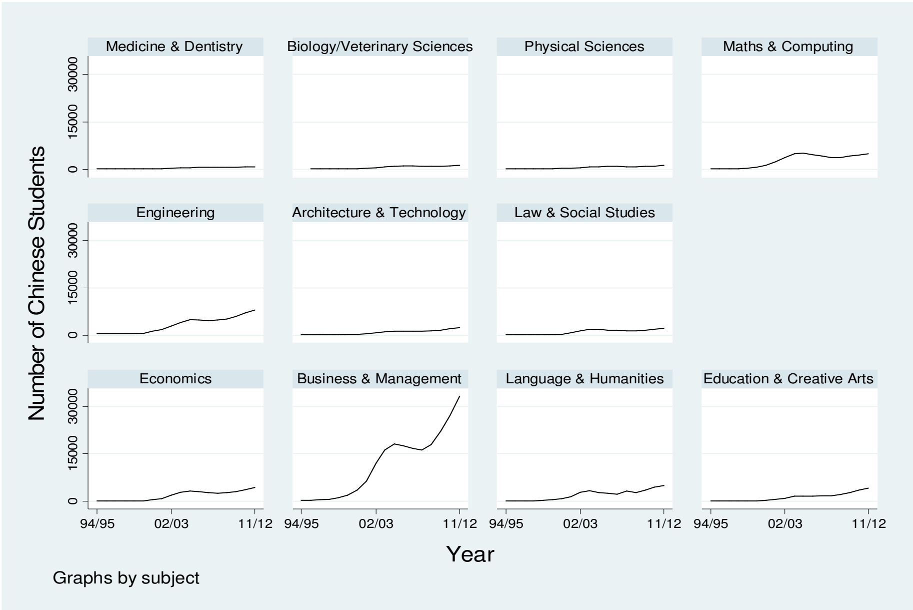


Figure A2: Total Numbers of Undergraduate, Taught Postgraduate and Research Postgraduate Chinese Students by Field of Study

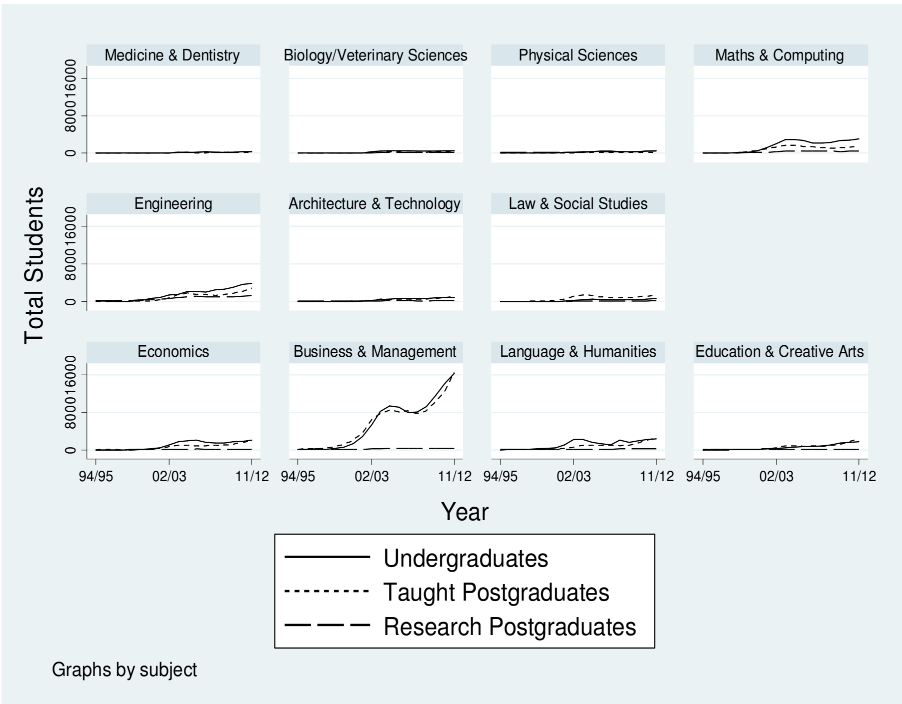
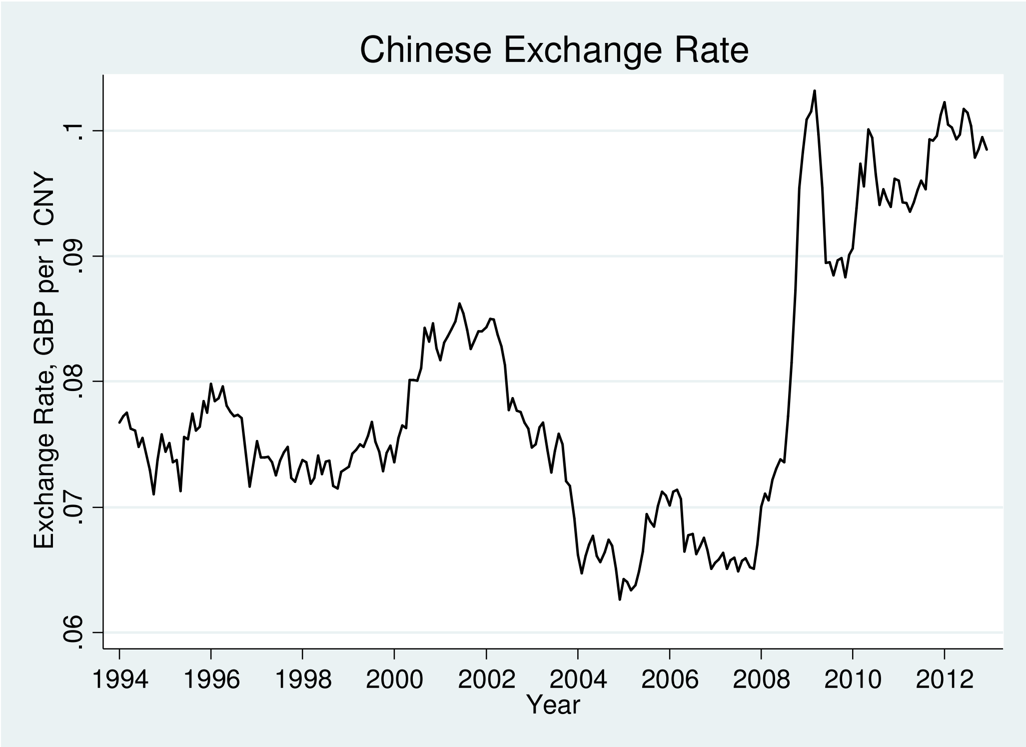


Figure A.3: Pound:Yuan Exchange Rate



Notes: Source International Monetary Fund. Note based on last day of month exchange rates up until 1999, when daily exchange rates are used.

CHAPTER 4.

TRADE UNIONS IN THE AGE OF LITIGATION

I thank Alex Bryson, Claudia Hupkau, Justin Gallagher, Steve Machin, Brendon McConnell, Emma Pickering, Imran Rasul, Felix Weinhardt, the participants of the CEP Labour Market Workshop and the Lancaster Departmental Seminar. All remaining errors are my own.

1. Introduction

Trade union density has been in decline across the developed world and across many sectors (Neumann and Rissman, 1984; Blanchflower and Bryson, 2008). However, there are some occupational groups that have seen an increase in demand for membership. This paper identifies a new source of demand for union representation, and how it has increased union density in these specific labour markets. The paper also provides an explanation for the longstanding free rider problem in the union literature: in occupations where pay and working conditions are determined centrally, and membership is not mandatory, why do individuals pay their monthly union dues (Freeman and Medoff 1984; Bryson and Forth, 2010)?

Society has become increasingly litigious and this could have many repercussions on workplaces and labour markets. These affects would be most acutely felt in occupations where employees have unsupervised interactions with vulnerable groups and are most at risk of accusations. A rational response by these employees to such changes would be an increase demand for insurance against these risks. Trade unions offer legal protection and advice to individuals who were members at the time when the allegation is made and when it was alleged to have occurred. In a vain similar to Blanchflower et al. (1990) who model union demand as a reaction to threat of unemployment, I model union demand as a reaction to the threat of accusations. Here union membership is a form of private legal insurance where individuals can chose to purchase it at an annual cost, and have better expected outcomes if an allegation is made against them. Therefore the decision to join is partly determined by the perceived threat level of having an allegation made against the risk averse agent.

I test this model by estimating the demand for union membership amongst UK teachers over the last two decades. As with most developed countries, the UK has seen a large decline in union membership. Total membership in 1979 stood at 13.2 million, twenty years later this had fallen to 7.9 million (DfB, 2009). The vast majority of occupational groups have experienced a fall in unionisation rates. However, the occupations that have bucked this trend the most, experiencing a rise in union membership, are characterised by having unsupervised employees working with vulnerable groups. The four occupations with the highest percentage point increases in union density since 1992 are Educational Assistants (28.7%), Secondary School

Teachers (12.5%), Primary School Teachers (8.5%) and the Police (6.7%).[1] Teachers were already one of the most unionised occupations making these additional gains even more remarkable.[2] In 1993 76.5% of teachers were unionised, by 2005 this had reached a peak of 87.0%, this 10.5 percentage point gain was whilst the rest of the UK workforce has seen a 6 percentage point decline in union density (Figure 1).

The UK teacher labour market is also a prime example of the trade union free-rider problem, in that pay and working conditions are determined centrally, and teachers are not required to be a member (or pay union dues) to teach. Why do they chose to be members?[3] Moreover, since 1987 teacher trade unions have had no official roll in the determination of pay, and could only state their recommendations. Despite this drop in union bargaining power, union membership has continued to increase, therefore there must be other benefits to joining a union. This is illustrated by the finding that from 2001 the proportion of teachers who thought that their pay and working conditions were affected by trade unions remained constant at, and then decreased from 75%, but the proportion of teachers that were in a union was higher at 83% and increased to 87% by 2005 (Figure 2).

A rational explanation for this increase in demand is an increase in the private benefits of union membership. One such benefit which is highly promoted by the unions is the legal advice and protection provided in the event of an allegation being made.[4] Teachers who are members of a union at the time of the incident and when the allegation is made receive an official representative for the internal disciplinary meetings and legal representation if it does escalate. The teacher trade unions themselves consider this service to be the major driver of union demand.[5] Moreover as part of the terms and conditions of membership many unions reserve the right to use the facts of successful cases to publicise their criminal representation scheme (NASUWT, 2014). It is not just

[1] Of all three digit occupational groups with at least 100 employees per year. The unionisation rate amongst the clergy also increased rapidly reaching a peak in 2005 of 14.3% up from a base of 2.8% in 1992 but had less than 100 observations for 5 of the 18 years.

[2] Educational Assistants 20.4% to 48.1%, Secondary School Teachers 76.1% to 88.6%, Primary School Teachers 82.3% to 90.8% and the Police 76.8% to 83.5%.

[3] The annual membership fees for a full time teacher in 2013 in the two largest teacher unions in the UK were £167 NASUWT and £170 NUT.

[4] The benefits of union membership now revolve more around servicing rather than organising. It is easier to exclude non-members from services such as: Continuing Professional Development, group discounts, group insurance offers.

[5] Paddy Marshal, Head Recruitment Officer NUT stated in a phone interview in relation to the legal insurance that “the safety net is the biggest potential benefit”, April 2009. Tracy Twist, Assistant General Secretary of NASUWT stated in a meeting with me that “a lot of teachers join because of these concerns”.

the unions that consider the threat of accusations as a determinate of union membership, teachers do also. In a survey of Hertfordshire teachers in 2010/11 I found that in answering the question *"What were the MAIN reasons why you initially joined a teacher union?",* the most popular response was "support in the event of allegations from pupils", which 85% of the respondents indicated (Appendix Table 1).

For fear of allegations to explain the rise in demand for union membership the threat of allegations also needs to have risen over this time period. There are no comparable records directly measuring the threat of allegations annually, however one teacher union reported dealing with 71 cases of alleged sexual or physical abuse in 1991, 134 in 1992 and 158 in 1993 (Independent, 1994) and then estimated dealing with 800-900 per year in 2009 (Keates, 2009).[6] To obtain a more detailed and comprehensive measure of threat against teachers, I use the number of national newspaper stories involving accusations of teachers. I treat this media coverage of allegations relating to local teachers as an exogenous shock to the perceived threat. The hypothesis is that in years and regions which have more stories concerning allegations involving teachers, the demand for teacher union membership will increase.

Similar approaches have seen media coverage change individuals' expectations regarding social security (van der Wiel, 2009), inflation (Carroll, 2003; Lamla and Lein 2008) and returns to education (McGuigan et al. 2012). These papers found that newspaper reporting in period t affected expectations of outcomes in period t+1. Relating media reports to the demand for insurance specifically, Gallagher (2013) found that demand for flood insurance increased in regions that experienced flooding in the previous year. Moreover in non-flooded communities that were in the same television media market that the flood took place, these effects continue to persist for five years after the event. This paper provides similar estimates but for the demand for legal insurance post media coverage and adds to the literature, by using the details of each story, to learn more about how individuals update their expectations according to how similar it is to their own situation.

The media coverage information comes from the LexisNexis database over two decades from 1991 to 2011. I use a rubric to codify all news stories from national newspapers in the UK relating to teachers, according to how relevant they would be to a teacher who may be concerned about having an allegation made against them. For

[6] NASUWT membership over this period increased by 63% (Certification Office, 2010) whilst the number of allegations against its members increased by 1167%.

example, a news story concerning false allegations would be extremely highly relevant, but a coverage of a teacher who admits guilt of a crime would not (details in Section 4). I have also extracted additional details from each story including the date, location, gender of the teacher and the school type. Whilst I cannot know how many or which newspapers individuals actually read, I expect that changes in the overall reporting levels to reflect general changes in perceived threat.

I find that unionisation rates increase with media coverage of relevant litigation at the regional and national levels. Ten relevant news stories in a region increase the probability of union membership by 5 percentage points. Additionally, the size of the effect is dependent on the relevance of the story to the teacher. Teachers from secondary schools react to stories involving other secondary school teachers but not to those involving primary school teachers. Similarly, the demand for union membership increases amongst male teachers when there is media coverage concerning other male teachers but not female teachers. For falsification tests, I confirm that media coverage of cases involving teachers which are not relevant to a teacher's decision to join a union are un-related to the unionisation rate. Moreover the impact of these stories decreases as the labour market becomes less similar to that of teachers. The model allows me to simulate demand from 1992 to 2010 with media coverage set to zero throughout the period. In this case predicted union membership remains stable at around 81% whilst observed membership actually rose to 87.5%.

In November 2005 new procedures were introduced by the government which restricted what could be published by newspapers before a case had gone to term (HM Government, 2006). Accordingly, post 2005 I observe a fall in the number of news stories involving teachers and a fall in the overall unionisation rate. I collected additional data on the actual number of allegations through Freedom of Information requests to all Local Authorities in England. There is a geographical correlation between the incidents of allegations and news coverage, but whilst the unionisation rate and amount of media coverage fell the total number of allegations continued to rise.[7] In a horserace between allegations and news coverage, over a shorter time period in a limited sample, I find that the incidents of news stories continue to have a significant relationship, but the number of actual allegations is not correlated. This is indicative that teachers are relating to the perceived threat of an allegation being made, rather than changes to the actual threat

[7] This actual allegations data is only available post 2006 and therefore could not be used for the main analysis.

level. This may not be due to teachers behaving irrationally. Rather news coverage is more salient and so is less costly to collect and use to update expectations, compared with actual allegation probabilities which may not be immediately available.

The contribution of this paper is that it identifies a new source of demand for union representation, and how it has increased union density in specific labour markets against a background of general union decline. The paper also provides an additional explanation for the longstanding free rider puzzle in the union literature, that employees should rationally free-ride on the union benefits and not pay the costs, but individuals do join unions. Therefore there must be a private excludable benefit form membership, this has been discussed previously, but is the first to empirically show the demand for a private service in the form of legal insurance. The policy implication is if governments or employers wanted to reduce the demand for union membership, they could do so by providing more support in the workplace against allegations being made against staff.

The rest of the paper is organised as follows. In Section 2, I formalise a model for union demand dependent on perceived threat of allegation. Section 3 describes the data sources and how the media coverage data was collected. Section 4 presents estimates of the impact of media coverage on demand for union membership. Section 5 provides some falsification exercises and explores the impact of actual versus reported allegations. Section 6 concludes.

2. Demand for Union Membership Model

2.1 Model and assumptions

Teacher unions provide a unique service in the form of legal advice and protection against allegations made by students. I model union membership as form of legal insurance that the teachers can chose to pay for with annual dues. The benefit is that the expected outcome once an allegation has been made is better if the teacher is a member of a trade union.

To formalise this decision process we need to make the following assumptions. There are multiple types of teachers that vary in their risk aversion, their actual risk of allegations being made against them and their other characteristics that can be correlated with the net benefits of union membership. These dimensions of teachers types are summarised by θ, which simultaneously represents risk aversion, riskiness and characteristics of a teacher[8]. This allows for variation in insurance take up across individuals with the same observable characteristics. Each teacher's utility is a function of income Y and type θ, $U(Y,\theta)$, which has decreasing marginal benefits from income. They are employed in schools which are 'open shops', where union and non-union members are both employed and earn the same wages $w > 0$. There is only one trade union and if a teacher decides to join the union they pay annual cost c>0. Therefore teacher wages can either be spent on union fees or left as disposable income $w = Y - c$. The perceived probability of an allegation being made against a teacher with characteristics x from region j, in year t, is $\delta(s_{xjt})$. This an increasing function of news stories s in the first derivative and negative in the second, reflecting the diminishing marginal impact of the news stories in a region.[9] If an allegation is made against a teacher they incur cost a, regardless of the subsequent outcome, reflecting the social costs and potential damage to career prospects. Similarly there is an additional cost l if a teacher is found guilty of an allegation, and that $l \gg c$. I can now rank utilities for any given state of the world for all types θ:

$$U^n(w,\theta) > U^u(w-c,\theta) > U^{nw}(w-a,\theta) > U^{uw}(w-a-c,\theta) > \\ U^{nl}(w-a-l,\theta) > U^{ul}(w-a-c-l,\theta) \quad (1)$$

where U^n and U^u are the utilities of non-members and members respectively with no allegation against them. U^w is the utility of winning a case, U^l is the utility of losing a case, which depend on union status. For union members $U^{uw} = U(w-c-a,\ \theta)$, and $U^{ul} = U(w-c-a-l,\ \theta)$, non-union members utilities U^{nw} and U^{nl} follow a similar structure, but do not incur membership cost c. The state with the highest utility is a non-member with no allegations against them U^n and the worst state is a union member who lost their case U^{ul}.

The probability of a teacher being exonerated is $r(x)$ which is increasing in the amount of resources devoted to their defence x. Teachers are not allowed to employ representation for internal school hearings, the only representation that teachers are

[8] Note that this allows for some types of teachers to potentially commit offences. All teachers were innocent all the time there would be no market for insurance as all teachers would be presumed non guilty.

[9] The perceived threat can also be a function of other factors in addition to news stories, such as the actual number of allegations. Section 5.4 investigates the use of this other less salient measure of threat.

allowed, apart from themselves, is that of the union. Given the restrictions that exist for teachers in employing private representation, I assume that the effective resources that can be devoted to a case are always higher that the cost of union dues (x>>c). This reflects that income cannot be easily translated into defensive resources.[10] Therefore expected utility of a teacher once an allegation is made Z is a convex combination of winning and losing utilities for their union status.

$$Z^n = r(x^n)U^{nw} + (1 - r(x^n))U^{nl} \tag{2}$$
$$Z^u = r(x^u)U^{uw} + (1 - r(x^u))U^{ul}$$

I can now model the teachers decision process regarding union membership. An individual of type θ^* is indifferent between joining a union or not, when there are no marginal benefits, $b = 0$, e.g. the expected utility from membership equals the subjective expected utility from non-membership.

$$b = EU(membership) - EU(non - membership) = 0$$
$$b = \left[\delta(s)Z^u + \left(1 - \delta(s)\right)U(w - c,\, \theta^*)\right] - \left[\delta(s)Z^n + \left(1 - \delta(s)\right)U(w,\, \theta^*)\right] = 0 \tag{3}$$

As $U(w - c,\, \theta^*) < U(w,\, \theta^*)$, and the perceived probability of an allegation is the same for an individual if they are a member or not, this means that the expected utility once an allegation is made for a union member, is greater than for a non-union member $Z^u > Z^n$. We must have that since the only difference between Z^u and Z^n comes from $r(x)$, for unions to have any members we need that unions provide more resources in for defence $r(x^u) > r(x^n)$.

Taking the first derivative of (2) with respect to the number of news stories, it can also be shown that the expected gain from membership for the marginal member is a function of news reports

$$\frac{db}{ds} = [\delta'(s)Z^{\,u} - \delta'(s)U(w - c,\, \theta^*)] - [\delta'(s)Z^{\,n} - \delta'(s)U(w,\, \theta^*)]$$
$$= \delta'(s)(Z^{\,u} - Z^{\,n}) + \delta'(s)[U(w,\, \theta^*) - U(w - c,\, \theta^*)] \tag{4}$$

---

[10] This is likely to be a large contributing factor why no private market for teacher legal insurance exists in the UK.

Given the assumptions that $\delta'(s) > 0$, $Z^{u} - Z^{n} > 0$ and $(w - c,\ \theta^{*}) < U(w,\ \theta^{*})$, then it follows that $\frac{db}{ds}$>0. For an indifferent teacher with taste for risk θ^{*}, the marginal benefit of unions is increasing the number of news stories, as this increases the expected probability of an allegation being made.

2.2 Comparative Statics

I now present comparative statics to illustrate the case of a teacher of type θ^{*} would chose to be a union member when the perceived risk of an allegation is high, but choose not to when the perceived risk is low. Panel A of Figure 3 shows her utility function $U(Y, \theta^{*})$, and the utility levels specified in (1). The chord linking the points U^{ul} and U^{uw} represent the convex combination of the two and represent the expected utility of a member once an allegation has been made (similarly for the points U^{nl} and U^{nw} for non-members). The point along the chord where the teacher will be is determined by the probability of success $r(x)$. As $r(x^{u}) > r(x^{n})$, the union member will be higher up the chord than the non-union member, and we can plot $Z^{u} > Z^{n}$. Note that Z^{u} and Z^{n} intersections both lie beneath the utility curve, reflecting the utility lost through risk aversion.

It follows that a non-union member with income w, would either receive $U(w, \theta^{*})$ or Z^{n}, the expected utility of a non-union member once an allegation has been made. Therefore the expected utility before an allegation is made is a combination of these two outcomes. The lower chord that links the intersection of the utility curve with Z^{n} to the intersection with w, represents this expected utility space of a teacher who is not a union member (Figure 3 Panels B and C). Similarly, before an allegation is made, a union teacher is at a point on the upper chord between $U(w - c, \theta^{*})$ and Z^{u}, reflecting that she is paying dues, but also has Z^{u} expected utility if an allegation is made.

This time point at which a teacher is along this new chord is dependent on their expectations of an allegation being made against them. Panel B of Figure 3 shows a high threat scenario $\delta(s)$=0.5, and the individual will be at the midpoint of each chord. With this high perceived threat level we can see that the expected utility from membership is greater than that of non-membership, $EU^{u} > EU^{n}$. In contrast Panel C shows the same teacher with the same taste for risk and type θ^{*} and same amount of union dues c, would chose not to be in a union if the risk level was low $\delta(s)$=0.1. In this case, although the teacher is at a higher point on each chord, the non-union option is more attractive as no dues are paid.

This basic example demonstrates that the demand for union membership is directly related to the perceived threat of allegations, $\delta(s_{xjt})$. When there is a low probability of an allegation being made, the cost of union membership outweighs the gains through better representation if an allegation was to occur and so teachers will not join the union.

2.3 Econometric Specification

I build on this basic model of rational decision making by the teacher to join the union as the basis of the estimation strategy. It stated that teacher *i* from region *j* in time period *t* will choose to join the union if the expected benefits of joining the union are positive, $EU_{ijt}^{u} - EU_{ijt}^{n} > 0$. Each of these terms will be a function of many factors in addition to perceived threat of an allegation being made and will be related to the teachers type θ. This can be summarised by the two following equations.

$$EU_{ijt}^{u} = \alpha^{u} + \rho^{u}\,\delta(s_{jt}) + \gamma^{u}X_{ijt} + \mu_{j}^{u} + \omega_{t}^{u} + \varepsilon_{ijt}^{u} \tag{5}$$

$$EU_{ijt}^{n} = \alpha^{n} + \rho^{n}\,\delta(s_{jt}) + \gamma^{n}X_{ijt} + \mu_{j}^{n} + \omega_{t}^{n} + \varepsilon_{ijt}^{n} \tag{6}$$

where $\delta(s_{jt})$ is the perceived threat in region j in time period *t*. The remaining parameters account for the other characteristics of a teacher type θ. α^{u} and $\alpha^{n}$are the general benefits for being a (non)union member for all individuals in all time periods. X_{ijt} is a vector of observable individual characteristics which affect the perceived benefits such age, qualifications, or gender. μ_{j}^{u} are the additional gains for being a union member in region j, this could reflect taste for unions in a particular region. ω_{t}^{u} allows for differential gains from union membership each year, which impacts all teachers in in the same way, such as any general perceived fall in union power. Individuals also have an idiosyncratic taste for (non)union membership which varies overtime, ε_{ijt}. The probability that individual *i* in region *j* at time period *t* will be a trade union member is $Pr(EU_{ijt}^{u} > EU_{ijt}^{n})$, using the standard result (McFadden, 1976) we can combine equations 5 and 6 into the following

$$Pr(EU_{ijt}^{u} > EU_{ijt}^{n}) = \frac{\exp(\alpha+\rho\theta(s_{jt})+\gamma X_{ijt}+\mu_{j}Region_{j}+\omega_{t}Year_{t})}{1+exp(\alpha+\rho\theta(s_{jt})+\gamma X_{ijt}+\mu_{j}Region_{j}+\omega_{t}Year_{t})} \tag{7}$$

where each parameter is now the marginal benefit for individual i to join the union ($\rho = \rho^u - \rho^n$). However, I am not able to separately identify the perceived threat from stories $\delta(s_{jt})$ and the marginal gain from a unit of threat ρ. Instead I will be estimating the combination of the two, the expected marginal gain per story. Given that by assumption $\delta(s_{jt})$ is a quadratic function, this will parameterised into the effect per story β_1, and it's square β_2. The demand for union membership can then be estimated using a logistic regression, where the parameters of interest is $\beta_1 + 2\beta_2\overline{s_{jt}}$ representing the marginal effect of an additional story at the mean news coverage on union membership.

$$U_{ijt} = \alpha + \beta_1 s_{jt} + \beta_2 s_{jt}^2 + \gamma X_{ijt} + \mu_j + \omega_t + \varepsilon_{ijt} \tag{8}$$

where U is an indicator variable if individual *i* in period *t* is a union member or not and s_{jt} is the number of stories in region *j* in time period *t*. This specification assumes that media coverage in region j has no impact on the perceived benefits of union membership in a different region. To allow for spill-overs and to obtain estimates of the total impact of news stories on union membership, I additionally include a measure for total news stories nationally each year, s_t and replace the year effects term with a more restrictive time trend term.

$$U_{ijt} = \alpha + \beta_1 s_{jt} + \beta_2 s_{jt}^2 + \beta_3{}^{Nat} s_t + \beta_4{}^{Nat} s_t^2 + \gamma X_{ijt} + \mu_j + \varphi Year_t + \varepsilon_{ijt} \tag{9}$$

Following similar reasoning that teachers are more likely to be affected by news stories originating in their region, one may expect certain stories to have a larger impact on certain teachers who share characteristics with the teacher involved in the media coverage. For example, a news story involving false allegations against a male teacher may be more relevant to other male teachers compared to female teachers. I investigate this by allowing the threat to vary by the characteristics of the teacher in the story s_{xjt} and estimate the impact when the characteristics of the teacher are the same or different to the characteristics of the story, $X_{ijt} = X_{sjt}$ and $X_{ijt} \neq X_{sjt}$. Any differences in the effect may be due to the threat that a given story generates is greater, $\delta(s_{xjt}) > \delta(s_{x'jt})$ when $X_{ijt} = X_{sjt}$ and $X_{ijt} \neq X_{sjt}$ or the expected marginal gain driven by the story is larger ($\rho_{X_{ijt}=X_{sjt}} > \rho_{X_{ijt}\neq X_{sjt}}$). Again, I cannot separately identify these effects but will instead estimate the marginal effect of a similar or less similar story.

$$Union_{ixjt} = \alpha + \beta s_{ixjt} + \beta s_{ixjt}^2 + \gamma X_{ijt} + \mu_j + \omega_t + \varepsilon_{ijt}$$

$$Union_{ixjt} = \alpha + \beta s_{ix'jt} + \beta s_{ix'jt}^2 + \gamma X_{ijt} + \mu_j + \omega_t + \varepsilon_{ijt} \qquad (10)$$

3. Data and Descriptive Statistics

3.1 Union Membership Data

The objective of this paper is to estimate the effect of the threat of allegations on the individual decision to join a trade union. The population of interest is individuals employed in occupations which are at high risk of having an allegation made against them, and where union membership (or purchasing of indemnity insurance) is non-compulsory.[11] I have chosen to focus this paper on the teacher labour market as it is a well-defined occupational group with a large number of employees and has also had considerable press attention regarding accusations from students over the last two decades.

To obtain information on teachers and their union status, I used data from the UK Quarterly Labour Force Surveys (QLFS) over the period 1992 through 2010. The QLFS is conducted by the Office of National Statistics and follows approximately 60,000 households every quarter. A rotating panel of households are surveyed over five quarters, and are then removed and replaced with a new household. Individuals are asked for employment and personal characteristics. This allows me to condition on factors that have been shown to be important determinants of union status (Machin, 2006); age, tenure, gender, region, occupation, public sector employee, qualifications and region. Information relating to union membership is only collected in the autumn quarter.[12] The estimates are generated over the period 1993 to 2010, as some individual characteristics are not available in 1992.

Teachers are identified through three digit occupation codes, which allows teachers who work in Primary Schools, Secondary Schools and Special Schools to be separately

[11] For example, UK doctors are required to become members of the British Medical Association which is the professional association and registered trade union for doctors in the UK. Similarly for Physiotherapists and Radiologists who each have a professional body which provide including insurance cover, professional and legal advice, and support for continuing professional development (Royal College of Radiologists, and Chartered Society of Physiotherapy).

[12] As the QLFS is a continuously rotating panel of households interview over five quarters and information relating to union membership is obtained every autumn, a quarter of individuals are asked twice about their union status. Unfortunately the number of repeated teacher observations is too small to run auxiliary analysis on this sample.

identified. This results in a final sample used for estimation of 30,392 teachers, with on average 1,782 teachers per year, 827 of which teaching at Primary Schools and 817 teaching at Secondary Schools. Summary statistics on teachers in comparison to the workforce in general can be found in Table 1. As one may expect the teacher labour market is considerably different, 88.6% are employed in the public sector compared with 24.6% in the wider economy. Moreover the teacher population is also more female (72.5% versus 47.5%) and has a higher proportion of graduates (74.3% versus 18%).

Regarding the main characteristics of interest, the unionisation rate of teachers is 84% compared with 27.6% for non-teachers and 59.4% in the public sector as a whole. This paper uses the twenty detailed Government Office Regions as the region of analysis, which is derived from Local Authority residence. These larger regions allow for news stories to have wider impacts outside of their immediate vicinity.[13]

3.2 Media Coverage

Many different factors may influence the perceived threat of an allegation being made against a teacher. This paper uses the impact of media coverage originating in the region teacher i is a resident of as an indicator for overall threat. To assume this to be exogenous that there is no moral hazard on the behalf of teachers, that they are not more likely to commit an offence if they are a union member. Moreover it also requires that cases involving union members are no more likely to be mentioned than not. It would be very difficult to have a measure of all news coverage e.g. television programmes, news websites, newspapers and magazines. Therefore, similar to other papers (Carroll, 2003; Lamla and Lein, 2008; van der Wiel, 2009) I will be using the number of articles in national newspapers as a proxy for all media coverage. The data on news stories is obtained from LexisNexis, an online database of media published in the UK.[14] I searched for all articles which contained the word ‘teacher’ in the headline and included any of the following terms (or derivatives) in the headline or main text; *teacher, damages, sued, litigation, allegation, jail, court, dismissed* or *fired*, over the period

[13] Use of the restricted access QLFS with Local Authority information is not possible before 2002 as these files have been converted to the new calendar framework and as union questions were only asked in the Autumn were removed from files.

[14] Some national newspapers were not included in the LexisNexis database throughout the entire period. To have a consistent measure of newspaper coverage over time these newspapers are excluded from the analysis. Their inclusion does not change the interpretation of the results. These newspapers are The Morning Star, The Express, The Daily Telegraph, Sunday Express, Sunday Telegraph, The Sun, The News of the World, The People.

September 1991 to August 2010. Using the date of the QLFS interview and media publication date, I allocate media coverage from the twelve months prior to the interview to the teacher.

As advised in Woolley (2000) I created a rubric by which to classify news stories before the search was conducted. This coding frame classified news stories into four levels according to how relevant they would be to influence the perceived benefit of joining a union; Extremely Relevant- teacher found innocent and case thrown out; Highly Relevant- teacher currently on trial; Little Relevance- Guilty of a lesser offence/on trial strong evidence; Not Relevant- Teacher admits guilt of extreme sexual abuse (See Appendix Table 2). Note, it is possible that a single case can appear in different levels as the newspaper stories develop over time. In total 1709 stories were coded, of which 623 were classified as extremely relevant and a further 548 as highly relevant.

The newspaper stories are also categorised by story type according to if they involve: *Allegations; Being Sued, Suing, Being Attacked, Criminal Activity, Being Sacked, Employment Tribunal* and *Teacher Union Activity*. For the main analysis the story types of interest are *'Allegations'*, *'Being Sued'* and *'Criminal Activity'*, with auxiliary analyses using all story types. The total number of stories of this type in the balanced panel of newspapers that are extremely or highly relevant are 439. Table 2 summarises the total number of stories by level and type.

The number of relevant stories has dramatically increased, between 1992 to 1998 the mean number of stories involving allegations against teachers per year was 6.6, from 1999 to 2007 this had increased to 37.9 (See Figure 3). Post 2005 there was a fall in the number of news stories in national newspapers which coincides which a change in the law which gave more protection to teachers to prevent their case being reported before the individual had been charged with a criminal offence.[15]

In addition to the relevance and region of the news stories, I have also extracted information on the teacher involved in the story. From the name of the teacher, or pronoun used in each story I was able to infer the gender of the teacher. Using references to the school name or the age of the pupils involved I was able to determine if the

[15] In accordance with the Association of Chief Police Officers (ACPO) guidance the police will not normally provide any information to the Press or media that might identify teacher who is under investigation, unless and until the person is charged with a criminal offence. In exceptional cases where the police might depart from that rule, e.g. an appeal to trace a suspect, the reasons should be documented and partner agencies consulted beforehand.

teacher was teaching in a Primary or Secondary school. In this way I was able to assign gender in 96.6% of stories and primary or secondary school level in 82.4% of stories. For stories where the gender or school level of the teacher were not mentioned, the story was not counted for either group. Stories relating to trade unions are only counted towards the total number of stories nationally. Stories involving secondary school teachers are the most common representing 66.3% of highly relevant stories. The balance between the genders is more equal with 50.7% of highly relevant stories involving male teachers and 46.5% involving female teachers (Table 3).

4.2 Actual Allegations

I obtained information on the actual allegations made against public sector employees who work with children and young people through use of the Freedom of Information Act. After contacting all 152 Local Authorities in England I received responses from 118 (See Appendix 1 for detailed list). This details which sector occupational group the allegation was made against, and the nature of the allegation.[16] Teachers received more than half of all allegations out of all occupations that work with children, with 52.6%. Of these, 56.9% are physical in nature and 23.9% sexual, which is comparable with allegations for all non-teacher occupational groups with 52.5% and 25.1% respectively (Table 4). These data also provides a count of the outcomes of allegations over the previous twelve months, which I have codified into four categories; 1) Not Upheld; 2) Police Involvement; 3) Disciplinary Procedures, and 4) Referral.[17] These outcomes cannot be connected to occupations, but in general 46.1% of these allegations are not upheld (Table 5). This means that for an innocent teacher there is still a risk of having an allegation being made against them.

The total number of allegations provides us with the size of actual risk. The shortcoming of these data is that it was not compulsory for Local Authorities to record this information and therefore the data are over a comparatively shorter period of time.

[16] There are 15 occupational groups: Social, Care, Health, Education, Foster Carers, Connexion, Police, YOT, Probation, CAFCASS, Secure Estate, NSPCC, Voluntary Youth Organisations, Faith Groups, Armed Forces, Immigration/Asylum Support Services, and Other. There are five abuse categories: Physical, Emotional, Sexual, Neglect and Other.

[17] The 16 outcome categories are: Not Upheld – No further action after initial consideration, Being unfounded, Being unsubstantiated, Being malicious, Acquittal ; Police Involvement – Criminal investigation, Conviction; Disciplinary Procedures – Disciplinary Action, Suspension, Dismissal, Resignation, Cessation of use, Inclusion on barred/restricted employment list; Referral - Section 47 investigation, Referral to DCSF, Referral to Regulatory Body.

The earliest reports start in 2007 up until the most recent in 2011. From these I have constructed a balanced panel of 66 Local Authorities over a three year period between 2008 and 2010 (Appendix 1).[18] Using this balanced panel the total amount of allegations made against public employees increased from 4691 to 6091 and for teachers this has also marginally increased from 2866 in 2008 to 2944 in 2010. To obtain a measure of relevant threat I normalised these totals by the number of teachers in each Local Authority taken from the School Workforce in England (2011). This growth in allegations equates to an increase of 1.49 to 1.5 allegations per 100 teachers per year. Assuming that these allegations are evenly concentrated over teachers over time, and that 46.1% of allegations are not upheld, this means an average teacher over a career of 35 years can expect a 24.2% chance to have an non-upheld allegation made against them.

4. Results

4.1 Aggregate Trends

Between 1993 and 2005 the union density amongst teachers increased by 10.5 percentage points, whilst amongst non-teachers it fell by 6 percentage points (Figure 1). This increase in unionisation rate has occurred across all teacher age groups, which implies that this growth rate is not solely due to improvement in recruitment rates amongst newly qualified teachers but a general demand in union membership across all teaching age groups (Figure 3).

During this same period the formal bargaining power of trade unions did not improve. Additionally one may have expected a fall in teacher union membership as the Burnham Committee, through which unions were directly involved in the decisions of setting of teacher pay, was replaced by the School Teacher Review Body (STRB) 1991.[19] In the QLFS teachers are asked *'are your pay and working conditions directly affected by agreements between a trade union and your employer?'* This shows a lower proportion of teachers agree with this statement than are union members (Figure 3),

[18] A smaller balanced panel of 32 local authorities over a longer period of 2007 to 2010 is also available and provides similar results.

[19] There was an interim period between 1987 and 1991 when the Minister of Education had the power to determine the size of the pay award directly.

which implies that teachers are members and pay their dues for other reasons. Moreover as pay and working conditions are set centrally for all employees there is a traditional free-rider issue that teachers could chose not to be a union member (and pay the dues) but receive the same pay and conditions. To explain this previous papers have modelled union membership as a form of insurance against becoming unemployed, using the variation in the local unemployment rate as an indicator for the risk (Blanchflower et. al. 1990). However there is indicative evidence that this is not the main driver of union membership amongst teachers as union density continued to fall during the start of the great recession. The hypothesis of this paper is that unions are now providing a private benefit in the form of legal insurance in the case of an allegation being made.

4.2 Main Results

In this section I present estimates of the effect of relevant news stories on union membership. I use the exogenous number of national news stories that originated in a region from the previous twelve months, as a shock to the perceived threat of an allegation being made and would expect an increase in probability in an individual teacher being a union member. Table 6 presents the marginal impacts of a news story from a logistic estimation of specification 8. To aid interpretation these have been transformed from the logistic parameters to the marginal effect multiplied by 100 and so can be thought of as percentage change in probability.[20]

Column 1 of Panel A shows a positive significant raw correlation of 0.548** between the number of Extremely Relevant stories involving an accusation in a region on the likelihood of union membership.[21] Column 2 conditions on individual characteristics that have been shown to be determinants of membership. There is little change in the coefficient which implies that there is little correlation between the incidence of news stories and these characteristics (0.588***). This is what one should expect if there was little sorting across region due to stories by teacher characteristics. Column 3 additionally allows for varying union demand in each region, and is therefore using the within region variation in news stories over time. The final column additionally includes year effects which allows for the average unionisation rate to

[20] Original estimates of the logit parameters available upon request. The parameters are transformed by $P(\widehat{\text{Union}}) * (1 - P(\widehat{\text{Union}}))$, changing them from odds ratios to probabilities at the mean.

21 I define an accusation story as coverage of the following types of stories: Allegations, Being Sued and Criminal Behaviour

increase over time, which is the smallest of the estimates at 0.498 but still significant at the 5% level. The quadratic term is negative and significant, implying that each additional story beyond the first has a smaller impact. Evaluating the marginal effect at the mean, I find that each additional highly relevant story increases the probability of being a union member by 0.428%. Panel B shows the same specifications on the same sample, but uses both extremely and highly relevant stories, instead of just the most relevant. As one may expect the marginal impact is smaller, at 0.380%**, but remains significant.[22]

One may be concerned that these effects could be generated from random fluctuations in the number of news stories by region. Therefore as a robustness check Panel C of Table 6 estimates the impact of stories of Little to No Relevance on union membership. Reassuringly I find that the incidence of these stories have no impact on union demand.

These estimates are the impacts within a region, this would not capture the total impact of news stories annually on national membership, as I am using the annual variation at the regional level whilst accounting for national year effects. To obtain a national impact I replace the 17 year effect terms with a single year trend variable. Total number of stories nationally per year is no longer absorbed by the year effect, and reflects the additional growth over the long run unionisation trend.[23] The corresponding estimates are found in Table 7. The number of the most relevant stories nationally has an additional impact above and beyond the number of regional stories. The impact is smaller than the regional impact (0.108, versus 0.481). Using the average number of stories locally and nationally I can calculate the mean total effect of newspaper stories on union demand. Compared to a year with none of the most relevant news stories, the mean number of stories in the past year increases the probability of union membership by 1.46 percentage points.

4.3 Media Impact by Relevance of Coverage

The model describes a teacher's rational decision process in choosing to become a teacher, highlighting the roll of the threat of litigation driven by news stories, on the

[22] I have run a parallel set of estimations which instead use a measure of news impact, derived from the number of words per story normalised by mean story length in that newspaper in in that year. These results mirror those found in this chapter and are available upon request.

[23] The stories additionally include relevant union activity at the national level that could not be allocated to a specific region.

marginal benefit of joining the union. If a teacher shares more characteristics with the teacher in the story one may expect that the story is more relevant in their updating process.

Table 8 presents results according to the school type of the potential union member works for (Primary School, Secondary School) and by the school type reported in the media. To simplify the table I report the final marginal impact of stories, accounting for the negative quadratic term, conditioning on individual characteristics, year and regional effects (original estimates appear in Appendix Table 3). Column 1 uses the subsample of Secondary School teachers, and Column 2 the sub-sample of Primary School teachers. The top panel estimates the impact of all relevant news stories, and shows that secondary school teachers react to media coverage but there is no significant reaction from primary school teachers. This coincides with there being more relevant stories involving secondary school teachers (relevant news: 285 Secondary stories, 90 Primary stories; Table 3). The lower two panels, Panels B and C, instead use only the stories involving Secondary and Primary School teachers respectively. I find that demand for union membership amongst Secondary School teachers significantly reacts to stories involving other secondary school teachers (0.907***) but not to stories involving Primary School teachers (0.131) (Column 1). For Primary School teachers neither effects are statistically significant, but the coefficient relating to Primary School stories is higher than the one for secondary schools.

These results are replicated in columns 3 and 4 which instead uses all relevant news stories, not just those relating to allegations, criminal activity or being sued. As before this produces similar results to the highly relevant stories, in which secondary school teachers react more in general and react more to secondary school stories than those set in primary schools. With this broader news story definition I now find a marginally significant effect of Primary news stories on Primary School teachers.

Table 9 has the same structure as Table 8 but focuses on the similarity of the teachers' gender to that of the story. Here we see that only female teachers react significantly to relevant news stories in general. However, once we examine the impact by story type, male teachers do significantly react to news stories involving other male teachers (0.591*) but not to those relating to female teachers (-0.056). Interestingly, I also find that female teacher react more to stories involving male teachers rather than female teachers (0.897***, 0.386). One could infer that female teachers, despite ostensibly having more in common with other female teachers mentioned in the press, may associate the incidence of false allegations to be higher in cases involving men and therefore react more to these types of stories. These findings are repeated using other story types (Columns 3 and 4), rather than those just relating to accusations against the teacher and produce similar results.

4.4 Media Impact on Other Occupations

To this point the paper has estimated the impact of media coverage of accusations against teachers on the unionisation rate of teachers. One may have more confidence in the estimates that these stories are reflecting the change in the perceived threat of teachers if they also have effects on other related occupations, and no impact on unrelated occupations. Table 10 shows the impact of these stories on occupations that are increasingly less similar to teaching; educational assistants, higher education professionals, non-teacher public sector graduates, and non-teacher graduates. The coefficients of interest are not significant for any of the other occupational groups. However there is indicative evidence of an effect on education assistants which has a larger marginal effect at the mean compared to the teachers, but is insignificantly determined (0.622 versus 0.428*). This could be reflective of there being only 10,022 Educational Assistants in the sample, compared with 30,392 teachers. Moreover, the marginal (insignificant) effects decrease in size as the occupations become less similar to teachers, with the effect of teacher news stories being a tenth of the size on non-teacher graduates in general.

4.5 Long Run Media Impact

All the estimates presented thus far have been estimating the impact of media coverage that occurred in the twelve months prior to the interview, thereby restricting the impact of news that occurred before this time to have no influence on an individual's decision. This section will vary the exposure length to examine the fade out of these media effects on union membership. Table 11 presents the impact of regional highly relevant media over periods of time increasing in six month periods from six months up to thirty months. I find that stories within the last six months have a similar impact to those over the last twelve, but once the period of time is extended to two years there is no significant impact of total news stories over that period.

This assumes all stories within this time period have the same impact, therefore Table 12 allows for individuals to be more affected by stories that happened more recently and shows the impact of news stories for each six month period up to thirty six months before the interview. Again I find that individuals react in a similar way to stories from the last six to twelve months, and there are effects from stories that happened between a year and eighteen months ago, but stories prior to that have not significant impact. This implies that for those marginal members who were otherwise indifferent to joining, being a union member is not an absorbing state. Alternatively, it could be interpreted that if a potential union member hasn't joined within the first eighteen months of a story being published then that story is not going to impact on their decision.[24]

As we have seen that there are effects of news stories up to eighteen months beforehand, I now estimate the total impact of media coverage on union membership over time. Allowing for separate effects for the amount of news stories in each six month period up to thirty six months prior to the QLFS interview, both nationally and regionally, I predict the probability of union membership for the years between 1993 and 2010. These estimates are plotted in Figure 6. The model is a good fit as these predicted probabilities fit very closely to the plotted series of actual union density, rarely diverging from the 95% confidence interval band. To estimate the aggregate impact the increased perceived threat has had on union membership, I use these estimates re-predict union membership for each year, but fix the total news coverage to zero. This provides a counterfactual time series of what would have occurred had there been no increase in the threat of allegations. The figure shows that without media coverage the union membership would have been relatively stable at around 81% from 1996 onwards, in contrast to it steadily rising and reaching a peak of 87.5%. In the period from 1999 through to 2009 the probability of union density is significantly greater than estimates where there was no media coverage and between 2002 to 2008 the estimated difference in union membership is 5% points.

[24] Despite observing approximately 25% of teachers twice a year apart, the sample is too small to separate these effects.

4.6 Actual allegations versus media impact

Thus far I have used media coverage as the determinant of the threat of an accusation being made against a teacher. But it may be the case that these news stories reflect a growing number of actual allegations being made against teachers, and the coverage itself has no impact. To the extent that any national increase in allegations would be accounted for with the individual year effects, and local differences in threat rate would similarly be accounted for with the regional effects, we may be convinced that this is the impact of media coverage. However, this makes an assumption that the incidence of a news story originating from a region in a given year is not strongly correlated with the number of incidents. I test this directly using the balanced panel of Local Authorities from 2008 to 2010, aggregating totals to be reflective of the region as a whole. The correlation between the incidence of highly relevant news stories and allegations per teacher is 0.27.[25] Regressing the number of news stories from a region on the number of allegations per hundred teachers conditional on year and regional effects, still leaves a significant correlation. For each additional allegation per hundred teachers 1.36 additional extremely relevant news stories appear in a national newspaper originating from that region.[26] Therefore I reject the assumption that in this limited period the number of news stories is not related to the number of allegations per teacher.

Ideally it would be possible to run a *horserace* between news stories and allegations over the entire sample, to determine which the more important factor is. However I only have the number of allegations for the last three years of the sample and for a limited number of Local Authorities, which limits any inference that can be made. With this in mind, I include the number of allegations per teacher in a region simultaneously with the number of news stories, in a specification similar to 8. This is a greatly reduced sample, but I continue to find that the number of news stories has a significant effect, but the actual rate of allegations is uncorrelated (Table 13). This is indicative evidence that it is the more salient news stories rather than the actual risk of

[25] The correlation with the total number of news stories in a region was 0.001 and insignificant, however the allegations per teacher would be the parameter of interest in a union demand specification.

[26] This result is from a OLS regression of extremely relevant news stories from a region on the number of allegations per hundred teachers in that region. This uses fifteen regions over a three year period with regional and year effects. This has a coefficient of $\beta_{allegations}$ =1.364 and standard error 0.4401. Using the broader classification of highly relevant stories, generates more stories $\beta_{allegations}$ =2.90 and standard error 0.794.

allegations that change an individual's demand for union membership. This may be expected as the underlying actual risk is already accounted for by the teacher and the media coverage are changes to this perceived threat.

5. Conclusions

This paper examines the role of the threat of accusations has had in the demand for trade union membership amongst teachers in the UK. I have found that teachers from regions in which news stories concerning accusations against other teachers originated are more likely to join a union in the following eighteen months. For every ten stories a teacher is 5% more likely to join. These effects are larger if teachers share characteristics with the teacher mentioned in the story, e.g. secondary school teachers react more to stories involving other secondary school teachers, similarly for male teachers. I show that the impact of stories are again larger the more relevant they are to an innocent teacher. In contrast occupations that a less like teachers do not react to these stories.

The specification accurately predicts the changes in union membership since 1993. Setting media stories to zero throughout the period, I forecast that union membership would remain steady at approximately 81% rather than increasing to 87% as seen in the data. This paper provides evidence as to why the unionisation rate amongst some occupational groups with direct and unsupervised interaction with vulnerable members of the public, has increased. Moreover, I provide a further answer to the puzzle of why individuals choose to join the union even if they could free ride and receive the higher pay and working conditions derived through union action without having to pay the union dues. Unions offering a private excludable service can maintain demand for membership, as long as demand for that service remains. The implication for policy is that there may be an increasing unmet demand for union membership in previously under-unionised service sectors. Moreover if regulations are introduced that protect individuals from allegation, then the demand for union services, and hence membership, will decline. Suggestive evidence for this can be seen in the fall in union density post the 2005 governmental reforms on newspaper reporting, which continued to fall despite the worsening of economic conditions, which is traditionally thought of as a key driver of union demand.

References

Blanchflower, D. and Bryson. A. (2008). Union Decline in Britain, IZA, Discussion Paper 3436, April 2008

Blanchflower, D. Crouchley. R. Estrin, S. and Oswald, A., (1990). Unemployment And The Demand For Unions, Centre for Labour Economics, Working Papers 372, London School of Economics.

Bryson. A. and Forth J. (2010). Trade Union Membership and Influence 1999-2009. Centre for Economic Performance, Discussion Paper 1003.

Chartered Society of Physiotherapists. (2013). Last accessed 27 May 2014, http://www.csp.org.uk/membership/membership-benefits

Carroll, C. D. (2003), Macroeconomic Expectations of Households and Professional Forecasters, Quarterly Journal of Economics, 118(1), pp.269-298.

Department for Business Enterprise & Regulatory Reform. (2009). Trade Union Membership 2008, Office of National Statistics

Department for Education, 2012, School workforce in England: November 2011, Last accessed 27 May 2014, https://www.gov.uk/government/publications/school-workforce-in-england-november-2011

Freeman, R. B. and Medoff. J. L. (1984). What Do Unions Do? New York: Basic Books, 1984

Gallagher, J. (2014). Learning about an Infrequent Event: Evidence from Flood Insurance Take-up in the US, American Economic Journal: Applied forthcoming.

HM Government. (2006). Working Together to Safeguard Children. The Stationery Office.

Keates C, (2009), General Secretary of NASUWT interview on “File on Four” BBC Radio 4, April 2009.

Lamla, M. and Lein, S. (2008), The Role of Media for Consumers' Inflation Expectation Formation, KOF Working Papers 201, KOF Swiss Economic Institute.

McMuigan. M. McNally. S. and Wyness G. (2012). Student Awareness of Costs and Benefits of Educational Decisions: Effects of an Information Campaign, Centre for Economic Performance, Discussion Papers 139

Murphy, R. (2011). Impact of UK Resilience Programme training on Staff. Hertfordshire Council, Programme Report.

NASUWT. (2006). "Legal Assistance Scheme: Alleged Criminal Offence", NASUWT Terms and Conditions, Last accessed 27 May 2014 http://www.nasuwt.org.uk/consum/groups/public/@legalandcasework/documents/nas_download/nasuwt_000791.pdf

Neumann. G. R., and Rissman E.R. (1984). Where Have All the Union Members Gone? Journal of Labor Economics, Vol. 2, No. 2, Essays in Honor of Melvin W. Reder (Apr., 1984), pp. 175-192.

van der Wiel, K. (2009). Have You Heard the News? How Real-Life Expectations React to Publicity, IZA Discussion Papers, 4064, Institute for the Study of Labor

Viscusi, W. K. and Hakes, J.: 2008, Risk Beliefs and Smoking Behavior, Economic Inquiry 46(1), pp 45-59.

Table 1: Descriptive Statistics of Employees

	Teachers		All Employees	
	Mean	Std. Dev.	Mean	Std. Dev.
	(1)	(2)	(1)	(2)
Union Member	0.840	*0.367*	0.276	*0.447*
Public Sector	0.886	*0.317*	0.246	*0.431*
Male	0.275	*0.447*	0.525	*0.499*
Full Time	0.786	*0.410*	0.738	*0.440*
University Qualification	0.743	*0.437*	0.180	*0.384*
A-Level Qualification	0.761	*0.426*	0.304	*0.460*
Age	42.67	*10.32*	40.29	*12.78*
Tenure				
less than 3 months	0.066	*0.249*	0.058	*0.235*
3 months but less than 6	0.016	*0.125*	0.047	*0.211*
6 months but less than 12	0.026	*0.158*	0.068	*0.252*
1 year but less than 2	0.082	*0.275*	0.107	*0.309*
2 years but less than 5	0.188	*0.390*	0.207	*0.405*
5 years but less than 10	0.205	*0.403*	0.193	*0.395*
10 years but less than 20	0.241	*0.428*	0.196	*0.397*
20 years or more	0.176	*0.381*	0.123	*0.329*
Government Region				
Tyne and Wear	0.015	*0.122*	0.018	*0.132*
Rest of North East	0.025	*0.155*	0.024	*0.154*
Greater Manchester	0.037	*0.190*	0.039	*0.194*
Merseyside	0.022	*0.145*	0.019	*0.138*
Rest of North West	0.049	*0.217*	0.050	*0.218*
South Yorkshire	0.021	*0.142*	0.021	*0.144*
West Yorkshire	0.038	*0.191*	0.037	*0.190*
Rest of Yorkshire & Humberside	0.028	*0.165*	0.029	*0.167*
East Midlands	0.073	*0.260*	0.074	*0.262*
West Midlands Metropolitan County	0.041	*0.198*	0.039	*0.193*
Rest of West Midlands	0.048	*0.213*	0.050	*0.218*
East of England	0.097	*0.296*	0.099	*0.299*
Inner London	0.030	*0.170*	0.034	*0.180*
Outer London	0.068	*0.252*	0.066	*0.248*
South East	0.145	*0.352*	0.147	*0.354*
South West	0.079	*0.269*	0.088	*0.283*
Wales	0.050	*0.217*	0.046	*0.208*
Strathclyde	0.039	*0.193*	0.035	*0.185*
Rest of Scotland	0.057	*0.232*	0.055	*0.228*
Northern Ireland	0.040	*0.195*	0.030	*0.170*
Observations	30,392		988,256	

Source: QLFS 1993-2010 Autumn Survey, sample of all employees 18-64
Notes: Teachers defined as Standard Occupational Classification codes (1993-2000):233, 234, 235 and Standard Occupational Classification codes (2001-2010): 2314, 2315, 2316

Table 2: Descriptive Statistics of News Coverage 1991-2010

Panel A: All Newspaper Stories

Relevance of Story	Allegations	*Being Sued*	*Suing*	*Being Attacked*	*Criminal Activity*	Sacked	*Employment Tribunal*	*Union Activity*	Total
	Story Type								
Extremely Relevant	322	45	100	4	12	15	61	64	623
Highly Relevant	179	28	52	45	53	36	43	112	548
Little Relevance	155	12	3	19	123	14	12	56	394
Not Relevant	55	1	2	10	68	4	0	4	144
Total	711	86	157	78	256	69	116	236	1709

Panel B: Balanced Newspaper Panel Stories

Relevance of Story	Allegations	*Being Sued*	*Suing*	*Being Attacked*	*Criminal Activity*	Sacked	*Employment Tribunal*	*Union Activity*	Total
Extremely Relevant	222	27	78	3	6	9	48	48	441
Highly Relevant	115	22	36	29	37	16	35	78	368
Little Relevance	95	5	1	10	77	8	9	46	251
Not Relevant	38	1	2	1	32	0	0	2	76
Total	470	55	117	43	152	33	92	172	1136

Source: LexisNexis 1991-2010. News search of national newspapers with the following term: *headline(teacher) and court or damages or sued or jail or litigation or dismissed or fired or allegations and #GC329#.* National Newspapers: Daily Mail, Daily Star, Mail on Sunday, Morning Star, The Express, Sunday Express, The Daily Telegraph, Sunday Telegraph, The Sun, The News of the World, The Guardian, The Independent, The Observer, The People, The Times, The Sunday Times. The Balanced Panel of Newspaper Stories: Daily Mail, Mail on Sunday, The Guardian, The Independent, The Mirror, Daily Star, Observer, The Times, The Sunday Times

Table 3: Total News Coverage by Story Subject

Panel A: All Newspaper Stories 1992-2010	Relevant Stories		Any Relevance Stories	
News Story Subject	Story Type Accusation	All Types	Story Type Accusation	All Types
By School Type				
Secondary School	435	661	706	975
	(68.1%)	(66.2%)	(67.0%)	(66.1%)
Primary School	126	186	184	249
	(19.7%)	(18.6%)	(17.5%)	(16.9%)
By Teacher Gender				
Male Teacher	327	469	591	762
	(51.1%)	(46.9%)	(56.1%)	(51.6%)
Female Teacher	303	521	455	705
	(47.4%)	(52.2%)	(43.2%)	(47.8%)
All Stories	639	999	1053	1476

Panel B: Balanced Newspaper Panel Stories 1992-2010	Relevant Stories		Any Relevance Stories	
News Story Subject	Story Type Accusation	All Types	Story Type Accusation	All Types
By School Type				
Secondary School	285	439	443	620
	(66.3%)	(63.9%)	(65.0%)	(63.7%)
Primary School	90	142	128	182
	(20.9%)	(20.7%)	(18.8%)	(18.7%)
By Teacher Gender				
Male Teacher	218	315	381	490
	(50.7%)	(45.9%)	(55.9%)	(50.4%)
Female Teacher	200	362	289	471
	(46.5%)	(52.7%)	(42.4%)	(48.4%)
All Stories	430	687	677	973

Source: LexisNexis 1991-2010 of National Newspapers, Balanced Panel

Note: Percentages in parentheses represent proportion of all stories of that type on that subject. Story Type: Accusation includes- Allegations, Being Sued and Criminal Activity. Union Activity not included under All Types as is only counted in national totals as not based in one region or reflect a specific teacher type. Total stories do not equal those from Table 3 as some stories are double counted when both male and female teachers are mentioned, or both primary and secondary schools are mentioned.

Table 4: Descriptive Statistics of Allegations by Employer and Type of Allegation

	Type of Allegation					
Panel A: All Reporting Local Authorities 2007-2011						
Employer	**Physical**	**Emotional**	**Sexual**	**Neglect**	**Other**	**Total**
Education	6,267	932	2,642	316	862	11,019
Foster Carers	1,512	305	388	255	70	2,530
Social Care	1,085	169	356	176	112	1,898
Secure Estate	384	15	26	0	6	431
Health	257	42	177	66	41	583
Voluntary Youth Organisations	203	34	342	23	48	650
Faith	177	8	96	1	12	294
Police	142	33	72	9	12	268
Immigration	39	2	39	6	0	86
Connexions	14	4	14	3	5	40
Youth Offending Teams	10	8	19	6	9	52
Armed Forces	6	0	25	1	0	32
Probation	5	0	2	1	0	8
NSPCC	4	1	2	0	1	8
CAFCASS	1	2	2	1	1	7
Other	1,380	247	941	233	247	3,048
Total by type	11,486	1,802	5,143	1,097	1,426	20,954
Panel B: Balanced Panel of Local Authorities 2008-2010						
Employer	**Physical**	**Emotional**	**Sexual**	**Neglect**	**Other**	**Total**
Education	2,440	284	1,123	111	224	4,182
Foster Carers	647	120	171	98	32	1,068
Social Care	486	76	167	85	48	862
Secure Estate	159	8	12	0	0	179
Health	129	27	90	24	19	289
Voluntary Youth Organisations	84	13	151	13	25	286
Faith	76	3	40	1	5	125
Police	76	23	22	3	3	127
Immigration	11	1	19	3	0	34
Connexions	6	3	7	0	2	18
Youth Offending Teams	3	3	6	6	4	22
Armed Forces	0	0	11	0	0	11
Probation	4	0	1	1	0	6
NSPCC	2	0	1	0	1	4
CAFCASS	0	1	2	0	1	4
Other	465	115	374	94	77	1125
Total by type	4,588	677	2197	439	441	8,342

Source: Freedom of Information Requests to English Local Authorities
Note: Lists of responding Local Authorities and balanced Panel of Local Authorities is in Appendix 1

Table 5: Recorded Outcomes of Allegations

	Allegation Outcome				
Panel A: All Reporting Local Authorities 2007-2011					
	Not Upheld	Police Involvement	Disciplinary Procedures	Referral	Total
Total	4,680	1,030	3,058	1,373	10,141
Percent of total	46.1%	10.2%	30.2%	13.5%	
Panel B: Balanced Panel of Local Authorities 2008-2010					
	Not Upheld	Police Involvement	Disciplinary Procedures	Referral	Total
Total	3,384	656	2,305	1,022	7,367
Percent of total	45.9%	8.9%	31.3%	13.9%	

Source: Freedom of Information Requests to English Local Authorities

Notes: Not Upheld – No further action after initial consideration, Being unfounded, Being unsubstantiated, Being malicious, Acquittal ; Police Involvement – Criminal investigation, Conviction; Disciplinary Procedures – Disciplinary Action, Suspension, Dismissal, Resignation, Cessation of use, Inclusion on barred/restricted employment list; Referral - Section 47 investigation, Referral to DCSF, Referral to Regulatory Body. Total outcomes do not equal total number of cases as not all cases had an outcome in the last 12 months.

Table 6: Effect of News Coverage on Union Membership

Panel A: Extremely Relevant News Stories of Accusations				
P(Union Membership)	(1)	(2)	(3)	(4)
Stories Regionally	0.548**	0.588***	0.674**	0.498**
	0.235	*0.206*	*0.325*	*0.251*
Stories Regionally	-0.024	-0.034**	-0.047**	-0.046***
Squared	*0.018*	*0.015*	*0.019*	*0.014*
Marginal Effect at	0.512**	0.535***	0.603*	0.428*
Mean	*0.234*	*0.207*	*0.326*	*0.252*
Panel B: All Relevant News Stories of Accusations				
P(Union Membership)	(1)	(2)	(3)	(4)
Stories Regionally	0.841***	0.783***	0.758***	0.449***
	0.158	*0.139*	*0.200*	*0.149*
Stories Regionally	-0.041***	-0.039***	-0.034***	-0.026***
Squared	*0.008*	*0.007*	*0.008*	*0.007*
Marginal Effect at	0.731***	0.679***	0.667***	0.380**
Mean	*0.160*	*0.140*	*0.201*	*0.150*
Panel C: Little/No Relevance News Stories of Accusations				
P(Union Membership)	(1)	(2)	(3)	(4)
Stories Regionally	0.098	0.169	0.222	-0.152
	0.202	*0.177*	*0.153*	*0.146*
Stories Regionally	0.015	0.003	-0.004	0.004
Squared	*0.012*	*0.010*	*0.005*	*0.005*
Marginal Effect at	0.120	0.173	0.216	-0.145
Mean	*0.202*	*0.177*	*0.153*	*0.146*
Teacher Characteristics		✓	✓	✓
Regional Effects			✓	✓
Year Effects				✓
Observations	30,392	30,392	30,392	30,392

Source: QLFS 1993-2010 *Notes*: This table presents estimates from 12 logit regressions of individual decision to join a union, four per panel. Reporting the marginal effects after transforming by $P(\widehat{Union}) * (1 - P(\widehat{Union}))$. All coefficients and standard errors are multiplied by 100 for ease of interpretation. Estimates can be read a percentage change in probability. Marginal effect at mean calculated by $\beta_1 + 2\beta_2 \overline{s_{jt}}$. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. Stories Regionally is a count for the number of news stories that originated in the region that the teacher resides. Standard errors in *italics*, clustered at the regional level.

Table 7: Effect of Regional and National News Coverage of Accusations on Union Membership

	Story Relevance		
P(Union Membership)	*Extremely Relevant Stories*	*Relevant Stories*	Little/No Relevance
Stories Regionally	0.481*	0.436***	-0.159
	0.263	*0.147*	*0.137*
Stories Regionally Squared	-0.042***	-0.022***	0.005
	0.015	*0.001*	*0.004*
Stories Nationally	0.108**	0.030	-0.205
	0.050	*0.043*	*0.236*
Stories Nationally Squared	-0.002*	0.000	0.005
	0.001	*0.001*	*0.001*
Marginal Effect at Mean	0.469*	0.403**	-0.182
	0.270	*0.158*	*0.160*
Total Effect at Mean	1.460	1.205	-2.003
Teacher Characteristics	✓	✓	✓
Regional Effects	✓	✓	✓
Time Trend	✓	✓	✓
Observations	30,392	30,392	30,392

Source: QLFS 1993-2010 Notes: This table presents estimates from three logit regressions of individual decision to join a union on news stories. Each column shows the impact of stories of varying relevance. Reporting the marginal effects. All coefficients and standard errors are multiplied by 100 for ease of interpretation. Estimates can be read a percentage change in probability. Marginal effect at mean calculated by $\beta_1 + 2\beta_2\overline{s_j}$. Total effect at mean $(\overline{s_j}\beta + \overline{s_j^2}\,\beta + \bar{s}\beta + \overline{s^2}\beta)$. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. *Relevant Stories* include both *Extremely* and *Highly Relevant* Stories. Stories Regionally is a count for the number of news stories that originated in the region that the teacher resides in the previous 12 months. Stories Nationally is a count for the number of all news stories in the previous 12 months, including stories that could not be allocated to a specific region. Standard errors in *italics*, clustered at the regional level.

Table 8: Effect of News Coverage on Union Membership by Teacher and Story School Type

P(Union Membership)	Stories of Accusations: Secondary School Teachers (1)	Stories of Accusations: Primary School Teachers (2)	All Story Types: Secondary School Teachers (3)	All Story Types: Primary School Teachers (4)
Panel A: Relevant Stories				
Total Marginal Effect	0.696***	0.042	0.437**	0.090
	0.241	*0.147*	*0.196*	*0.136*
Panel B: Relevant Secondary School Stories				
Total Marginal Effect	0.907***	0.048	0.389*	0.127
	0.229	*0.306*	*0.239*	*0.218*
Panel C: Relevant Primary School Stories				
Total Marginal Effect	0.131	0.627	0.057	0.632*
	0.664	*0.672*	*0.461*	*0.403*
Teacher Characteristics	✓	✓	✓	✓
Regional Effects	✓	✓	✓	✓
Year Effects	✓	✓	✓	✓
Observations	13,949	14,076	13,949	14,076

Source: QLFS 1993-2010 *Notes*: This table presents estimates from 12 logit regressions of individual decision to join a union. Columns 1 and 3 use the sub sample of Secondary School Teachers, columns 2 and 4 use the sub sample of Primary School Teachers. Panel A uses all estimates the impact of all relevant stories, Panel B the impact of all relevant secondary school stories and Panel C all relevant Primary School stories. Reporting the marginal effects at mean after accounting for quadratic terms. All coefficients and standard errors are multiplied by 100 for ease of interpretation. Estimates can be read a percentage change in probability for an additional story. Marginal effect at mean calculated by $\beta_1 + 2\beta_2\bar{s_j}$. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. Relevant Stories include both Extremely and Highly relevant stories. Stories Regionally is a count for the number of stories that originated in the region that the teacher resides in the previous 12 months. Standard errors in *italics*, clustered at the regional level.

Table 9: Effect of News Coverage on Union Membership by Teacher and Story Gender

	Stories of Accusations		All Story Types	
	Male Teachers	Female Teachers	Male Teachers	Female Teachers
P(Union Membership)	(1)	(2)	(3)	(4)
Panel A: Relevant Stories				
Total Marginal Effect	0.038	0.51**	0.147	0.294
	0.154	*0.201*	*0.18*	*0.136*
Panel B: Relevant Male Teacher Stories				
Total Marginal Effect	0.591*	0.897**	0.564*	0.473*
	0.305	*0.374*	*0.428*	*0.363*
Panel C: Relevant Female Teacher Stories				
Total Marginal Effect	-0.055	0.386	0.086	0.128
	0.305	*0.412*	*0.220*	*0.221*
Teacher Characteristics	✓	✓	✓	✓
Regional Effects	✓	✓	✓	✓
Year Effects	✓	✓	✓	✓
Observations	8,361	22,031	8,361	22,031

Source: QLFS 1993-2010 Notes: This table presents estimates from 12 logit regressions of individual decision to join a union on news stories. Columns 1 and 3 use the sub sample of Male Teachers, columns 2 and 4 use the sub sample of Female Teachers. Panel A uses all estimates the impact of all relevant stories, Panel B the impact of all relevant stories involving male teachers and Panel C all relevant stories involving female teachers. Reporting the marginal effects at mean after accounting for quadratic terms. All coefficients and standard errors are multiplied by 100 for ease of interpretation. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. Relevant Stories include both Extremely and Highly relevant news stories. Stories Regionally is a count for the number of news stories that originated in the region that the teacher resides in the previous 12 months. Standard errors in *italics*, clustered at the regional level.

Table 10: Effect of News Coverage on Union Membership by Occupation

Occupation Group	Teachers	Education Assistants	Higher Education	Non Teacher Public Sector Graduates	Non Teacher Graduates
P(Union Membership)	(1)	(2)	(3)	(4)	(5)
Relevant Stories	0.498**	0.688	0.235	0.161	0.052
Regionally	*0.251*	*0.577*	*0.422*	*0.202*	*0.090*
Relevant Stories	-0.046***	-0.021	-0.018	-0.001	-0.000
Regionally Squared	*0.014*	*0.023*	*0.018*	*0.010*	*0.005*
Marginal Effect at	0.428*	0.622	0.185	0.133	0.051
Mean	*0.252*	*0.582*	*0.425*	*0.204*	*0.091*
Teacher Characteristics	✓	✓	✓	✓	✓
Regional Effects	✓	✓	✓	✓	✓
Year Effects	✓	✓	✓	✓	✓
Observations	30,392	10,022	9,007	49,671	154,932

Source: QLFS 1993-2010 Notes: This table is presenting estimates from five logit regressions of teacher news stories of different occupational groups. I am reporting the marginal effects after transforming by $P(\widehat{Union}) * (1 - P(\widehat{Union}))$. All coefficients and standard errors are multiplied by 100 for ease of interpretation. Estimates can be read a percentage change in probability. Marginal effect at mean calculated by $\beta_1 + 2\beta_2\bar{s_j}$. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. Relevant Stories include both Extremely and Highly relevant news stories. Stories Regionally is a count for the number of news stories that originated in the region that the teacher resides in the previous 12 months. SOC codes: Educational Assistants 652, 6124; Higher Education 230, 231, 2311, 2312. Standard errors in *italics*, clustered at the regional level.

Table 11: Effect of News Coverage on Union Membership by News Coverage Period

News Coverage period P(Union Membership)	In last 6 months (1)	In last 12 months (2)	In last 18 Months (3)	In last 24 Months (4)	In last 30 Months (5)	In last 36 Months (6)
Relevant Stories	0.540**	0.449***	0.521***	0.331***	0.141	0.065
Regionally	*0.260*	*0.149*	*0.107*	*0.119*	*0.103*	*0.104*
Relevant Stories	-0.058***	-0.026***	-0.024***	-0.014**	-0.005	-0.001
Regionally Squared	*0.019*	*0.007*	*0.004*	*0.005*	*0.004*	*0.004*
Marginal Effect at	0.472*	0.380**	0.429***	0.258**	0.112	0.055
Mean	*0.261*	*0.150*	*0.425*	*0.204*	*0.106*	*0.109*
Teacher Characteristics	✓	✓	✓	✓	✓	✓
Regional Effects	✓	✓	✓	✓	✓	✓
Year Effects	✓	✓	✓	✓	✓	✓
Observations	30,392	30,392	30,392	30,392	30,392	30,392

Source: QLFS 1993-2010 Notes: Reporting the marginal effects after transforming by $P(\widehat{Union}) * (1 - P(\widehat{Union}))$. All coefficients and standard errors are multiplied by 100 for ease of interpretation. Estimates can be read a percentage change in probability. Marginal effect at mean calculated by $\beta_1 + 2\beta_2\overline{s_j}$. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. Relevant Stories include both Extremely and Highly relevant news stories. Stories Regionally is a count for the number of news stories that originated in the region that the teacher resides in the previous X months. Standard errors in *italics*, clustered at the regional level.

Table 12: Effect of News Coverage on Union Membership by News Lag Period

News Lag period P(Union Membership)	Marginal Effects (1)	Total Marginal Effect (2)
Stories Last 6 Months	0.487**	0.424**
	0.214	*0.215*
Stories Last 6 Months Squared	-0.053***	
	0.019	
Stories 7-12 Months Previous	0.508***	0.453***
	0.148	*0.148*
Stories 7-12 Months Previous Squared	-0.037***	
	0.008	
Stories 13-18 Months Previous	0.948***	0.861**
	0.348	*0.349*
Stories 13-18 Months Previous Squared	-0.078***	
	0.030	
Stories 19-24 Months Previous	-0.182	-0.166
	0.217	*0.218*
Stories 19-24 Months Previous Squared	0.010	
	0.013	
Stories 25-30 Months Previous	-0.319	-0.280
	0.326	*0.328*
Stories 25-30 Months Previous Squared	0.035	
	0.032	
Stories 31-36 Months Previous	-0.385	-0.348
	0.296	*0.296*
Stories 25-30 Months Previous Squared	*0.015*	
Teacher Characteristics	✓	✓
Regional Effects	✓	✓
Year Effects	✓	✓
Obs	30,392	30,392

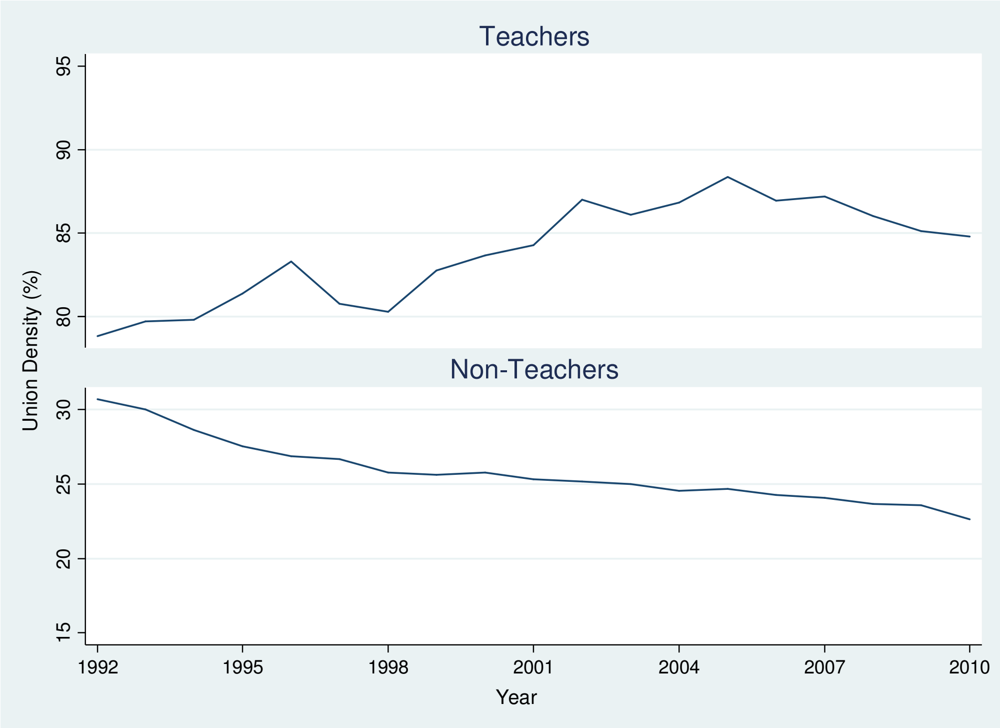
Source: QLFS 1993-2010 Notes: This table presents estimates from one logit regression. Column 1 reports the marginal effects. Column 2 reports the total marginal effect at mean. All coefficients and standard errors are multiplied by 100 for ease of interpretation. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. Relevant Stories include both Extremely and Highly relevant news stories. Standard errors in *italics*, clustered at the regional level.

Table 13: Effect of News Coverage of Accusations and Actual Allegations on Union Membership

News Relevance	Extremely Relevant Stories			Relevant Stories		
P(Union Membership)	(1)	(2)	(3)	(4)	(5)	(6)
Stories Regionally		0.348**	0.335***		0.220	0.197
		0.140	*0.109*		*0.173*	*0.168*
Stories Regionally Squared		-0.055**	-0.040**		-0.050	-0.046
		0.021	*0.019*		*0.093*	*0.065*
Allegations Per 100 Teachers	-0.310*		-0.355	-0.310*		-0.388
	0.227		*0.215*	*0.227*		*0.253*
Teacher Characteristics	✓	✓	✓	✓	✓	✓
Regional Effects	✓	✓	✓	✓	✓	✓
Year Effects	✓	✓	✓	✓	✓	✓
Observations	3,399	3,399	3,399	3,399	3,399	3,399

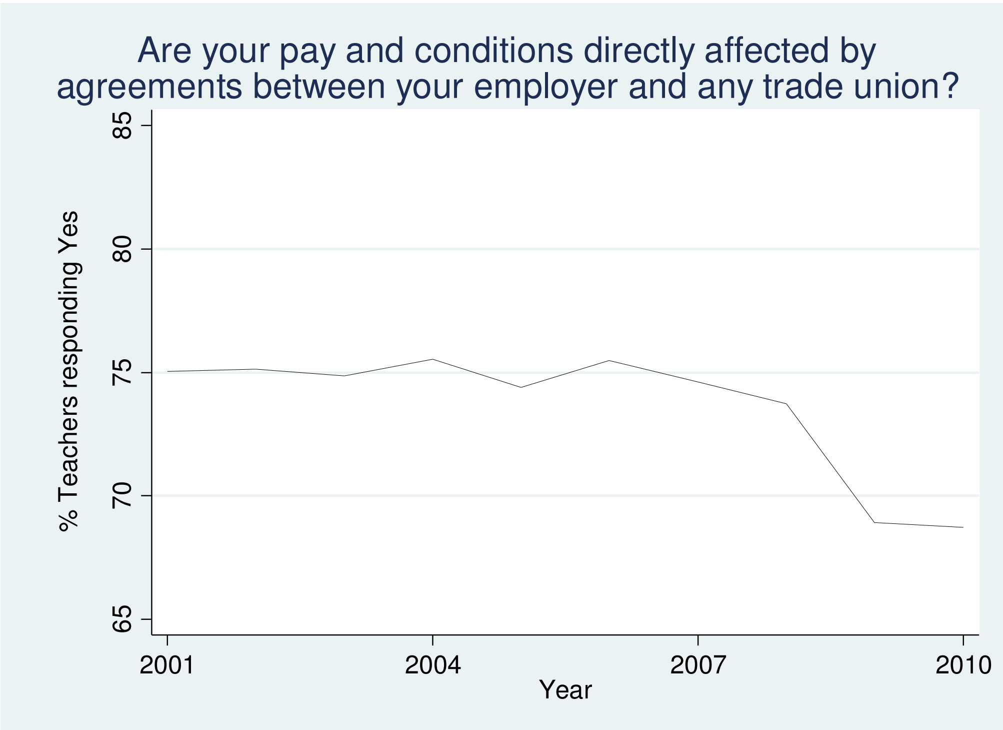
Source: QLFS 2008-2010 Notes: Estimates from a logit regression. Reporting the marginal effects All coefficients and standard errors are multiplied by 100 for ease of interpretation. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. Relevant Stories include both Extremely and Highly relevant news stories. Standard errors in *italics*, clustered at the regional level.

Figure 1: Union Density Time Series by Occupation



Source: QLFS 1992-2010

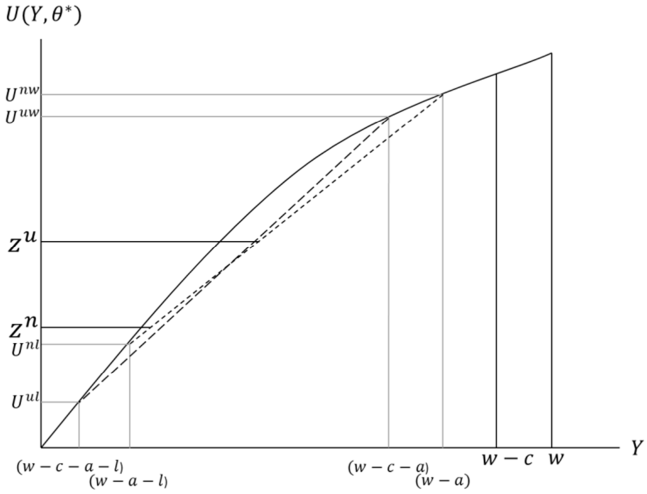
Figure 2: Teacher Perception of Union Power



Source: QLFS 2001-2010: All employees, regardless of union status were ask "Are your pay and conditions directly affected by agreements between your employer and any trade union?

Figure 3: Illustration of Union Membership Decision

Panel A: Utility curve of teacher $U(Y, \theta^*)$ with wages w, union dues c.

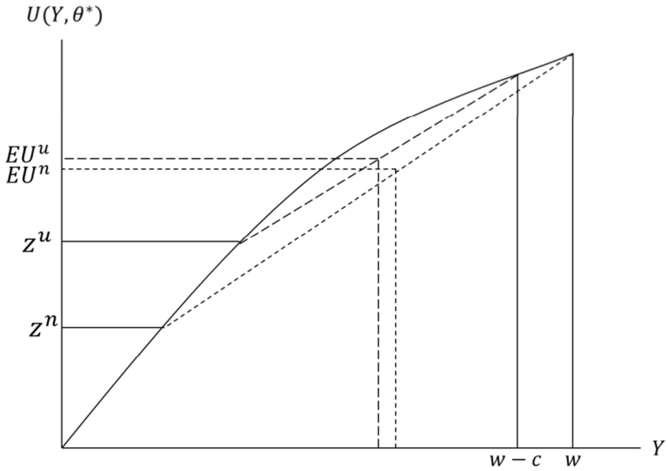


Notes: U^{uw} (U^{ul}) is the utility of a union member who has had an allegation made against them and won (lost) their case. Similarly for non-members U^{nw} (U^{nl}). Z^u (Z^n) represents the expected utility once an allegation has been made for a (non) union member. a represents the cost of an allegation and l the additional cost of being found guilty.

– – – – *Union Member*

- - - - - - *Non Union Member*

Panel B: Expected utilities of teacher $U(Y, \theta^*)$ with a high perceived risk $\delta(s)$=0.5



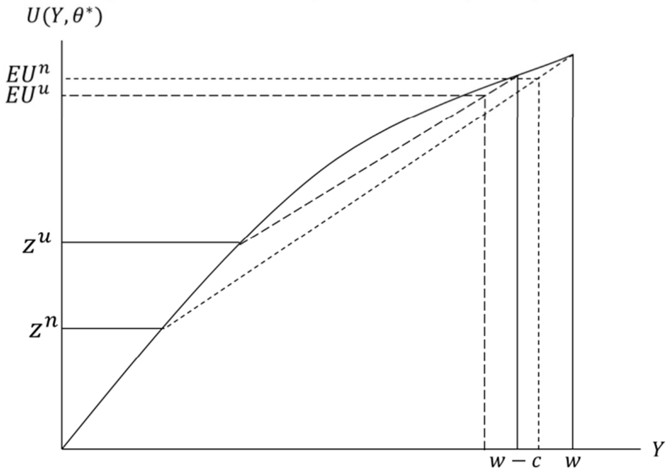
Notes: EU^u (EU^n) represents the expected utility of a (non) union member for a given threat level $\sigma(s)$.

When the risk is high δ(s)=0.5, at the midpoint of each cord, then $EU^u > EU^n$.

– – – – *Union Member*

- - - - - - *Non Union Member*

Panel C: Expected utilities of teacher $U(Y, \theta^*)$ with a low perceived risk $\delta(s)$=0.1



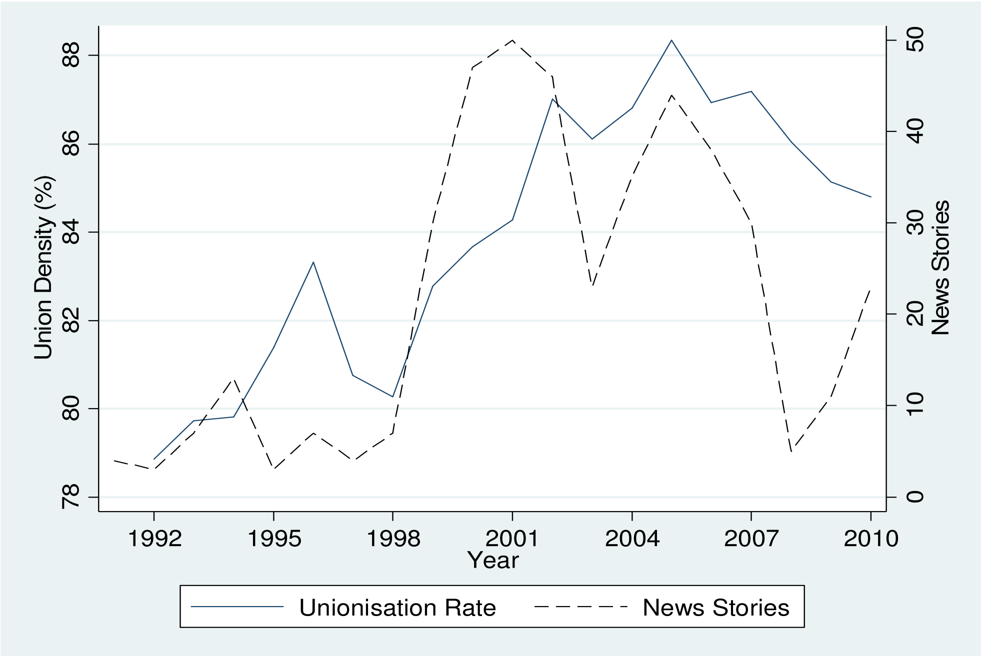
Notes: EU^u (EU^n) represents the expected utility of a (non) union member for a given threat level $\sigma(s)$.

When the risk is low δ(s)=0.1, teachers are at a higher point on each cord and then $EU^u > EU^n$.

– – – – *Union Member*

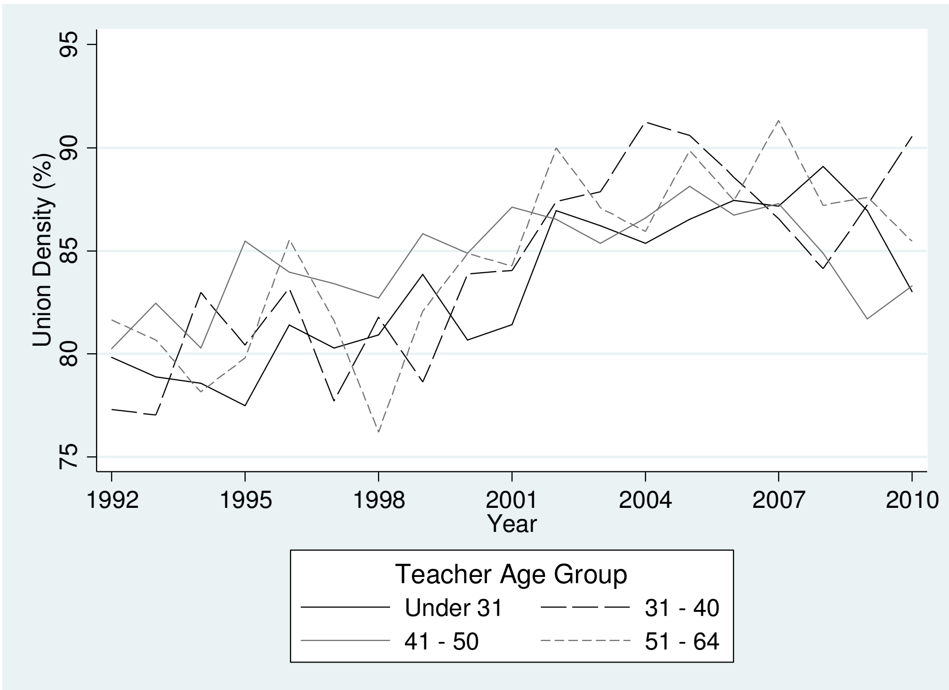
- - - - - - *Non Union Member*

Figure 4: Union Density and Relevant News Stories over Time



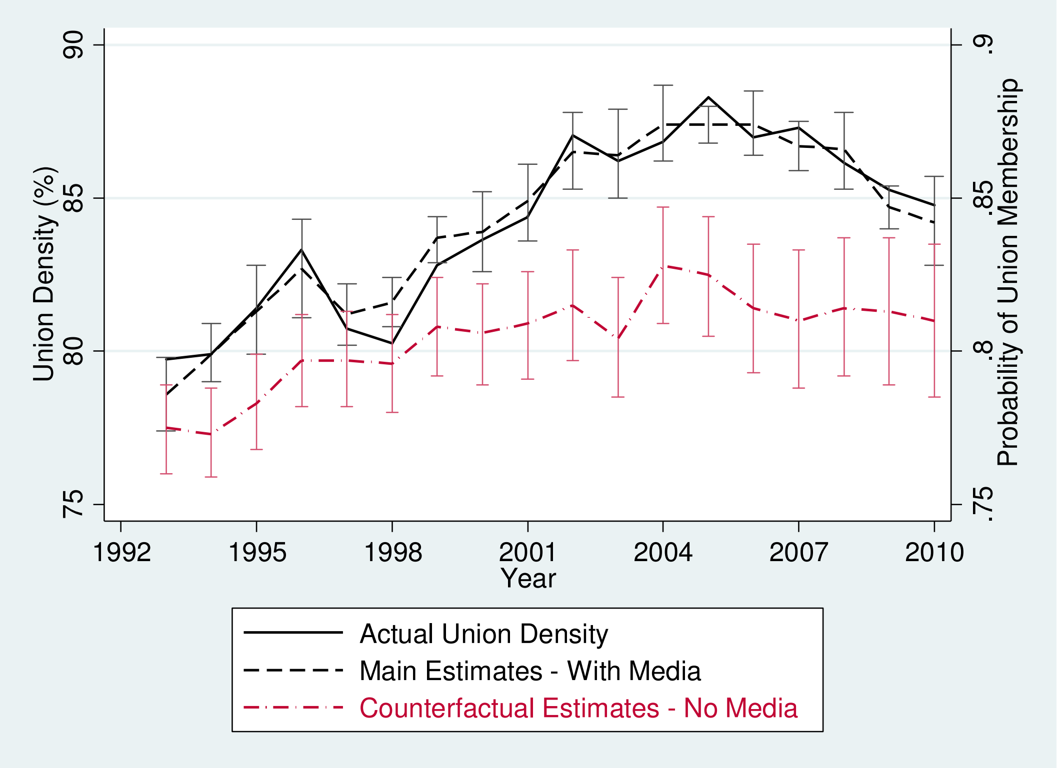
Source: QLFS 1992:2010, Lexis Nexus 1992-2010 *Notes*: Annual union density based on mean union membership of teachers based on QLFS reporting year. News story total based on total relevant news stories about teachers concerning Allegations; Being Sued, and Criminal Activity over a calendar year.

Figure 5: Union Density by Age of Teachers over Time



Source: QLFS 1992:2010 *Notes*: Annual union density based on mean union membership of teachers based on QLFS reporting year.

Figure 6: Predicted Union Density With and Without Media Reports



Source: QLFS 1993-2010 *Notes*: Predictions of probability union from a logit regression for each year. Allow separate effect of news stories regionally and nationally (and their square), for each six month period up to thirty six months prior to the interview. The counterfactual estimates generated with the same parameters apart from setting the media terms to zero. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. Stories Regionally is a count for the number of news stories that originated in the region that the teacher resides in the previous 6 months, 7-12 months, 13-18 months, 19-24 months, 25-30 months, 31-36months. Similarly Stories Nationally is a count for the number of all news stories, including stories that could not be allocated to a specific region. Standard errors in *italics*, clustered at the regional level.

Appendix 1: Local Authorities who responded to the Freedom of information request regarding allegations

All Local Authorities who responded (Years of data):

Local Authority (Years), Barnet (2) Barnsley (3), Bath and North East Somerset (3), Bedford (1), Bexley (2), Blackburn with Darwen (3), Bolton (3), Bracknell Forest (2), Bradford (3), Brent (4), Bristol City (3), Bromley (3), Buckinghamshire (4), Calderdale (3), Cambridge (2), Camden (3), Central Bedfordshire (1), Cheshire East Council (1), Cheshire West and Chester (2), Cornwall (1), Croyden (3), Cumbria (3), Derby (1), Derbyshire (3), Devon (1), Doncaster (3), Dorset (3), Dudley (3), Durham (3), East Riding of Yorkshire (4), East Sussex (2), Essex (4), Gateshead (3), Gloustershire (2), Greenwich (4), Hackney (1), Hammersmith and Fulham (2), Hampshire (3), Haringey (2), Havering (4), Hertfordshire (2), Hillingdon (3), Hounslow (2), Isle of Scilly (4), Isle of Wight (3), Islington (4), Kensington and Chelsea (2), Kent (4), Kingston Upon Hull (3), Kingston Upon Thames (4), Kirklees (3), Knowsley (3), Lancashire (4), Leeds (4), Leicester (3), Lewisham (4), Lincolnshire (1), Liverpool (1), Luton (2), Manchester (2), Medway (3), Milton Keyenes (1), Newham (1), Norfolk (3), North East Lincolnshire (3), North Lincolnshire (1), North Somerset (4), North Yorkshire (3), Northumberland (4), Nottingham City (4), Nottingham County (2), Oldham (4), Oxfordshire (4), Peterborough (1), Plymouth (4), Poole (3), Reading (4), Redbridge (3), Richmond (1), Rochdale (3), Rotherham (1), Rutland (4), Salford (4), Sandwell (3), Scilly Isles (4), Sheffield (2), Shropshire (1), Slough (2), Solihull (4), Somerset (4), South Glouster (2), Southampton (2), Southend (3), St Helens (4), Stockport (4), Suffolk (3), Surrey (2), Sutton (2), Swindon (2), Telford and Wrekin (2), Thurrock (4), Torbay (3), Trafford (2), Wakefield (3), Walsall (4), Waltham Forest (3), Wandsworth (4), Warrington (2), West Berkshire (2), West Sussex (3), Wigan (2), Wiltshire (2), Winsor and Maidenhead (2), Wirral (4), Wokingham (2), Wolverhampton (2), Worcestershire (4), York (3), All (323)

Balanced Panel of Local Authorities 2008-2010:

Barnsley, Bath and North East Somerset, Blackburn with Darwen, Bolton, Bradford, Brent, Bristol City, Bromley, Buckinghamshire, Calderdale, Camden, Croydon, Cumbria, Derbyshire, Doncaster, Dorset, Dudley, Durham, East Riding of Yorkshire, Essex, Greenwich, Hampshire, Havering, Hillingdon, Isle of Scilly, Isle of Wight, Islington, Kent, Kingston Upon Hull, Kingston Upon Thames, Kirklees, Lancashire, Leeds, Leicester, Lewisham, Medway, North East Lincolnshire, North Somerset, North Yorkshire, Northumberland, Nottingham City, Oldham, Oxfordshire, Plymouth, Poole, Reading, Redbridge, Rutland, Salford, Sandwell, Sicily Isles, Solihull, Somerset, Southend, St Helens, Stockport, Suffolk, Thurrock, Torbay, Wakefield, Walsall, Waltham forest, Wandsworth, West Sussex, Wirral, Worcestershire

Appendix Tables

Table A1: Reasons for Union Membership

"What were the MAIN reasons why you initially joined a teacher union?"	
Belief in the union movement	40%
To improve job security	44%
To improve terms and conditions	56%
For solidarity with other workers	24%
Advice/opinion on educational policy	62%
Support in the event of allegations from pupils	85%
No particular reason	3%
Observations	*176*

Source: Online Survey of Hertfordshire Teachers 2010/11 for unrelated evaluation of UK Resilience Programme on teaching staff (Murphy 2011)

Table A2: Media Rubric

	Allegations	***Being Sued***	***Suing***	***Being Attacked***	***Criminal Activity***	**Sacked**	***Employment Tribunal***	***Union Activity***	**Total**
Extremely Relevant	Found innocent, case thrown out	Teacher sued for school activity	Sues for damages/ libel	Pupil attacks teacher in classroom	Manslaughter of pupil charges	For health and safety or allegations	Legitimate Unfair dismissal	Discuss threat of allegations/ being sued	
Stories	*322*	*45*	*100*	*4*	*12*	*15*	*61*	*64*	*623*
Highly Relevant	Currently on trial, no verdict	May be sued, could be sued	Lose case, indirectly related to school	Parent-Pupil attacks teacher outside of school	Criminal accusations from pupil	Inappropriate behaviour, not up to standards	Other Unfair dismissal, inappropriate behaviour	As above but brief mention or union demands	
Stories	*179*	*28*	*52*	*45*	*53*	*36*	*43*	*112*	*548*
Little Relevance	Guilty of lesser offence, on trial of hard offence	School/ Council sued	Threats to sue for indirect teaching	Attacked by ex pupil	School related crime	Miscellaneous school related activity	Union back the dismal	Comment on education policy	
Stories	*155*	*12*	*3*	*19*	*123*	*14*	*12*	*56*	*394*
No Relevance	Admits guilt of extreme sexual abuse	Non school related activity	Non school related activity	Non school related activity	Child pornography /murder	Non-school related activity	Non-school related activity	Anti-union members	
Stories	*55*	*1*	*2*	*10*	*68*	*4*	*0*	*4*	*144*
Total	*711*	*86*	*157*	*78*	*256*	*69*	*116*	*236*	*1709*

Table A3: Union Membership by Teacher and Story School Type – Showing Quadratic Terms

Probability of Union Membership	Accusation Stories: Secondary School Teachers (1)	Accusation Stories: Primary School Teachers (2)	All Story Types: Secondary School Teachers (3)	All Story Types: Primary School Teachers (4)
Panel A: Relevant Stories				
Stories Regionally	0.795***	0.066	0.524***	0.127
	0.239	0.146	0.193	0.134
Stories Regionally Squared	-0.035**	-0.009	-0.020***	-0.009
	0.010	0.007	0.007	0.006
Panel B: Relevant Secondary School Stories				
Secondary Stories	1.000***	0.063	0.447*	0.151
	0.228	0.304	0.238	0.217
Secondary Stories Squared	-0.049**	-0.008	-0.021**	-0.009
	0.001	0.015	0.010	0.010
Panel C: Relevant Primary School Stories				
Primary Stories	0.11	0.685	-0.009	0.701*
	0.663	0.671	0.457	0.352
Primary Stories Squared	0.036	-0.108*	0.068	-0.079*
	0.069	0.060	0.064	0.035
Observations	13,949	14,076	13,949	14,076

Source: QLFS 1993-2010 *Notes*: This table presents estimates from 12 logit regressions of individual decision to join a union. Columns 1 and 3 use the sub sample of Secondary School Teachers, columns 2 and 4 use the sub sample of Primary School Teachers. Panel A uses all estimates the impact of all relevant stories, Panel B the impact of all relevant secondary school stories and Panel C all relevant Primary School stories. Reporting the marginal effects at mean. All coefficients and standard errors are multiplied by 100 for ease of interpretation. Estimates can be read a percentage change in probability for an additional story. Marginal effect at mean calculated by $\beta_1 + 2\beta_2\bar{s_J}$. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. Relevant Stories include both Extremely and Highly relevant stories. Stories Regionally is a count for the number of stories that originated in the region that the teacher resides in the previous 12 months. Standard errors in *italics*, clustered at the regional level.

Table A4: Union Membership by Teacher and Story Gender – Showing Quadratic Terms

	Accusation Stories		All Story Types	
Probability of Union Membership	Male Teachers (1)	Female Teachers (2)	Male Teachers (3)	Female Teachers (4)
Panel A: Relevant Stories				
Stories	0.075	0.594***	0.218	0.363***
	0.154	*0.200*	*0.179*	*0.135*
Stories Squared	-0.015**	-0.031**	-0.018	-0.016**
	0.006	*0.009*	*0.006*	*0.005*
Panel B: Relevant Male Teacher Stories				
Male Stories	0.714*	1.090***	0.684**	0.602*
	0.401	*0.367*	*0.319*	*0.347*
Male Stories Squared	-0.105**	-0.151***	-0.069	-0.067
	0.047	*0.055*	*0.051*	*0.056*
Panel C: Relevant Female Teacher Stories				
Female Stories	-0.036	0.426	0.130	0.152
	0.304	*0.411*	*0.219*	*0.220*
Female Stories Squared	-0.016	-0.028	-0.020	-0.009
	0.015	*0.022*	*0.008*	*0.009*
Observations	8,361	22,031	8,361	22,031

Source: QLFS 1993-2010 Notes: This table presents estimates from 12 logit regressions of individual decision to join a union on news stories. Columns 1 and 3 use the sub sample of Male Teachers, columns 2 and 4 use the sub sample of Female Teachers. Panel A uses all estimates the impact of all relevant stories, Panel B the impact of all relevant stories involving male teachers and Panel C all relevant stories involving female teachers. Reporting the marginal effects at mean after accounting for quadratic terms. All coefficients and standard errors are multiplied by 100 for ease of interpretation. Estimates can be read a percentage change in probability for an additional story. Marginal effect at mean calculated by $\beta_1 + 2\beta_2\overline{s_j}$. Accusation stories are stories involving *Allegations, Being Sued and Criminal Activity*. Relevant Stories include both Extremely and Highly relevant news stories. Stories Regionally is a count for the number of news stories that originated in the region that the teacher resides in the previous 12 months. Standard errors in *italics*, clustered at the regional level.

CHAPTER 5.

ILL COMMUNICATION: TECHNOLOGY, DISTRACTION & WORKPLACE PRODUCTIVITY

I would like to thank my co-author Louis-Philippe.Beland for his contribution. We would like to thank Andriana Bellou, Vincent Boucher, David Card, Andrew Eyles, Baris Kaymak, David Kiss, Ismael Yacoub Mourifie, Daniel Parent, Shqiponja Telj and Felix Weinhardt and seminar participants at RES and University of Montreal for comments and discussions. We would also like to thank Vlad Khripunov and Guillaume Cote for excellent research assistance. Any remaining errors are our own.

1 Introduction

Information Technology is commonly viewed as increasing productivity. This is the reason that both governments and firms invest in research and new technologies. This paper estimates the impact of the removal of a modern piece of Information Technology (IT) on productivity.

Since the development of information technology numerous studies have documented its benefits on productivity in the workplace (Kruger, 1993; Chakraborty and Kazarosian, 1999; Aral et al., 2007; Bartel et al 2007; Ding et al., 2009) and on human capital (Malamud and Pop-Eleches, 2011). However, since Solow’s (1987) infamous statement “You can see the computer age everywhere but in the productivity statistics”, this literature has been paralleled by another that has argued that these IT improvements have not improved productivity (Roach, 1987; Strassmann, 1990; Pischke & DiNardo 1997).

This productivity paradox was summarised by Brynjolfsson (1991) who stated that the lack of findings could be summarised with four reasons. The first two of which relate to the research methods. Firstly mis-measurement, it is hard to get good measures of inputs, outputs and most importantly of the technological innovation itself. Secondly, what are the appropriate time lags for measuring the impact of technology? Should we expect firms to increase productivity within a month, year, or decade? It may take time before the investment into new technology to pay off, so using an incorrect period of analysis could depress any impacts. The second two relate to the adoption of IT in the workplace. The technology may have redistributive effects, making some individuals more productive, but at the expense of others and so there is no net gain. Finally mismanagement, managers do not know how to implement new technology and therefore it is either not fully utilized or that it is only used to create slack instead of increasing productivity.

We are going to discuss a new factor, distraction. With improving communication technology, it has become a lot easier to be distracted. New technologies are increasingly multifunctional, having both a primary purpose and abilities to connect to the internet, social media and play games. This has lowered the transition cost for users shifting from work to procrastination to near zero. Having multifunctional technology present in the workplace is common to many workforces today and so leads to the question; how much productivity is lost through this type of distraction? We will address this question of technology and workplace productivity in the context of the school classroom. We use the banning of mobile phones on school grounds as a restriction on the potential distraction caused by technology.

To do this we focus the experience of students from English classrooms as this setting addresses both the methodological issues highlighted by Brynjolfsson (1991). In England detailed, comparable and externally marked achievement test scores of students are regularly obtained to monitor their progress. This provides us with comprehensive information on the inputs and outputs of individuals and therefore provides a clean measure of individual productivity. Moreover over 90% of teenagers in England owned a mobile phone, therefore a ban will be likely to effect the vast majority of teenagers (E-Marketer 2013). Figure 1 shows the percentage of teenagers and adults who owned a mobile phone in England between 2000 and 2012, showing a steady increase in ownership up to 94% in 2012. Using the removal of a technology rather than it's marginal improvement also sidesteps the issue of how to measure this improvement. Secondly, the debate over suitable lag length is muted as the students affected by the ban can be directly identified and hence the appropriate outputs can be analysed.

The case of teenagers and mobile phones in schools typifies the modern technology productivity paradox as they provide access to chat software, texting, games, social media and the Internet in the classroom. However they could also be used by teachers as an educational aid to improve productivity. There is debate in many countries as to how schools should address the issue of mobile phones. However, this debate has been dominated by the media, with some advocating for a complete ban (Telegraph 2012; Childs, 2013) while others promote the use of mobile phones as a teaching tool in classrooms (Barkham and Moss, 2012; Drury, 2012; O'Toole. 2011; Johnson, 2012; and Carroll, 2013). Despite their prevalence and high profile debate the consequences of mobile phones for high school student performance has not yet been academically studied.

The distracting nature of mobile phones has previously been examined by Bhargava and Pathania (2013), by estimating their use on the incidents of road accidents. They exploit a pricing discontinuity in call plans and show that there is a large jump in phone use post 9pm. However, they find that this jump in phone usage is not followed by an increase in car accidents. Possible explanations put forward for this lack of affect are that; drivers compensate by improving their driving behaviour, a "Peltzman Effect" (Peltzman, 1975); or some are risk loving and so are just substituting away from another distracting activity such as talking to passengers; or that there may be heterogenous effects and that the local estimated effect may be zero but could still be negative for some drivers.

Fairlie and Robinson (2013) conducted the largest-ever field experiment that randomly provides free home computers to students, they find no effects on a wide range of educational outcomes and are precise enough to rule out even modestly-sized positive effects. Looking towards new technology in the classroom Berlinski et al. (2011) implement a large scale randomised control trial in Costa Rica, providing large quantities of computer equipment to rural schools. Despite the large nature of the treatment, they find limited to negative results of the treatment. One interpretation of their results is that when technology is not used appropriately it can be detrimental to student outcomes.

Machin et al (2006) estimate the impact of ICT investment on student outcomes in England, using changes in funding rules as an exogenous shock to investment. They find positive effects of ICT investment on student achievement in English and Science, but not for Mathematics (where computers were rarely used). Combining mobile phones and education, Bergman (2012) used the technology to inform the parents of students of homework assignments through texting. The students of parents who were receiving messages gained in test scores. However, none of these papers have estimated the impact of removing a technology on individual productivity.

In this paper, we estimate the effect of a mobile phone ban on student test scores within schools that implemented them. The lack of consensus of the impact of mobile phones means that there is no government policy relating to their use in schools. This means that schools have complete autonomy in their mobile phone policy, and have differed in how they have reacted. We exploit these differences through a difference in differences estimation strategy, using a two-way fixed-effect model, comparing the gains in test scores across and within schools before and after a phone ban was introduced.

We generate a unique dataset on the history of mobile phone policies from a survey of high schools in the three largest cities in England (Birmingham, London, Manchester) and Leicester, carried out in spring of 2013. This is combined with administrative data of the complete student population from the National Pupil Database (NPD). Given the retrospective nature of the NPD we can establish the academic attainment of students at these schools from 2001 onwards and so can use differences in implementation dates of these policies to measure the impact on student performance. Moreover these data tracks student over time, so that we may account for prior attainment and a set of pupil characteristics including, gender, race, free school meals eligibility, and special educational needs.

We find that after a phone ban the gains in student test scores improve by 5% of a standard deviation (0.05σ). These gains are driven by the most deprived children and that students with high prior attainment are neither positively or negatively affected by their ban. Furthermore, our results indicate that there are no significant gains in attainment if the bans are not widely complied with. Accordingly these estimates are of the average treatment on the treated effect and not an average treatment effect on all schools.

Returning to the productivity paradox literature, these results reflect the redistribution and mismanagement arguments put forward by Brynjolfsson (1991). Redistribution, in that not all students are equally effected by the presence of mobile phones. Our results suggest that non-disadvantaged students are not negatively affected, whilst students eligible for Free School Meals or Special Educational Needs gain significantly once phones are barred from school premises. We also see mismanagement, as given these large gains made by students, schools could have improve test scores earlier by introducing the ban at an earlier date. This paper shows that for schools that enforced bans on mobile phones, the achievement of disadvantaged students significantly improves. Schools could significantly reduce educational inequality by prohibiting mobile phone possession in schools.

The rest of the paper is organized as follows: Section 2 presents the empirical strategy; Section 3 provides a description of the data, survey and descriptive statistics; Section 4 is devoted to the main results, robustness checks and heterogeneity; and Section 5 concludes with policy implications.

2 Empirical Strategy

We estimate the impact of a mobile phone ban, using the differences in timing of the introduction of policies across different schools using a two-way fixed effect model. Equation (1) represents our baseline specification.

$$Y_{ist} = \beta_0 + \beta_1 Ban_{st} + \mu_s + \gamma_t + \varepsilon_{ist} \tag{1}$$

where Y_{ist} is the test score of student i in school s in year t.[1] We assume there are three components to the error which are unobservable; μ is the difference in student attainment due to unobservable school effects; γ represents common shocks to all schools in a particular year; and ε which is the idiosyncratic error and contains all of the variation of individual outcomes within a school year. There may be a concern that only high achieving schools introduced mobile phone bans, and so without accounting for these long-run achievement differences between schools this would lead to an upward bias estimates of the ban. Similarly if there was a positive trend in student test scores, in that they were growing over time and mobile phone bans were only introduced in the later periods, some of this growth would be wrongly attributed to the bans. We can easily account for these two components by allowing for school and year mean achievement to vary through fixed effects. The inclusion of these fixed effects allows for the introduction of mobile phone bans to be non-random e.g. more likely to occur in schools with low test scores, as it allows for covariance between Ban_{st} and μ_s as well as γ_t. Note it does not allow for the effect of the ban to vary across schools or student types.[2]

Ban_{st} is the variable of interest and is an indicator variable for whether school s is prohibiting mobile phones from the premises in time period t. It captures the impact of the introduction of the mobile phone ban on student performance. Accordingly, the coefficient of interest is β_1 which represents the increase in test scores due to the ban. This is estimated using the within school variation in test scores over time.

Another major threat to the basic specification (1) is of pupil sorting to schools according to mobile phone policy. If the most able students changed schools to attend those with bans, again this would lead to an upward bias in the results. Conditioning on pupil characteristics and prior achievement ostensibly accounts for sorting on the basis of observables and a range of unobservables given that prior attainment is a noisy measure of all unobservable and observable inputs up until this point in time for the student e.g. ability, parental investments.

Specification (1) is very restrictive as it does not allow for differences in student outcomes apart from through ε_{ist}. The individual level panel aspect of the NPD allows for us to account for each students prior attainment, which is a large determinant of

[1] We also estimate impacts on achievement level at age 14 and can be seen in Appendix A. These are very similar to our main findings, but with a smaller sample size as we stop the analysis in 2009 as this assessment changed from externally marked to teacher assessed.

[2] Standard errors are clustered at the at the school-year level.

future attainment. Our prior attainment measure, Yist−1 , is student test scores at age 11 which is at the end of primary school and therefore represent all the inputs up until the student starts secondary school. This changes the interpretation of the β_1 parameter from the increase in test scores due to the ban, to the increase in the gains in test scores due to the ban. In addition to prior achievement we also condition on other observable permanent student characteristics, thereby allowing the growth rate of test scores to vary by each of them. Xi represents the vector of student characteristics; SEN, FSM, gender, and ethnicity. The inclusion of these individual controls are to account for student sorting, the extent to which β_1 changes with their inclusion provides us with a gauge for how much sort to school according to phone bans occurs.

$$Y_{ist} = \beta_0 + \beta_1 Ban_{st} + \beta_2 Y_{ist-1} + \beta_3 X_{ist} + \mu_s + \gamma_t + \varepsilon_{ist} \tag{2}$$

A final potential threat to validity arises if there were other positive changes to the school that are correlated with the introduction of the mobile phone ban. Up to this point we had assumed the school effects were invariant over time, if schools introduced other policies that improved test scores at the same time as the phone ban, this again would lead to an upward bias. To address this, we use survey information on whether any leadership or policy changes occurred during the period of analysis and so we control for such changes ($Head_{st}$). This is open to recall bias, but we would expect that head teachers would be very familiar with school-level policies and leadership changes.[3] In our most demanding specification we also account for linear in mean peer effects for each student. We know which students were in the same school year as student i, and it is possible that the other students also effect the growth in attainment of the student through peer effects. Therefore we additionally condition on the mean characteristics and prior attainment of all the other students, in school s in year t, $\bar{X}_{-ist}$. The inclusion of peer characteristics and information on other policy and leadership changes allows us to account for time variant characteristics of the school.

$$Y_{ist} = \beta_0 + \beta_1 Ban_{st} + \beta_2 Y_{ist-1} + \beta_3 X_{ist} + \beta_3 Head_{st} + \beta_5 \bar{X}_{-ist} + \mu_s + \gamma_t + \varepsilon_{ist} \tag{3}$$

[3] Six schools had a change of leadership during this time period. One school had a change in school uniform policy at the same time as the change in mobile policy. Our results are robust to omitting this school from the estimations.

Finally we estimate the heterogeneity of the impact of mobile phone bans by student characteristics in a triple differences framework. β_{1c} is the additional difference in student outcomes by binary student characteristic *c* within schools that had implemented the ban in period *t*. We use our most flexible specification (3) above for these estimates and obtain the additional effect of a ban on: SEN students, FSM students, males, minorities and by achievement level at age 11.

$$Y_{ist} = \beta_0 + \beta_1 Ban_{st} + \beta_{1c} Ban_{st} * Characteristic_i + \beta_2 Y_{ist-1} \\ + \beta_3 X_{ist} + \beta_3 Head_{st} + \beta_5 \bar{X}_{-ist} + \mu_s + \gamma_t + \varepsilon_{ist} \quad (4)$$

3 Administrative and Survey Data

3.1 Administrative Student and Performance Data

The National Pupil Database (NPD) is a rich education dataset of the complete state school population of England. It contains information on student performance, schools attended plus a range of as student characteristics; gender, age, ethnicity, Free School Meals (FSM) eligibility and Special Educational Needs (SEN) status. Each student is allocated an individual identifier which allows for them to be tracked over time and across schools. We generated a dataset which follows students from the end of primary school at age 11 through to the end of compulsory school education at age 16, and so allowing us to condition on prior attainment. Moreover, we use the combination of these characteristics and school attended information to generate peer characteristics, which are a mean of other students characteristics in that school year.

All students in publicly funded schools follow the National Curriculum. This progresses through a series of five Key Stages. Our paper focuses on secondary school students. Students start secondary school at age 11 after completing Key Stage 2 in primary school. Key Stage 3 covers the first three years of secondary school and Key Stage 4 leads to subject-specific exams at age 16, called a General Certificates of Secondary Education (GCSEs).

GCSE test scores from 2001 to 2011 are our main measure of student achievement. Each GCSE is graded from A* to G with an A* being worth 58 points and decreasing in increments of six down to 16 for G grade (details of the coding and their equivalence can be found in DfE, 2011). Students take multiple GCSEs, the mean number of GCES or equivalents taken in the sample is 9. We use a total sum of these GCSE points, standardized each year so that it has mean zero and standard deviation 1. This is to ease interpretation and to account for any grade inflation that may have occurred during this time period.[4]

We also use alternative measures of student performance to examine the robustness of the results, these are 1) a points score which reflects the differences in difficulty in attaining certain grades as recommended by Eyles and Machin (2014)[5]; 2) a standard measure of achievement recognized by employers and in school league tables is if a student achieved five GCSEs at grade C or above including English and Maths; 3) the impact of this student performance at age 14; and 4) a school level measure of the proportion of students achieving 5 GCES at grade C or above including English and Maths, which may reflect gaming on behalf of the school.

3.2 Mobile Phone Survey Data

There is no official policy or recommendation from Department of Education in England to schools regarding mobile phone usage in schools. Therefore schools' mobile phone policies are decided at the school level by the head teacher and the school's governing body, which has resulted in a large variation in mobile phones policies across schools over time.

As information relating to school policies are not collected centrally, in the spring of 2013 we conducted a survey of high schools about their mobile phone policies in four large cities in England; Birmingham, Leicester, London and Manchester. Before approaching the schools we obtained permission from the relevant Local Authorities[6]. Every secondary school from these Local Authorities were then contacted. This consisted of two personalized email, and a follow up phone call seven days after the last, had we not yet received a reply. The email invited the head teacher or school administrator to complete an online survey or reply to the questions via email[7]. The phone call also invited the head teacher to complete it online or immediately over the phone.

[4] Grade inflation would not affect the final results as the inclusion of year effects would account for them. However, standardising by year does make the summary statistics easier to interpret.
[5] This allocated points to GCSEs and their equivalents on a 1-10 scale: A*=10, A=8, C=6, D=4, E=3, F=2, G=1.
[6] From within London we did not obtain permission five Local Authorities (Hackney, Lewisham, Newham, Redbridge and Tower Hamlets) which combined have 77 secondary schools. The City of London Authority does not contain any public schools and therefore was not approached. The remaining 27 London Local Authorities gave permission with 337 secondary schools being approached.
[7] The survey questionnaire is presented in Appendix. Survey website : http://mobilephoneatschool.weebly.com

The survey contained questions about the current policy toward mobile phones, when it was implemented, whether there was a previous mobile phone policy and, if so, when it was implemented. This was repeated until we could construct a complete mobile phone policy history at the school to 2000. These questions were complemented with questions relating to punishments for violating the policy and how well the head teacher considered the policy was complied with. They were also asked if there were any other policy or leadership changes occurring over the same time period, to account for any general shifts in educational policy at the school (Dhuey and Smith, 2013).

We received completed surveys from 90 schools, which represents 21% of the high schools in the four cities in our sample. Table 1 presents statistics on when mobile phone policies were put into effect and how well it was enforced. We define a school as introducing a phone ban if they answered; A) Complete ban of mobile phones on school grounds; or F) Other - Students hand all phones in at the start of school. Headteachers were asked to rate 'to what extent would you say the policy is adhered to by students?' on a seven point scale with 1-'Not at all' to 7 'Completely'. A school was considered to have a high-compliance ban if the response was greater than four. The table shows that the peak years of implementing a ban were between 2005 and 2010, and that most of the bans are complied with.

Table 3 uses the NPD to illustrate the representativeness of the schools in our sample compared to schools in the cities and to England as a whole, over the entire period. Comparing standardized age 16 test scores, we see that schools in these cities score slightly more than the national average but that the schools in our sample achieve significantly higher scores than other schools within these cities (0.069σ). In contrast the cities have slightly lower age 11 achievement than the national average, and that the sampled schools have an even lower intake quality (-0.072σ) although not statically significant at the 10% level. Taken together these imply that the schools in our sample over the 2001-2011 period have a higher gain in test scores than the average school. Despite this the sample schools have a significantly more disadvantaged population than other schools in the city and nationally, having more minority, Special Educational Needs and Free School Meal eligible pupils. There is no difference in the proportion of males students nationally, in the cities or in the sample.

Table A.10 presents descriptive statistics for the same characteristics of the surveyed schools pre- and post-policy introduction. It shows that the average age 16 attainment is significantly higher post-policy compared to pre (0.093σ), but that there was no significant improvement in the prior attainment of the intake students to these schools.

This implies that there is minimal sorting by parents according to mobile phone policy or any other changes that occurred in the school. Other permanent student characteristics change slightly pre- and post-ban, with a 5.3% decrease in the proportion of minority students and a 5.2% and 6.8% increase in the proportion of SEN and FSM students. As these variable are not standardized each year these differences could be reflecting general trends in the population. Once the changes over time and the difference across schools are taken into account there are no significant differences in these variables before and after the ban was introduced.[8]

4 Results

4.1 Main Results

Table IV presents estimates of the impact of a mobile phone ban on individual student performance in five specifications which account for more potential biases moving from left to right. Column 1 is the most basic specification that only accounts for the across school and across year mean differences in test scores. Here we find that the introduction of a mobile phone ban improves student test scores by 5.93% of a standard deviation.

However we still may be concerned that student sorting by observable or unobservable characteristics may be driving this estimate. The columns 2 and 3 successively include student characteristics in order to account for this. Conditioning on prior attainment indicates that the growth in test scores is 0.064σ and that this marginally increases when additionally controlling for student characteristics to 0.067σ. These estimates did not change significantly from the basic specification implying that sorting is not driving the results.

The last two columns account for time varying school characteristics. Including an indicator variable which denotes if there was a leadership change at the school in year t and onwards, decreases the estimate[9]. Results of our preferred specification (5), which allow for linear in mean peer effects, continue show an improvement in student performance after a school bans mobile phones. After a ban has been introduced an

[8] Appendix Table estimates these variables on an indicator variable for if a policy has been introduced at that school conditional on year and school effects. Each characteristics is tested separately and none were found to be significantly correlated.

[9] The coefficient for leadership changes is large and positive, which is what would be expected if new head teachers brought in new beneficial policies/management. β_3=20.4

average student attending that schools has 5.06% of standard deviation higher gains in test scores compared to a school which did not introduce the ban.

These estimates assume the ban to have constant impact after it was introduced through time. Figure 3 relaxes this assumption and also allows for us to check for pre-trends in student attainment before the introduction of the ban. This plots the impact of the ban by exposure length, with the reference year being the year prior to the phone ban being introduced. Estimates with negative exposure implies the years prior the ban where we would not expect to find and impact of the ban. Using our most preferred specification we find a significant impacts of the ban after two years. There is a general upward trend in the impact of the ban, this reflects that these students have experienced more time in a school which has had a phone ban in place.[10] Moreover there is little evidence that the schools were generally improving before the phone introduction as all the years prior to the ban do not have impacts significantly different from zero.

4.2 Placebo and robustness checks.

Before exploring the heterogeneity of the effects, it is important to test a key assumption of the model. We are obtaining unbiased estimates of β_1 as long as we have $Cov(Ban_{ist}, \varepsilon_{ist}) = 0$. If schools that introduced a mobile phone ban were improving regardless then these gains could be falsely attributed to the policy. Whilst this is partially addressed with the event study show in Figure 3, we test this directly in two different ways. Firstly with a placebo treatment and secondly testing for common trends amongst schools that introduced a ban in different years.

The placebo ban is generated by imposing that the ban occurred two years before it was actually initiated. This placebo intervention should have no significant impact on the gains in student test scores. If there is a significant relationship then there are correlations between the trend and the intervention. Table V presents a parallel set of results as Table IV, but with the effects of a placebo intervention. The analysis is restricted up until the actual ban was introduced, thereby allowing for two years of artificial growth in these schools. We find that there are no significant gains in test scores associated with a placebo ban.

Unlike a traditional difference in differences setting, the vast majority of the schools receive the treatment. This would cause a problem when attempting to show common

[10] Estimations that directly estimated this additional positive trend failed to find a significant relationship.

trends for the treated and control groups as all expect two schools would form the treatment group. Instead we compare the pre-treatment trends of schools that adopted the bans in different years. This is treating each school who adopted the ban in a specific year as a different treatment group, and we test that the early adopters have similar trends to the late adopters. If schools that adopted the policy in different years have similar pre-trends then we can be more confident in our results. Each school was assigned a cohort dependent on the first year they introduced a ban (Cohort). This cohort indicator was then interacted with a year trend variable, and estimated on age 16 test scores before the ban was in place in that school.

$$Y_{ist} = \beta_0 + \sum_{c=1}^{12} \beta_c Year_t * Cohort_s + \beta_2 Y_{ist-1} + \beta_3 X_{ist}$$
$$+\beta_3 Head_{st} + \beta_5 \bar{X}_{-ist} + \mu_s + \gamma_t + \varepsilon_{ist} \quad (3)$$

We test for significant differences in the pre-trends e.g.$\beta_{c=1} \neq \beta_{c=2}$. We find no significant differences in pre trends (Table A.2).

Thus far we have use age 11 test scores as a measure of prior achievement for attainment at age 16. However there is another statutory exam that takes place between these ages. We replicate Table 4 in appendix Table A.3 instead using achievement at age 14. This has the advantage that it is a more recent measure of student ability, but has the disadvantage that these exams are conducted at secondary school and therefore could also be affected by the ban. Therefore we only use the age 14 test scores of students attending schools who have not yet implanted a ban. As there is only two years between the age 14 and age 16 exams this reduces the sample significantly, but also examines the short run impact of phone bans. The estimates are very similar to before with a mobile phone ban improving test scores by 5.3% of a standard deviation. These results also in part address the issue of pre-trends, as we see that there are significantly larger gains between 14 and 16 in test scores for students who were attending schools who introduced a ban during that time. This is a small window of time for other effects to incur. If a positive trends were in place in schools previous to this, the age 14 tests scores would be higher and then the gains in test scores would be accordingly lower.

One may be additionally concerned that these results are dependent on the outcome measure that we are using, therefore for the remainder of this section replicates the previous results using a different set of outcome variables to establish the robustness of the estimates. The age 16 measure of achievement used so far in this paper is the standardized point score over all exams taken at the end of compulsory schooling using the scoring system from the official government league tables. An alternate scoring system which accounts for the different difficulties for attaining grades finds very similar results and associated tables can be found in the appendix (Tables A.5, A.6 and A.7).

However, there is another measure that is widely used by the government and employers, this is whether the student achieved at least five GCSEs above a C including English and Maths. We derive a binary variable representing this for each student in our sample. This is used as the outcome of interest in the same specifications, and so assumes a linear probability model. In our most demanding specification, we find that a ban improves the probability of a student attaining 5-GCSEs at A*-C level by 1.92% points on a baseline of 38% students in our sample attaining this level (Table A.8). Finally we present equivalent results at the school level, using the proportion of students achieving this level (Table A.10), which shows schools improve after the introduction of the ban.

Overall, results are robust to alternative specifications and set of student characteristics including different measures of prior achievement and peer effects. These numerous robustness checks provide confidence that mobile phone bans play a role in determining school and student performance.

4.3 Heterogeneity

Table VI studies the heterogeneity of a ban on students with different characteristics, under a triple differences frameworks, estimating the additional impact on SEN, FSM, male, minority students and by prior attainment. This is in addition to any baseline effects of the ban under specification 5. The results indicate that a mobile phone ban has a positive and significant impact on FSM-eligible students (column 1), SEN students (column 2) and males (column 3). In columns (1), (2) and (3) the baseline effect of a mobile phone ban is not significant when controlling for student characteristics and ban interaction, which indicates that results are driven by certain students and that not all students are significantly affected by mobile phone bans.

The interaction of the ban with prior achievement is negative (column 5), implying that it is predominantly low ability students who are gaining from the ban. Prior achievement is measured in percentiles, with the top percentile having a score of 1. Therefore the coefficient of -5.89 means that students in the top percentile nationally would lose 0.059σ with the introduction of a ban compared to a student at the bottom. However there is a general positive effect of the ban of 0.058σ and so overall are not harmed by the ban. This is tested formally in the next table. Column 7 additionally includes the interactions with ability, FSM and SEN simultaneously and we find that all three interaction terms are significant. This is in line with the heterogeneity results, with the most at risk students gaining the most.

Table 7 examines the linearity of the impact of mobile phone bans by prior achievement in more detail. Students are grouped into five quantiles based on their achievement level at age 11, where group 1 means lowest achievement group and group 5 is the highest achievement group. This time the coefficients are representing the total effect of the ban on an ability quintile, as the main effect of the ban is not included. Again we see that the low achieving students are gaining the most from the ban. Those in the lowest quintile gain 0.133σ more after a ban has been introduced. Only the top two quintiles do not significantly gain from the policy, but they are also not negatively affected.

One would expect the impact of a mobile phone ban to vary according to how well it was enforced. We replace the single Ban_{st} variable with two one for bans with high compliance $Ban - High_{st}$ and one with low compliance $Ban - Low_{st}$, as reported by the head teacher. Table 7 shows estimates impact of the ban by compliance, as expected we find much larger and significant effects in schools that reported a high compliance to the ban compared to where it was not enforced.

The heterogeneity of these results are replicated conditional on age 14 ability. Table A.4, shows the estimates by ability have slightly smaller positive effect for the least able students, but these effects are not significantly different to before. Tables A.6 and A.7 also replicate the heterogeneity using the alternative age 16 test score measures. Finally Figure 3 shows the density of standardized student test scores before and after a mobile phone ban. It shows that the density of test scores shifts right after the imposition of a ban.

5 Conclusions

Modern information communication technology has the potential to improve individual productivity, however given it's multifunctional abilities can also be distracting and reduce productivity. Whether the gains outweigh the losses is a question facing many workplaces today. This paper examines the impact on productivity from removing a common form of information technology (mobile phones) in an education setting where individual inputs and outputs can easily be measured.

We combine survey data on mobile phone policies in schools in four cities with administrative data on student achievement to create a history of student attainment in schools from 2002-2011. Using a two way fixed effects estimations, we estimate the effect of mobile phone bans on student performance, using the variation in implementation dates. Balancing tests find no difference in pre-policy trends across early and late adopters.

Our results indicate that in schools that have introduced a mobile phone ban there is an increase in student performance of 5.1% of a standard deviation. We find that banning mobile phones improves the outcomes of the low-achieving students (12.4%) the most and has no significant impact on high achievers. The results suggest that low-achieving students are more likely to be distracted by the presence of mobile phones, while high achievers can focus in the classroom regardless of the mobile phone policy. Given heterogeneous results, banning mobile phones could be a low cost way for schools to reduce educational inequality.

References

Aral, S., Brynjolfsson, E. and Van Alstyne, M. (2007). Information, technology and information worker productivity: Task level evidence. *National Bureau of Economic Research* Working Paper.

Barkham, P. and Moss, S. (November 2012). Should mobile phones be banned in schools. The Guardian.

Bartel, A., Ichniowski,C. and Shaw, K. (2007). How does information technology affect productivity? plant-level comparisons of product innovation, process improvement, and worker skills. *The Quarterly Journal of Economics*, 122(4), pp.1721–1758.

Bergman, P. (2012). The more you know: Evidence from a field experiment on parent-child information frictions and human capital investment. *Working Paper* – Teachers College Columbia

Berlinski, S., M. Busso, J (2013). Pedagogical Change in Mathematics Teaching: Evidence from a Randomized Control Trial. *Columbia Working Papers*

Bhargava, S. (2013). Driving under the (cellular) influence. *American Economic Journal: Economic Policy* 5(3), 92–125.

Brynjolfsson, E. (1991). Information technology and the re-organization of work: theory and evidence. *Ph.D. Dissertation*, MIT Sloan School of Management, 1991

Carroll, D. (September 2013). Cleveland teacher uses cell phones as classroom learning tools. WRCB.

Chakraborty, A. and Kazarosian, M. (1999). Product differentiation and the use of information technology: New evidence from the trucking industry.

DfE (2011) Test and Examination Point Scores Used in The 2011 School and College Performance Tables. *Department for Education* 2011. Last Accessed May 27th 2014 http://www.education.gov.uk/schools/performance/secondary_11/PointsScoreAllocation2011.pdf

Dhuey, E. and Smith, J. (2013). How important are school principals in the production of student achievement? *Canadian Journal of Economics* Forthcoming

DiNardo, J. E. and Pischke J.-S. (1997). The returns to computer use revisited: Have pencils changed the wage structure too? *The Quarterly Journal of Economics* 112(1), pp.291–303.

Ding, W. W., Levin, S. G. Stephan, P. E. and Winkler A. E. (2009). The impact of information technology on scientists' productivity, quality and collaboration patterns. *Management Science*, 56(9) pp.1526-5501.

Drury, E. (September 2012). Mobile phones in the classroom: teachers share their tips. The Guardian, Guardian Professional.

Eyles, A. and Machin, S. (2014), 'Academy schools and their introduction to English education', Centre for Economic Performance, mimeo.

Fairlie, R. W. and Robinson,J. (2013). Experimental evidence on the effects of home computers on academic achievement among schoolchildren. American Economic Journal: Applied Economics 5(3), 211–240.

Johnson, S. (October 2012). Teachers 'tactically' ignore mobile phone use in classroom. The Telegraph.

Krueger, A. B. (1993). How computers have changed the wage structure: evidence from microdata,1984:1989. *The Quarterly Journal of Economics* 108(1), pp.33–60

Machin, S., McNally, S., and Silva, O. (2007). New technology in schools: Is there a payoff?. *The Economic Journal*, 117(522), pp.1145–1167.

Malamud, O. and Pop-Eleches, C. (2011). Home computer use and the development of human capital. *The Quarterly Journal of Economics* 126(2), pp.987–1027.

Murphy, R. and F. Weinhardt (2013). Top of the class: The importance of rank position. *Centre for the Economics of Education* Discussion Paper 1241.

Ofcom (2011). A nation addicted to smartphones. Ofcom report.

O'Toole, M. (2011). Toronto district school board to lift cell phone ban. *National Post*.

Roach, S. S. (1987). The technology slump: Dim prospects from the bottom. *Economics Perspective Memorandum.*

E-Marketer (2013). UK teens far outshine us counterparts in smartphone usage. E-Marketer newsletter.

Solow, R. (1987), We'd better watch out. New York Times (July 12) Book Review, 36.

Strassmann, P. A. (1990). The business value of information technology. Strassmann Inc, New Canaan, Conn.

Table 1: Mobile Phone Ban Policies in Effect over Time

Year	Mobile Bans	High-compliance Bans	Low-compliance Bans
2000	0	0	0
2001	0	0	0
2002	3	2	1
2003	6	5	1
2004	9	7	2
2005	19	13	6
2006	29	20	9
2007	43	31	12
2008	58	38	20
2009	71	47	24
2010	85	54	31
2011	88	55	33
2012	90	56	34

Source: School Survey. *Notes:* Mobile Phone implementation each year. Headteachers were asked what their phone policy is and when it was introduced. A phone ban is classified as 1) A complete ban of mobile phones on school grounds; or 2) Students hand all phones in at the start of school. Headteachers were asked to rate 'to what extent would you say the policy is adhered to by students?' on a seven point scale with 1-'Not at all' to 7 'Completely'. A school was considered to have a high-compliance ban if the response was greater than four. Source: Mobile phone policy survey of schools in four cities in England: Birmingham, Leicester, London and Manchester

Table 2: Descriptive Statistics - Representativeness of Sample

Student Characteristics	All Students in English	Students in Sampled Cities	Students in Responding Schools	Difference Between Responding Schools and Cities
Test Scores: Age 16	0.00	0.01	0.07	0.07*
	(1.00)	(1.02)	(0.94)	(0.04)
Test Scores: Age 11	0.00	-0.03	-0.09	-0.07
	(1.00)	(1.01)	(1.01)	(0.04)
Male	0.51	0.50	0.47	-0.04
	(0.50)	(0.50)	(0.50)	(0.03)
Minority	0.48	0.66	0.74	0.10***
	(0.50)	(0.48)	(0.44)	(0.03)
SEN	0.15	0.17	0.18	0.02**
	(0.35)	(0.37)	(0.38)	(0.01)
FSM	0.16	0.24	0.31	0.09
	(0.37)	(0.43)	(0.46)	(0.02)
Total Students	5,576,276	789,638	130,482	

Table 2 presents descriptive statistics on key variables for all schools, schools in city surveyed, schools in sample and difference between schools in sample and in city surveyed. SEN means Special Educational needs student and FSM means Free School Meal students. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table 3: Descriptive statistics of Pre and Post Policy

Student Characteristics	Students in Responding Schools	Pre Phone Ban	Post Phone Ban	Pre-Post Difference	No Ban
Test Scores: Age 16	0.07	0.02	0.12	0.09**	0.14
	(0.94)	(0.96)	(0.92)	(0.04)	(0.93)
Test Scores: Age 11	-0.09	-0.11	-0.08	0.02	0.02
	(1.01)	(1.01)	(1.01)	(0.04)	(0.95)
Male	0.47	0.47	0.47	-0.00	0.53
	(0.50)	(0.50)	(0.50)	(0.02)	(0.50)
Minority	0.74	0.77	0.72	-0.05**	0.79
	(0.44)	(0.42)	(0.45)	(0.03)	(0.41)
SEN	0.18	0.15	0.21	0.05***	0.20
	(0.38)	(0.36)	(0.40)	(0.01)	(0.40)
FSM	0.31	0.28	0.35	0.07***	0.25
	(0.46)	(0.45)	(0.48)	(0.02)	(0.43)
Total Students	130,482	62,214	66,266		2002

Notes: Table 3 presents descriptive statistics on key variable Pre and Post policy and for all schools and schools in city surveyed. SEN means Special Educational needs student and FSM means Free School Meal students. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table 4: Effect of Mobile Bans on Age 16 Student Performance

Age 16 Test Scores	(1)	(2)	(3)	(4)	(5)
Mobile Ban	5.93**	6.35**	6.70**	6.90**	5.06*
	2.91	*2.91*	*2.94*	*2.92*	*2.86*
Prior Test Scores: Age 11		✓	✓	✓	✓
Student characteristics			✓	✓	✓
Leadership changes				✓	✓
Peer characteristics					✓
School effects	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓
Observations	130,595	130,595	130,595	130,595	130,595
Adj -R-squared	0.148	0.415	0.455	0.456	0.460

Note: Table 4 presents regression estimates for student performance. Outcome variable is standardized test score at age 16. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school-year level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status. Key Stage 2 is standardized student test score at age 11 (before high school). *** p<0.01, ** p<0.05, * p<0.1. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table 5: Effect of Placebo Mobile Bans on Age 16 Student Performance

Age 16 Test Scores	(1)	(2)	(3)	(4)	(5)
Panel A: Placebo Only					
Placebo Mobile Ban	2.88	2.96	3.08	4.02	3.50
	3.20	3.14	3.22	3.21	3.17
Panel B: Placebo and Actual Ban					
Placebo Mobile Ban	2.45	2.48	2.48	3.44	3.12
	3.19	3.12	3.2	3.19	3.16
Mobile Ban	5.49***	6.01*	7.58**	6.89**	4.81*
	2.91	2.91	2.94	2.92	2.87
Prior Test Scores: Age 11		✓	✓	✓	✓
Student characteristics			✓	✓	✓
Leadership changes				✓	✓
Peer characteristics					✓
School effects	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓
Observations	130,595	130,595	130,595	130,595	130,595
Adj -R-squared (Panel A)	0.147	0.414	0.452	0.453	0.457

Note:Table 5 presents regression estimates for student performance. Placebo ban is introducing the ban two years before it was actually introduced. Outcome variable is standardized test score at age 16. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status.. Key Stage 2 is standardized student test score at age 11 (before high school). *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table 6: The Effect of Mobile Phone Bans on Student Performance by Student Characteristics

Age 16 Test Scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mobile Ban	3.98	4.26	4.17	8.96***	5.87**	4.49	3.61
	(2.98)	(2.99)	(2.95)	(3.43)	(2.91)	(2.96)	(3.02)
Mobile Ban * FSM	7.37***					4.88**	4.63**
	(2.02)					(1.94)	(1.93)
Mobile Ban * SEN		10.89***					5.85**
		(2.36)					(2.39)
Mobile Ban * Male			4.13**				
			(2.10)				
Mobile Ban * Minority				-4.90*			
				(2.63)			
Mobile Ban * Prior Test					-5.89***	-5.42***	-4.55***
Scores: Age 11					(1.05)	(1.02)	(1.06)
Prior Test Scores: Age 11	✓	✓	✓	✓	✓	✓	✓
Student characteristics	✓	✓	✓	✓	✓	✓	✓
Leadership changes	✓	✓	✓	✓	✓	✓	✓
Peer characteristics	✓	✓	✓	✓	✓	✓	✓
School effects	✓	✓	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓	✓	✓
Observations	130,595	130,595	130,595	130,595	130,595	130,595	130,595
Adj R-Squared	0.495	0.495	0.495	0.495	0.497	0.498	0.498

Note: Table 6 presents regression estimates for student performance. Outcome variable is standardized test score at age 16. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status.. Key Stage 2 is standardized student test score at age 11 (before high school). *** p<0.01, ** p<0.05, * p<0.1. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table 7: The Effect of Mobile Phone Bans on Student Performance by Prior Attainment

Age 16 Test Scores	(1)	(2)	(3)	(4)
Impact by age 11 Test Scores				
Mobile Ban* 1st Quintile	13.29***	14.40***	13.68***	12.36***
	(3.11)	(3.14)	(3.13)	(3.05)
Mobile Ban* 2nd Quintile	8.59***	9.94***	9.30***	8.02***
	(3.07)	(3.11)	(3.08)	(3.05)
Mobile Ban* 3rd Quintile	5.78*	6.72**	6.22**	4.97
	(3.12)	(3.15)	(3.13)	(3.07)
Mobile Ban* 4th Quintile	2.98	2.68	2.14	1.06
	(3.15)	(3.18)	(3.16)	(3.11)
Mobile Ban* 5th Quintile	-0.82	-1.97	-2.63	-2.70
	(3.42)	(3.47)	(3.49)	(3.38)
Test Scores: Age 11 Categorical	✓	✓	✓	✓
Student characteristics		✓	✓	✓
Leadership changes			✓	✓
Peer characteristics				✓
School effects	✓	✓	✓	✓
Year effects	✓	✓	✓	✓
Observations	130,595	130,595	130,595	130,595
Adj R-Squared	0.422	0.442	0.453	0.457

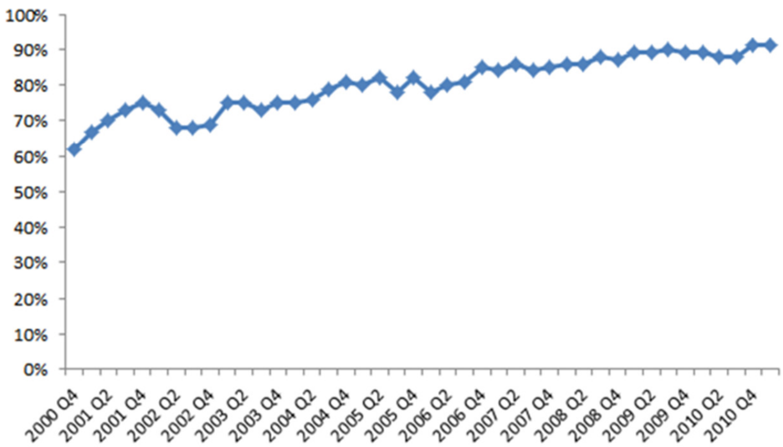
Note: Table 7 presents regression estimates for student performance. Outcome variable is standardized test score at age 16. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status.. Key Stage 2 is standardized student test score at age 11 (before high school). In this table, student performance at age 11 (Key Stage 2) are grouped in 5 category based on their achievement level at age 11, where group 1 means lowest achievement group and group 5 are highest achievement group. *** p<0.01, ** p<0.05, * p<0.1. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table 8: The Effect of Mobile Phone Bans on Student Performance by Ban Compliance

Age 16 Test Scores	(1)	(2)	(3)	(4)	(5)
High Compliance- Mobile Ban	6.42**	6.81**	7.12**	7.45**	6.60*
	(3.00)	(2.99)	(3.03)	(3.04)	(2.96)
Low Compliance- Mobile Ban	2.12	2.67	3.42	5.87	8.60
	(6.58)	(6.53)	(6.56)	(5.29)	(5.40)
Prior Test Scores: Age 11		✓	✓	✓	✓
Student characteristics			✓	✓	✓
Leadership changes				✓	✓
Peer characteristics					✓
School effects	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓
Observations	130,595	130,595	130,595	130,595	130,595
Adj R-Squared	0.168	0.427	0.456	0.460	0.466

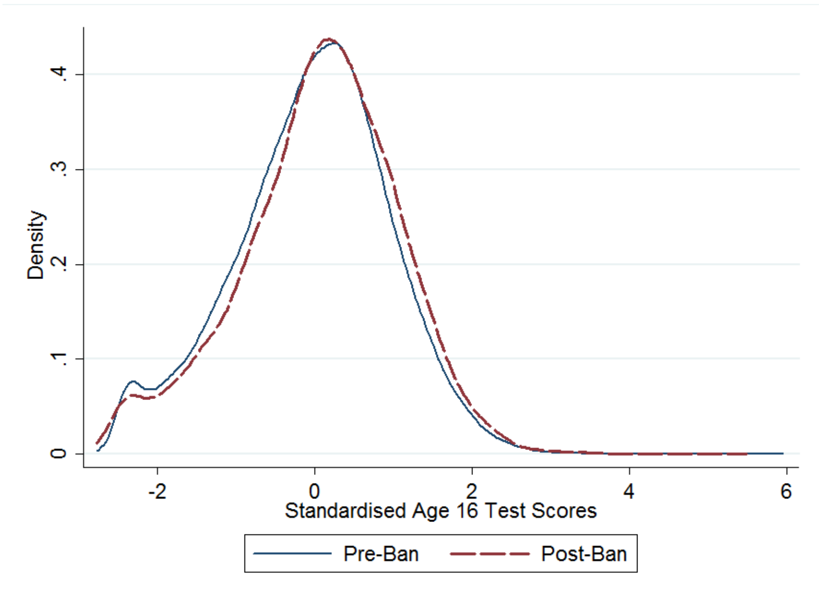
Note: Table 8 presents regression estimates for student performance. It separates ban into high-compliance (principal assessment score above or equal to 4 out of 7) and low-compliance mobile ban. Outcome variable is standardized test score at age 16. We use robust clustered standard error at the school-year level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status.. Key Stage 2 is standardized student test score at age 11 (before high school). *** p<0.01, ** p<0.05, * p<0.1 Source: National Pupil database (NPD) and mobile phone policy from survey.

Figure 1: Mobile Phone Ownership Rates in England



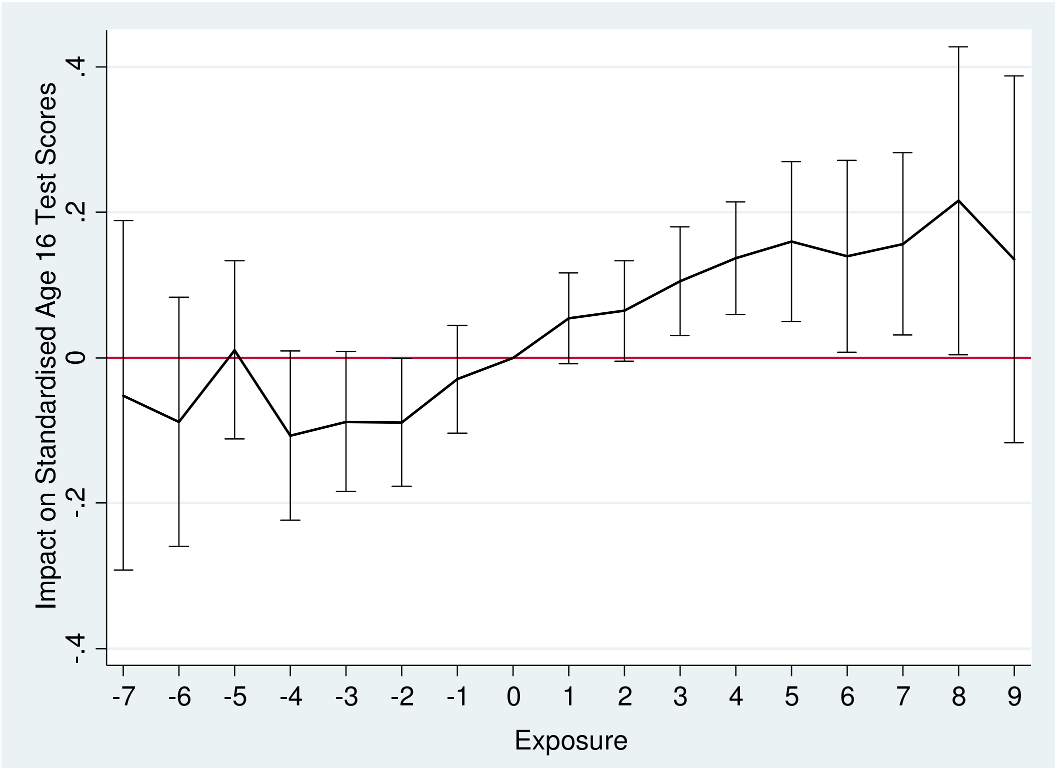
Notes: Phone ownership rates in England amongst individuals 13 years and over. Source: Oftel/ Ofcom Based on face to face survey data, 2011

Figure 2: Age 16 Student Test Scores density Pre and Post Phone Ban



Notes: Density of standardised age 16 test scores for all years in sample Pre/Post Phone Ban. Schools which never introduced a ban not included. Source: National Pupil Data Base

Figure 3: Impact of Phone Ban by Years of Exposure



Source: National Pupil Data Base, School Survey
Notes: Estimated impact of time since the year prior to policy conditional on school, and year effects, prior test scores, pupil characteristics, leadership changes and peers effects. Reference year is the year prior to introduction. Error bars represent the 95% confidence intervals with robust standard errors clustered and the school year level.

Appendix 1 - School Survey

Mobile Phone survey questionnaires

Question 1.1) What best describes the school's current mobile phone policy?

a) Complete ban of mobile phones on school grounds

b) Allowed on grounds, but must be turned off

c) Allowed on grounds, but must be turned to silent and off during classes

d) Allowed on grounds, but must be turned to silent at all times

e) Allowed on grounds, but must be considerate with use

f) Other -Yes

g) None

Question 1.2) If Other, could you please briefly describe current policy.

Note: Only Answer: Hand into reception, and collected at end of day.

Question 1.3) When was the current policy first introduced?

Question 1.4) What are the punishments for misuse of phones on school grounds?

Question 1.5) Out of 7 to what extent would you say the policy is adhered to by students? [With 7 being 'Completely' and 1 being 'Not at all']

Question 2) Was there a different policy in place before this? - Yes/No

If Yes, please answer the following.

If No please skip to question 4.

In the space below please answer questions 1.2 to 1.5 for this pervious policy (brief description of policy/introduction date/punishments/adherence).

Question 3) Was there a different policy in place before this? - Yes/No

If Yes, please answer the following.

If No please skip to Question 4.

In the space below please answer questions 1.2 to 1.5 for this pervious policy (brief description of policy/introduction date/punishments/adherence).

Question 4) Were there any other policy or leadership changes at the same time as the mobile policy change?

Question 5) Do you have any other comments?

Table A.1: Balancing Test

Variables	Prior Attainment	Male	Minority	SEN	FSM
Mobile Ban	-.074	-0.39	0.01	0.81	1.07
	(1.25)	(0.44)	(0.72)	(1.00)	(0.70)
Observations	130,595	130,595	130,595	130,595	130,595

Note: Table A.1 presents regression estimates for different outcome variables to investigates if schools that impose a ban are different and if students are sorting into schools based on student characteristics. SEN means fraction of Special Educational Needs student, FSM means fraction Free School Meal students. Key Stage 2 means standardized average test score at age 11 of the school. Male and Minority are fraction of students that are male and from a minority group respectively. We use robust clustered standard error at the school level with school and year fixed effect. *** p<0.01, ** p<0.05, * p<0.1 Source: National Pupil database (NPD) and mobile phone policy from survey.

Table A.2: Descriptive statistics on School Growth Trends by Adoption Cohort

Age 16 Test Scores	(1)	(2)
2003 Cohort	-1.21	-2.28**
	(1.11)	(0.91)
2004 Cohort	-1.23	-2.30**
	(1.11)	(0.91)
2005 Cohort	-1.23	-2.29**
	(1.11)	(0.91)
2006 Cohort	-1.24	-2.30**
	(1.11)	(0.91)
2007 Cohort	-1.23	-2.29**
	(1.11)	(0.91)
2008 Cohort	-1.23	-2.30**
	(1.11)	(0.91)
2009 Cohort	-1.23	-2.29**
	(1.11)	(0.91)
2010 Cohort	-1.23	-2.29**
	(1.11)	(0.91)
2011 Cohort	-1.18	-2.25**
	(1.11)	(0.91)
Student Characteristics		✓
School Effects	✓	✓
Year Effects	✓	✓
Observations	62,214	62,214

Table A.2 compares the average growth rate in schools before the introduction of a mobile phone ban by year the ban was introduced. The standard errors are clustered by school year. Student characteristics are Key Stage 2 test scores and student's gender, ethnicity, SEN and FSM eligibility. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table A.3: The Effect of Mobile Bans on Student Performance Conditioning on age 14 Test Scores

Age 16 Test Scores	(1)	(2)	(3)	(4)	(5)
Mobile Ban	6.86**	5.51**	6.14**	6.38**	5.30**
	(2.71)	(2.59)	(2.60)	(2.64)	(2.67)
Prior Test Scores: Age 14		✓	✓	✓	✓
Student characteristics			✓	✓	✓
Leadership changes				✓	✓
Peer characteristics					✓
School effects	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓
Observations	83,211	83,211	83,211	83,211	83,211

Note: Table A.3 presents regression estimates for student performance. Outcome variable is standardized test score at age 16 and control for standardized test score at age 14. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school-year level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status. Key Stage 2 is standardized student test score at age 11 (before high school). *** p<0.01, ** p<0.05, * p<0.1. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table A.4: The Effect of Mobile Bans on Student Performance by Age 14 Prior Achievement Quintile

Age 16 Test Scores	(1)	(2)	(3)	(4)
Impact by age 14 Test Scores				
Mobile Ban* 1st Quintile	10.43***	11.21***	11.39***	10.06***
	(3.05)	(3.02)	(3.04)	(3.03)
Mobile Ban* 2nd Quintile	9.28***	10.58***	10.79***	9.64***
	(3.11)	(3.08)	(3.12)	(3.16)
Mobile Ban* 3rd Quintile	5.53*	6.21**	6.44**	5.20
	(3.05)	(3.08)	(3.13)	(3.13)
Mobile Ban* 4th Quintile	2.49	2.67	2.92	1.75
	(3.09)	(3.09)	(3.13)	(3.14)
Mobile Ban* 5th Quintile	-0.22	0.39	0.68	-0.07
	(3.42)	(3.47)	(3.49)	(3.49)
Test Scores: Age 14 Categorical	✓	✓	✓	✓
Student characteristics		✓	✓	✓
Leadership changes			✓	✓
Peer characteristics				✓
School effects	✓	✓	✓	✓
Year effects	✓	✓	✓	✓
Observations	83,211	83,211	83,211	83,211

Note: Table A.4 presents regression estimates for student performance. Outcome variable is standardized test score. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school-year level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status. Age 14 test scores are Key Stage 3 results before they were teacher assessed (2008/9)/ In this table, results are grouped in 5 category based on their achievement level (Key Stage 3) at age 14, where group 1 means lowest achievement group and group 5 are highest achievement group. *** p<0.01, ** p<0.05, * p<0.1 Source: National Pupil database (NPD) and mobile phone policy from survey.

Table A.5: The Effect of Mobile Bans on Alternate Age 16 Student Performance

Age 16 Alternate Test Scores	(1)	(2)	(3)	(4)	(5)
Mobile Ban	5.57**	6.01**	6.33**	5.77**	4.81**
	(2.69)	(2.68)	(2.71)	(2.69)	(2.63)
Prior Test Scores: Age 11		✓	✓	✓	✓
Student characteristics			✓	✓	✓
Leadership changes				✓	✓
Peer characteristics					✓
School effects	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓
Observations	130,595	130,595	130,595	130,595	130,595
Adj R-Squared	0.159	0.464	0.499	0.500	0.504

Note: Table A.5 presents regression estimates for student performance. Outcome variable is standardized test score at age 16 accounting for differences in difficulty of attaining grade. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school-year level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status. Key Stage 2 is standardized student test score at age 11 (before high school). *** p<0.01, ** p<0.05, * p<0.1. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table A.6: The Effect of Mobile Bans on Alternate Age 16 Student Performance by Student Characteristics

Age 16 Alternate Test Scores	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mobile Ban	3.82	4.40	4.20	8.33***	5.58**	4.27	3.75
	(2.74)	(2.76)	(2.74)	(3.11)	(2.68)	(2.72)	(2.78)
Mobile Ban * FSM	6.80***					4.61***	4.46**
	(1.86)					(1.78)	(1.78)
Mobile Ban * SEN		8.13***					3.44
		(2.10)					(2.15)
Mobile Ban * Male			3.37*				
			(1.95)				
Mobile Ban * Minority				-4.37*			
				(2.43)			
Mobile Ban * Prior Test Scores: Age 11					-5.21***	-4.76***	-4.31***
					(0.99)	(0.96)	(1.01)
Prior Test Scores: Age 11	✓	✓	✓	✓	✓	✓	✓
Student characteristics	✓	✓	✓	✓	✓	✓	✓
Leadership changes	✓	✓	✓	✓	✓	✓	✓
Peer characteristics	✓	✓	✓	✓	✓	✓	✓
School effects	✓	✓	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓	✓	✓
Observations	130,595	130,595	130,595	130,595	130,595	130,595	130,595
R-Squared	0.500	0.500	0.500	0.500	0.502	0.501	0.501

Note: Table A.6 presents regression estimates for student performance. Outcome variable is standardized test score at age 16 accounting for differences in difficulty of attaining grade. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school-year level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status. Key Stage 2 is standardized student test score at age 11 (before high school). *** p<0.01, ** p<0.05, * p<0.1. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table A.7: The Effect of Mobile Bans on Alternate Age 16 Student Performance by Prior Age 14 Achievement Quintile

Age 16 Alternative Test Scores	(1)	(2)	(3)	(4)
Impact by age 14 Test Scores				
Mobile Ban* 1st Quintile	11.05***	12.11***	11.44***	10.22***
	(2.85)	(2.88)	(2.86)	(2.79)
Mobile Ban* 2nd Quintile	9.03***	10.30***	9.72***	8.53***
	(2.84)	(2.87)	(2.84)	(2.81)
Mobile Ban* 3rd Quintile	6.00**	6.89**	6.44**	5.27*
	(2.9)	(2.92)	(2.91)	(2.86)
Mobile Ban* 4th Quintile	2.86	2.56	2.06	1.05
	(2.94)	(2.96)	(2.94)	(2.9)
Mobile Ban* 5th Quintile	-0.78	-1.87	-2.47	-2.57
	(3.21)	(3.25)	(3.27)	(3.18)
Test Scores: Age 14 Categorical	✓	✓	✓	✓
Student characteristics		✓	✓	✓
Leadership changes			✓	✓
Peer characteristics				✓
School effects	✓	✓	✓	✓
Year effects	✓	✓	✓	✓
Observations	130,595	130,595	130,595	130,595
R-Squared	0.442	0.482	0.483	0.487

Note: Table A.7 presents regression estimates for student performance. Outcome variable is standardized test score in student 8 best subject. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school-year level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status. Key Stage 2 represents test score at age 11. In this table, results are grouped in 5 category based on their achievement level at age 11, where group 1 means lowest achievement group and group 5 are highest achievement group. *** p<0.01, ** p<0.05, * p<0.1. Source: National Pupil database (NPD) and mobile phone policy from survey.

Table A.8: The Effect of Mobile Bans on Student Performance Probability of Achieving 5 GCSES Including English and Maths

Age 16 Alternate Test Scores	(1)	(2)	(3)	(4)	(5)
Mobile Ban	1.98**	2.17**	2.24**	2.23**	1.92**
	(0.93)	(0.91)	(0.91)	(0.92)	(0.89)
Prior Test Scores: Age 11		✓	✓	✓	✓
Student characteristics			✓	✓	✓
Leadership changes				✓	✓
Peer characteristics					✓
School effects	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓
Observations	130,595	130,595	130,595	130,595	130,595
R-Squared	0.195	0.436	0.446	0.446	0.446

Note: Table A.8 presents regression estimates for student performance. Outcome variable is passing GCSE - EM. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school-year level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status. Key Stage 2 means test score at age 11. *** p<0.01, ** p<0.05, * p<0.1 Source: National Pupil database (NPD) and mobile phone policy from survey.

Table A.9: The Effect of Mobile Bans on Student Performance at age 14

Age 14 Test Scores	(1)	(2)	(3)	(4)	(5)
Mobile Ban	0.99	1.59	2.53*	2.98*	2.34
	(1.77)	(1.49)	(1.50)	(1.54)	(1.53)
Prior Test Scores: Age 11		✓	✓	✓	✓
Student characteristics			✓	✓	✓
Leadership changes				✓	✓
Peer characteristics					✓
School effects	✓	✓	✓	✓	✓
Year effects	✓	✓	✓	✓	✓
Observations	112,339	112,339	112,339	112,339	112,339
R-Squared	0.195	0.745	0.756	0.757	0.757

Note:Table A.9 presents regression estimates for student performance at age 14. Outcome variable is standardized test score at age 14. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school-year level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status. Key Stage 2 means standardized test score at age 11. *** p<0.01, ** p<0.05, * p<0.1 Source: National Pupil database (NPD) and mobile phone policy from survey.

Table A.10: The Effect of Mobile Bans on School Performance at Age 16

School Performance % Students 5 A-C inc Eng+Maths	(1)	(2)	(3)	(4)
Mobile Ban	1.88*	2.08**	2.04**	2.07**
	(1.02)	(1.00)	(0.97)	(0.98)
Prior Test Scores: Age 11		✓	✓	✓
Mean Student characteristics			✓	✓
Leadership changes				✓
School effects	✓	✓	✓	✓
Year effects	✓	✓	✓	✓
Schools	90	90	90	90
Observations	816	816	816	816
R-Squared	0.350	0.876	0.881	0.885

Note: Table A.10 presents regression estimates for proportion of student who pass the GCSE-EM test. All estimates and standard errors are multiplied by 100 to ease interpretation. We use robust clustered standard error at the school-year level with school and year fixed effect. Students characteristics are control for indicator for male, minority, Special Educational Needs, Free School Meal status. Key Stage 2 is standardized student test score at age 11 (before high school). The leadership changes variable control if there was a leadership or policy changes occurring at the time of the introduction of the policy. *** $p<0.01$, ** $p<0.05$, * $p<0.1$ Source: National Pupil database (NPD) and mobile phone policy from survey