Higher education policy in England:

Three empirical studies of influences on enrollment behavior

by

Benjamin Matthew Alcott

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Higher Education)
in the University of Michigan
2015

Doctoral committee:

Professor Brian P. McCall, Chair
Professor Stephen L. DesJardins
Assistant Professor Julie R. Posselt
Professor Anna Vignoles, University of Cambridge, UK

Acknowledgements

I could not have asked for more from a doctoral committee. It has been a real privilege to receive such thoughtful feedback on my work. Beyond the thesis itself, each committee member has offered guidance at critical points during this program, and its hard to imagine my progress without such support. I consider each committee member to be a role model: I aim to be in a similar position one day, and I only hope that I can reflect such generosity, kindness and integrity in my work with students.

The support of other faculty has also helped me to develop during my time in Ann Arbor. Of the many aspects of the University of Michigan that impress me, I have been particularly taken with the willingness of faculty members to give time to students, no matter how unclear the link to their own work. I am grateful to Pamela Moss, Susan Dynarski, Deborah Faye Carter, Lisa Lattuca, Donald Freeman, Frank Yates, and Robert Axelrod, all of whom have helped broaden my research. In addition, I would like to thank Melinda Richardson and Linda Rayle. I do not recall working with such excellent administrative staff before – their warmth and support for my non-conventional path to completing this doctorate are greatly appreciated.

Friends made in Ann Arbor helped me adapt to a program and work culture that felt far more alien than I had anticipated. Both Julio Cardona and Gloryvee Fonseca-Bolorín took great care with the new arrival to their research team, helped me to grow, and kept me on track. Sarah Okuyama, Peter Keane, Marcy Epstein, Josh, their friends and family showed me another side to Ann Arbor and offered a counterbalance to stop me falling too far down the academic rabbit hole. And I still cannot believe my good luck at the cohort that I joined. We created a support network that we each have contributed to and relied upon, and I hope the bonds we have formed will prove lasting.

Finally, family has made all the difference. I am grateful for the blessing of my parents, Gill Buszard and Drew Alcott, and my suegros, David Fishman and Helia Lopez Zarzosa. Charles Lenth has been instrumental, especially at the toughest stages. My Coloradan family offered a home from home, making sure I never missed a family holiday. Lastly, I offer the greatest thanks to my strongest source of support, who has borne an equal share of this doctorate, and who is the love of my life, Leo.

Table of contents

List of Tables

List of Figures

Introduction

England's higher education system has expanded greatly in the last half century, and its universities rank among the world's best. But England also remains a country with persistent socio-economic disparities in access to higher education. Rectifying this remains an avowed priority of policymakers in all major political parties. In one sense, this challenge appears straightforward: the enrollment gap maps directly onto an achievement gap, as students from different socio-economic classes who perform similarly in high-school exit exams enroll in undergraduate degrees at similar rates. Consequently, policymakers have focused on reducing disparities in attainment and high-school completion. However, implementing effective policy mechanisms has proven difficult to achieve.

This thesis consists of three stand-alone papers, each of which assesses a policy incentive that might alter the socio-economic gap in university enrollment; these incentives are encouragement, grade labels, and money, respectively. The rationale for examining these influences comes from a diverse range of theoretical backgrounds. In contrast, the analytical procedure used in each paper comes from the same empirical tradition, namely the Neyman-Rubin causal model, which quantifies the impact of a given incentive by determining counterfactual scenarios to estimate the difference in outcomes stemming from either the absence or the presence of a given incentive.

The first chapter's incentive of interest is teacher encouragement. Bourdieu's theorization of habitus argues that students from socio-economically deprived backgrounds perceive formal education as a hostile environment in which they cannot succeed, thus deterring them from progressing to non-compulsory education and so exacerbating access disparities. However, more recent adaptions of this theoretical tradition suggest that teachers' interactions with such students may help to lessen these disparities. I use propensity-score matching to estimate the impact of teachers encouraging students to continue to the non-compulsory years of high school on future enrollment behavior, both in those high school years and subsequently in university degree programs. To date, educational research on encouragement is generally qualitative in nature. To the best of my knowledge, there are not yet any quantitative inferential studies that assess its longer term effects, and this study helps to fill that gap in the literature.

In the second chapter, I estimate the impact of grades in school assessments (independent of performance in the given assessment) on students' progression to higher education. A variety of theoretical traditions suggest that information about assessment performance is likely to influence the subsequent behavior of school students, their parents, and their teenagers, and to lead to a bifurcation of higher and lower performers. In addition, Bourdieu's sociological theories hypothesize that this information is especially influential on students from socio-economically deprived backgrounds, hence exacerbating disparities. To test these claims, I use a regression-discontinuity design to compare students who performed almost identically in national examinations at age 14 yet received different grades. Similar to the topic of encouragement, prior studies on this subject have used qualitative methods to explore students' perceptions of summative grading and its potential pigeonholing effect on subsequent academic outcomes, but none have tested these theories with inferential statistical techniques or linked the effect of grading to outcomes years after the examination.

The third chapter's incentive of interest is the Education Maintenance Allowance, a conditional cash transfer program for students from low-income households who remained in non-compulsory education at ages 16–19. Human capital theory identifies short term credit constraints as a barrier to students from deprived households being able to progress in non-compulsory education. The Education Maintenance Allowance program, which pays students up to US$45 per week for remaining in education, was intended to lessen these constraints. I assess the program's impact on targeted students' enrollment in education immediately after the end of compulsory schooling, i.e., the end of tenth grade. Continuation in education after tenth grade is relevant to higher education access because it is a prerequisite for entering university and a key point of socio-economic stratification in England's education system. Unlike the first two studies, the theoretical rationale underpinning this incentive – human capital theory – has an established history in quantitative inferential studies. Yet, while strong quasi-experimental research exists on the policy's impact on enrollment levels at the pilot stage, there is a lack of studies that are of commensurate quality on the policy's impact during its national rollout.

This thesis advances research on education policy by spanning previously unconnected theoretical and empirical traditions (chapters 1 and 2) and providing a more accurate assessment of the impact of one of England's largest access policies (chapter 3). Consequently, its component studies can improve our understanding of student motivations and some of the determinants of socio-economic disparities that it may be possible for education policies to ameliorate. Hopefully, these can enable researchers, policymakers, and practitioners to make more equitable what is currently a prestigious but stratified higher education system.

Chapter 1: Does Teacher Encouragement Influence Students' Educational Persistence?

Summary

In this study, I investigate whether teacher encouragement to continue to the non-compulsory years of high school influences the likelihood of students enrolling in (1) A-levels, (2) any university degree course, and (3) a degree course at a prestigious university. I do this by using propensity-score matching on data from the Longitudinal Study of Young People in England. Further, I explore whether the impact of encouragement is greatest for those students in the middle third of academic achievement, as well as those with lower levels of parental education. Model estimates suggest that teacher encouragement increases the likelihood of both A-level and university enrollment. These findings have important policy implications, especially since it appears that teacher encouragement has the greatest impact on those most likely to be on the margin for university attendance.

Introduction

England possesses one of the world's elite university systems, with its prestigious institutions among the most widely recognized and highly ranked (Shanghai Ranking Consultancy, 2013; Times Higher Education, 2012; U.S. News and World Reports, 2012). But this should not mask the fact that participation in higher education, particularly at the most prestigious institutions, is highly stratified (Archer et al., 2003; Ball, 2008; Reay et al., 2005). Between the 1960s and 1990s, the link between parental income and university enrollment grew (Blanden and Machin, 2004; Galindo-Rueda et al., 2004; Glennerster, 2001; Machin and Vignoles, 2004). Research indicates that this class gap has not eroded in more recent years: although it may be true that more students of low socio-economic status are staying in formal education beyond the tenth grade (the point at which school attendance in England ceases to be compulsory), this closing of the socio-economic enrollment gap has not carried over to the high school/university nexus (Chowdry et al., 2008; Jackson, 2011).

Socio-economic disparities in enrollment map onto achievement disparities. Among students with similar high-school qualifications, socio-economic disparities in university attendance are small (Vignoles, 2013). Consequently, Britain's incumbent government has focused on schoolteachers as key to redressing access inequalities (Reay, 2013), whereby teachers are judged primarily by their ability to teach students the core content necessary for them to pass national examinations (Stevenson and Wood, 2014).

In this chapter, I analyze the role of teachers in influencing university access in a manner that is distinctive from the current norms of education policy in England. In addition to framing the role of teachers as teaching course material, I consider them as sources of support and advice for students. More specifically, I analyze the impact of teacher encouragement on students' persistence in education beyond the compulsory years of high school. Encouragement has been discussed in past research on education, but most typically in relation to sociology and educational psychology rather than policy development (for example, McHarg et al. 2007; Moogan, 2011; Reay et al., 2001).

Two dimensions of this study offer a better fit to the needs of policy research. First, I analyze a large-scale dataset that enables a quantitative methodological framework. Past studies of encouragement of school students often undertake constructivist analytical approaches on small student samples, but, while important, these approaches tend to encounter concerns about the generalizability of their findings to the national scale. Second, I connect teacher encouragement to student outcomes that occur years later. Past studies of teacher encouragement commonly examine effects within the same academic year, but the longitudinal nature of the data used in this study enables inferential claims about sustained effects.

Theory

The theory of habitus, as applied to educational access in England, provides this study's conceptual underpinning. The utilization of habitus as a conceptual tool dates back to the Classical Mediterranean, with Aristotle using the term *hexis* to contrast the characteristic of a stable disposition with changing disposition (*diathesis*) and accident (*pathos*) (Nash, 1999; Simeoni, 1998; Wacquant, 2005). While research fields vary with regards to whom they credit for recent applications of the concept – such as Claude Lévi-Strauss in anthropology or Jean Piaget in developmental psychology (Lizardo, 2004) – educational researchers in England focus on the work of Pierre Bourdieu.

For Bourdieu, it is crucial to recognize that humans operate within social structures (Reay, 2004); in his own words, he claims, "the cognitive structures which social agents implement in their practical knowledge of the social world are internalized, 'embodied' social structures" (Bourdieu, 1986, p. 468). However, he also eschewed the types of deterministic, structuralist approaches that deny any modicum of human agency (Ovenden, 2000), arguing that social science "must make room for a sociologically grounded phenomenology of the primary experience" (Bourdieu and Wacquant, 1992, p. 127).

Habitus serves to balance these factors in its depiction of human behavior. Bourdieu defines it as:

> a socialised body. A structured body, a body which has incorporated the immanent structures of a world or of a particular sector of that world – a field – and which structures the perception of that world as well as action in that world. (Bourdieu, 1998a, p. 81)

Habitus is thus both the attitudes and tendencies that one develops as a result of interactions with others (Nash, 2005), as well as one's ensuing actions (Reay, 2004). Cultivated over time, an individual's habitus forms a range of dispositions and behaviors indicating what she considers to be appropriate, desirable and possible. While an individual's decision-making is not fixed across the lifespan, it is bounded at a given time according to prior opportunities and constraints.

Further, Bourdieu (1990) argued that individuals' experiences and interactions occur within a hierarchical social structure imbued with the historical experiences of groups. While habitus occurs at the individual level as a result of the individual's cognitive processes, these and social processes mediate one another throughout the lifespan (Archer, Hollingworth, and Halsall, 2007; Hollway and Jefferson, 2000). Hence, analyses of the social environment are paramount.

While acknowledging that the aspects of identity that shape an individual's habitus are numerous, Bourdieu and co-authors' (Bourdieu, 1998b, 2005; Bourdieu and Passeron, 1977, Bourdieu and Wacquant, 1992) work on educational structures emphasized social class. Bourdieu and Wacquant (1992) argue that upper- and middle- class students are more likely to encounter scenarios in which: "when habitus encounters a social world of which it is the product, it is like a 'fish in water': it does not feel the weight of the water and it takes the world about itself for granted" (p. 127). In contrast, students from the working classes face a greater onus to "engage in rational computation in order to reach the goals that best suit their interests" (Bourdieu, 1990, p. 108). These students are also vulnerable to stigmatization should they behave in line with their own habituses because "taste classifies, and it classifies the classifier" (Bourdieu, 1986, p. 56), and, in formal education, working-class tastes are classified as vulgar (Bourdieu and Passeron, 1979).

Exponents of habitus in the English context echo Bourdieu's claims about the importance of social class in formal education. Ball (2012) asserts that education policies "legitimate and initiate practices in the world, and they privilege certain visions and interests. They are power/knowledge configurations *par excellence*" (p. 22). Reay (2001) argues that the founders of England's mass-education system in the nineteenth century sought to control working-class children. Both Reay (2001) and Ball (2012) claim that this approach was challenged by the growth in welfare spending post-World War II, but was then decisively consolidated in the decades of the 1970s and 1980s. In a documentary analysis, Ball (2012) charts this consolidation in Cox and Dyson's (1969) pamphlets, *The Black Papers*. In the wake of the previous year's student protests in France, Germany, and Mexico, *The Black Papers* reiterated the importance of social order, hierarchy, and traditional values in education. Cox and Dyson also voiced concerns about slipping academic standards, a policy trope still common in the present day (Burke, 2012), as well as claiming that accommodating aspects of working-class culture and speech in school curricula might keep working-class children "literary primitive[s]" (Thornbury, 1978, pp. 136-7).

Among studies of current conditions, a common line of argument is that schools develop processes that reflect the socio-economic composition of their student bodies (Thrupp, 1999), forming institutional priorities wherein "certain sorts of choices or considerations take on an obviousness that is difficult to evade" (Ball, Davies, David, and Reay, 2002, p. 58). In other words, school staff instill in students a sense of which post-school options are feasible and desirable. In Ball et al.'s (2002) study, a student at one fee-paying school tells interviewers,

> I thought about would I go to Cambridge or not, because quite a lot of people, you know always think – am I going to Cambridge or not? I don't know why, that just seems to be the question a lot of people ask themselves about higher education. (p. 58)

In contrast, one working-class respondent recalls from a visit to Cambridge, "it was like a proper castle, and I was thinking – where's the moat, where's the armor? Save me from this" (Ball et al., 2002, p. 68). This is corroborated by Archer and Hutchings (2000), who found that young working-class adults from ethnically diverse communities in London held clearly defined identities according to speech and dress that separated them from the tastes of young people who would go to university. Consequently, many working-class students do not aspire to university because they do not see it as a place for them (Archer et al., 2003; Reay, 2006); in contrast, middle- and upper-class students tend to come from families with a history of higher education participation, and so the decision to apply to and attend a university is relatively smooth (Ball, Reay, and David, 2002).

This class-based disparity in whether to attend university also extends to choice of institution. Ball et al. (2002) argue that the choice process does not rest on prior cognitive performance alone because it is also the case that "university is a choice of lifestyle and a matter of 'taste', and, further, that social class is an important aspect of these subtexts of choice. In other words, this is choice as 'class-matching'" (p. 53). Given the more automatic decision to attend, upper- and middle-class students are thus likely to have more time to select a university that better suits them, consolidating the advantages stemming from the guidance mechanisms already available to them (Brooks and Everett, 2008). In contrast, working-class students who decide to attend university face differing constraints. For example, locality may be of greater importance to working-class students and institutional prestige to middle-class students (Ball et al., 2002; Pugsley, 1998). This adds another barrier to working-class students' enrollment decisions: their likelihood of attending university has been undermined by their experiences with formal education, and this is exacerbated when they only consider a small range of universities as feasible for what "people like us do" (Reay et al., 2005, p. 67), foregoing "tastes of luxury" for "tastes of necessity" (Bourdieu, 1986, pp. 177-178).

Scholars who have applied Bourdieu's concepts to the UK education system have tended to theorize, much as Bourdieu did himself, that teachers are a complicit, rather than a resistant, component of an education system that perpetuates socio-economic disparities. This engenders two primary concerns about the compatibility between (a) these concepts and their past empirical applications and (b) the hypothesis that teacher encouragement may be of particular benefit to socio-economically disadvantaged students. The first concern is whether Bourdieu's theories are so deterministic as to deny the potential for such students to overcome the systemic barriers that they face. This concern is attributable both to theory and to empirical methods. By arguing that the entire education system is designed to reproduce inequality, exponents of habitus may neglect the potential for agency among working-class students that defies these macro-structures. As a consequence, this literature risks susceptibility to broader accusations of determinism made against Bourdieu's theorization of habitus (see, for example, Archer, 2000; Jenkins, 1982, 1992; Mouzelis 1995). This is exacerbated by the predominant empirical approaches in this literature: researchers tend to explore students' dispositions to higher education at the time of interview; because most studies are cross-sectional, this gives more limited scope for how much students' dispositions may vary over time or in response to different stimuli.

At the conceptual level, these accusations of determinism are somewhat ironic given that Bourdieu's intention was to develop a critical theorization of behavior that would amend the determinism of past sociological theory by balancing agency and structuralism (Bourdieu, 1990, 1999; Reay, 2004). Critiques claiming that he fails to strike this balance have received spirited rejoinders in the sociological theory literature, namely on the basis that they fail to recognize Bourdieu's assertion that structure is influenced in turn by counteracting instances of agency (see, for example, Crossley, 2001; Holdsworth, 2006) or that they conflate objectivism with structuralism (see, for example, Lau, 2004; Sweetman, 2003). Nonetheless, the structuralist nature of habitus compels researchers choosing to employ this theory to be aware of similar accusations of determinism in their own work.

Researchers of education in the UK have shown a keen awareness of the challenge of determinism. In their discussions of theory, responses to Bourdieu's own rejoinders to his critics have varied. While the most common approach has been to revisit and accept these rejoinders (see, for example, Allen, 2013; Cochrane, 2011; Pimlott-Wilson, 2011; Reay, 1995, 2004), others have claimed that the coherence of Bourdieu's writing fluctuated, while demonstrating a keen desire to steer clear of determinism in their own work (see, for example, Ball, 2003; Hernandez−Martinez and Williams, 2013). Although discordant, these researchers demonstrate a willingness to engage with the challenge of determinism and to avoid it in their own research.

This care is reflected in analyses of empirical studies, which tend to treat habitus's theorizations of structural inequality as powerful but not irresistible. In their studies of students navigating educational pathways, both Archer et al. (2012) and Pugsley (1998) draw attention to students who were exceptions to the principles of habitus, i.e., working-class students who were successful in school, in order to offer a counterbalance to the majority who corroborate these principles. Other research has been devoted to the experiences of such exceptional students as they progress towards the further and higher education opportunities from which they should be theoretically excluded (see, for example, Byrom and Lightfoot, 2012; Reay et al., 2010; Redmond, 2006). For example, Baker and Brown (2008) attempt to identify factors, such as community cohesion and a shared sense of identity, that have enabled successful working-class students to overcome the barriers that they face. Such evidence suggests that exponents of habitus have been careful to promote sources of inequality as constraining, but not defining, opportunities for working-class students (Vincent and Martin, 2002).

The second concern is whether teacher encouragement to such students might help them to overcome the stigmatization that they encounter in formal education. Bourdieu and Passeron (1990) support Durkheim's (1973) assertion that teachers hold a position of recognized authority; hence, teachers serve a "gatekeeper" role in education because their ability to distribute positive and negative feedback from a position of authority make them instrumental in students' educational progress (Redmond, 2006). In Bourdieu's writing though, efforts by teachers to rectify disparities receive little attention; instead, "as former model pupils who would like to have no pupils except future teachers, teachers are predisposed by their whole training and all their educational experience to play the game of the institution" (Bourdieu and Passeron, 1990, 132). Given the aforementioned tendency of schooling to celebrate cultural norms of the middle- and upper-classes, teachers are complicit in the stratifying role of education, as their judgments of students inevitably reflect and so consolidate social prior classifications (Bourdieu, 1990). Thus, in Bourdieu's original conception of habitus, the role of teacher encouragement in challenging inequalities is minimal.

Much of the research to date on conditions in the UK corroborates the notion that teachers do more to perpetuate than to challenge patterns of inequality. In the latter years of school, teachers may give differential guidance to students according to social class (Preston, 2003; Reay, 1998). Qualitative fieldwork indicates that teachers tend to under-appreciate the achievements of working-class students (Reay, 2005; Redmond, 2006), typically placing them in lower ability streams and less academically challenging subjects. Researchers in this field have attributed this tendency among teachers to conflate middle-class behaviors with cleverness and working-class behaviors with stupidity (Ball, 2003; Reay, 2005). Such designations are likely to have a lasting impact on students' academic confidence and important repercussions for their university applications (Boaler, 1997; Steedman, 1988; Thomas et al., 2012).

Whether intentional or not, such discrimination by teachers is further exacerbated by households, as distinctive class strategies are apparent in families' reactions to such judgments (Ball, 2003): since the formal school system reflects the norms of middle-class behavior, working-class families are more dependent on teachers' opinions of their children's achievement and behavior (Gunn, 2005; Lareau, 1997). In part, this is because parents of higher social classes are often more comfortable engaging with schoolteachers, applying pressure to ensure favorable outcomes for their children (Cochrane, 2007, 2011; Giddens, 1991; Reay, 1995). For example, at schools that stream by ability, such parents may be more willing to contest a teacher's decision to put their child in a lower stream. Pugsley (1998) contrasts middle-class parents who are willing to demand that teachers provide advice on A-level subject choices with working-class parents who are reluctant to initiate any contact, as exemplified by one interviewee who notes, "you don't like to interfere really. You can't, can you?" (p. 79).

However, there is also a competing body of research on conditions in the UK that works from Bourdieu's theorizations but instead depicts teachers as "agents of transformation rather than reproduction" (Mills, 2008, 80). Such work has moved beyond Bourdieu's arguments against accusations that habitus is deterministic, not only accommodating exceptional acts of resistance to the prevailing forces of habitus, but arguing that individuals frequently choose to consciously resist such forces (Crossley, 2001), making these resistant behaviors intrinsic to the formation of a given student's habitus (Sayer, 2005). Oliver and Kettley (2010) argue that teachers' promotion of university applications is key to whether students from underrepresented backgrounds apply, and, across the range of government schools that they surveyed, they find contrasting instances of proactive encouragement and reticence that were not defined by students' socio-economic status.

Student-teacher relationships are key to forming a student's attitude towards formal education (Hollingworth and Archer, 2009), and Reay et al. (2009) emphasize the importance to student plans of relations with individual teachers rather than institutional culture. While also working from Bourdieu's conceptual foundations, this literature provides competing qualitative evidence of the potential for teachers' social interactions to widen access for underrepresented groups. Thus, unlike Bourdieu's conceptualization of habitus in its original form, subsequent theoretical reinterpretations and empirical applications of habitus make it an approach amenable to the notion that teacher encouragement might play a role in reducing socio-economic disparities in students' educational progress.

Aims of the current study. Within the literature to date, most studies of teacher encouragement (for example, McHarg et al. 2007; Moogan, 2011; Reay et al., 2001) are based on convenience samples at either a single or small set of education institutions. Consequently, the extent to which their circumstances are representative of students more broadly is ambiguous. Without more extended discussions of why a given site or sample group is of particular research interest, many of these studies forego the type of theoretical generalization that Eisenhart (2009) argues is so important to qualitative educational research methods. As a consequence, while habitus-based empirical studies frequently provide illuminating perspectives, the absence, to date, of inferential quantitative studies to corroborate their assertions weakens this literature's capacity to make claims about students' experiences at the national scale, which is an important consideration for policymakers.

In this study, I provide a constructive application of the principles of habitus by testing whether teacher encouragement has a positive impact on students' enrollment decisions and whether any effects vary by students' socio-economic status and achievement levels. My use of quantitative methodology extends the current literature because it facilitates the construction of a credible counterfactual, and the use of longitudinal data makes it possible to assess the impact of social interactions on observed enrollment behaviors years later. The study's aims are articulated by the following research questions:

1. If a teacher encourages a student to progress to the non-compulsory years of high school, does this increase the likelihood of that student enrolling in A-levels?
2. Does any impact extend to future enrollment in a) any university and b) a high-prestige university?
3. For all outcomes, does the impact of teacher encouragement vary by parental education or students' prior academic achievement?

Although I use Bourdieu's concept of habitus to theorize class disparities in access and choice, my empirical strategy is markedly different from the norms of the habitus literature in England. More specifically, I use a quantitative estimation approach rather than the more common qualitative approaches of interviews, focus groups, and ethnographic observation, and, as a consequence, I relinquish affordances of those approaches. For example, with qualitative methods, researchers have been able to provide phenomenological studies that represent students' decision-making in a nuanced, heterogeneous manner.

In contrast, I use a straightforward measure of students' perceived encouragement from teachers, and link the impact of this encouragement to observed enrollment behavior in subsequent years. The measure is a closed survey question: *did they [the teacher] tell you [the student] that they thought you should stay on in full-time education?* Thus, it can only examine a single dimension of teacher interactions with students, and in a narrow manner.

In the research literature though, no standard approach has yet been accepted for operationalizing either teacher encouragement or support. Past studies have varied in how they operationalize encouragement, although most if not all focus on students' perceptions. Reay et al. (2001) define teacher encouragement as persuading students that there are benefits to staying in higher education. Moogan (2011) discusses teacher encouragement in terms of teachers sharing information about their own experiences in higher education, as well as information about the type of school attainment necessary in order to make feasible a particular choice of subject major at university. McHarg et al. (2007) treat support as a teacher telling students they should continue to post-compulsory education and provides information about whether they are likely to get the requisite grades. Among quantitative studies, Berzin (2010) uses a scale based on 11 questions asked of students about teacher support, but does not specify what these questions are. In their study of student aspirations, Anders and Micklewright (2013) use the same question from the Longitudinal Study of Young People in England (LSYPE) survey to measure teacher encouragement as I do in this study.

Students answered this survey question during the same school year to which the question relates (i.e., tenth grade). Still, it is worth noting that this measure of encouragement is vulnerable to potential sources of bias. One challenge is that students' interpretation of the question might vary systematically according to prior characteristics. For example, students who feel more positively than others about school prior to tenth grade may be more likely to interpret the same interaction with a teacher as encouragement. A similar but distinct challenge is response bias: students may be consistent in how they interpret an interaction but differ in their reporting of it. Again, this might differ according to background characteristics such as self-worth, aspirations, and prior attainment. Both challenges increase the risk of spurious estimations of teacher encouragement when systematic differences occur according to past characteristics likely to correlate with enrollment in post-compulsory education.

The Empirical Model

I work from the assumption that it is implausible that teachers are randomly choosing which students to encourage to continue to the final years of schooling. Such non-random provision of support is likely to create a selection bias in the 'treatment' of interest. Further, this selection bias is likely to be apparent in treatment status across a range of covariates that the research literature indicates are predictors of educational attainment. For example, on average teachers are more likely to provide encouragement to students with higher prior attainment (Franklin, 1995; McHarg et al., 2007). Nonetheless, my hypothesis is that, even after accounting for selection bias, students receiving encouragement from teachers are more likely to progress to university than those who do not.

I attempt to account for selection bias in the receipt of encouragement by establishing the counterfactual. The counterfactual compares two outcomes for any individual: first, receiving the ‘treatment’ (in this paper, teacher encouragement), and second, not receiving the treatment (Rubin, 1974; Holland, 1986). Either outcome can potentially occur for any individual in the analysis (Holland, 1986), but only one actually does occur, providing the observed outcome, while the other, the unobserved outcome, does not.

Assessing the difference between the observed and unobserved outcomes can ideally be achieved using random assignment of ‘treatment’ and ‘non-treatment’ across a sufficiently large group, since all other characteristics that might influence the outcome of interest will be indistinguishable, on average, between the groups. However, random assignment is rarely feasible for many ‘treatments’ of interest to social science researchers. This has led to the development of quasi-experimental methods that seek to statistically emulate features of random assignment. Among these, the most common are various forms of regression. These parametric methods control, or hold constant, other characteristics that influence the outcome of interest and would be equally distributed across treated and untreated groups under random assignment. Yet, although popular, the assumptions underlying regression analyses are stringent and will not hold in many cases (Reynolds and DesJardins, 2009). For instance, it is important to control for all potential confounding factors. Further, these controls must either have a linear relationship with the outcome of interest or the researcher must successfully predict the characteristics of the non-linear relationship.

Matching provides a non-parametric alternative to regression analyses. Matching individuals in treatment and control groups according to identical values for confounding characteristics would allow for treatment and control groups to be identical (at least across observed characteristics) except for receipt of the treatment, thus providing a strong approximation of a randomized trial. However, matching across numerous characteristics increases the likelihood of encountering the “curse of dimensionality”: as the number of characteristics increases for which a match is needed, the probability of finding individuals in both groups with matching characteristics goes to zero (Reynolds and DesJardins, 2009).

Propensity-score matching offers a solution to this challenge. Instead of comparing individuals across multiple characteristics, the approach matches on a single dimension: the probability of receiving the treatment of interest (Rubin, 1997). The propensity score is created for each individual by estimating a logistic regression on observed characteristics to calculate the probability of receiving the treatment. As Rosenbaum and Rubin (1983) have demonstrated, matching on the probability of covariate occurrence is equivalent to directly matching on covariate existence. The impact of the treatment is then estimated by calculating the average difference in the outcome of interest between pairs with matching propensity scores across the treatment and control groups.

My goal is to identify the average treatment on the treated (ATT). In other words, I aim to assess the average impact of encouragement on the subset of students who received encouragement. This is a different goal from the average treatment effect (ATE), for which I would calculate the average impact of the treatment on both treated and untreated individuals. The reason for targeting the ATT over the ATE is that teacher encouragement is not randomly assigned, and so an estimation of the impact of encouragement on all students may be less plausible and so have fewer practical implications. In addition, it may not be feasible or desirable for teachers to encourage all students to continue to the non-compulsory years of high school.

The ATT is formally identified as:

$$\text{ATT} = E(Y_1 \mid T = 1) - E(Y_0 \mid T = 1), \quad (1)$$

where $E(Y_1 \mid T = 1)$ represents the conditional expectation of Y_1 given T equals 1. This is readily observable in the data by finding the enrollment rate among students who received encouragement. The second part of the right-hand side, $E(Y_0 \mid T = 1)$, represents the conditional expectation of Y_0 given T equals 1. This is the counterfactual, i.e., the unobserved outcome.

I attempt to provide a reliable estimate of this counterfactual through the use of propensity score matching. Using a set of predictors, I estimate a logistic regression model to determine each student's propensity score, i.e., the probability of receiving the treatment:

$$\log \frac{P_s}{1-P_s} = \alpha + \beta_1 X_{1s} + \beta_2 X_{2s} + \ldots + \beta_k X_{ks} + \epsilon_s \quad (2)$$

where P_s is the estimated propensity score for student s, α and β_1 through β_k are estimated coefficients, and ε_s represents a random error term that is logistically distributed. X_1 to X_k are all of the observed background characteristics that I discuss subsequently. This model is estimated with the LSYPE's survey weights, and standard errors are clustered at the level of schools. Each student's propensity score is estimated once and then used in all subsequent matching models.

After estimating each student's propensity score, I derive the ATT through kernel-based matching as developed by Heckman, Ichimura and Todd (1998). Evidence suggests that the kernel approach to matching is more precise than the most common alternatives, radius and one-to-one matching (Frolich, 2004). Kernel matching is a non-parametric estimation approach that uses multiple untreated participants to generate the counterfactual for any treated individual. This offers a trade-off in comparison to one-to-one matching, as matching on additional observations with more dissimilar propensities reduces variance, but at the risk of increased bias (i.e., non-comparability between a treatment unit and its matched observations) (Caliendo and Kopeinig, 2008). In comparison to radius matching, another approach that uses multiple observations, kernel-based matching reduces the risk of bias because it uses a weighting function that weights more heavily those observations with propensity scores closest to that of the given treated individual.

As Reynolds and Desjardins (2009) note, the ATT for matching methods is represented by

$$ATT = \frac{1}{n_1} \sum_{i \in (T=1)} \left(Y_{i1} - \sum_{j \in (T=0)} w(i,j) Y_{0j} \right) \tag{3}$$

where n_1 is the number of treated units, j is the untreated unit, and $w(i, j)$ is the weight placed on each untreated unit j for a treated unit i. I define this weight according to a kernel function,

$$K(\varphi), \quad \varphi = \frac{\left(\frac{P_i(X)}{1 - P_i(X)} \right) - \left(\frac{P_j(X)}{1 - P_j(X)} \right)}{h} \tag{4}$$

in which φ represents the quality of the match. This quality of match is calculated by taking the difference between the odds ratio of the propensity score for the treated observation (S_i), and the odds ratio of the propensity score for the untreated observation (S_j), as a proportion of a bandwidth h. I match on the odds ratio of the propensity scores rather than the propensity scores themselves because this provides more robust estimates when some respondents are oversampled in relation to their proportion in the overall population (Smith and Todd, 2005a), as is the case with LSYPE. This is particularly important when using kernel matching approaches, because in these approaches multiple observations are used to provide each counterfactual and are weighted on the basis of their absolute difference from the treated observation (Caliendo and Kopeinig, 2008).

Using this kernel function then, the weight for a given untreated observation is

$$w(i, j) = \frac{K(\varphi)}{\sum_j K(\varphi)} \tag{5}$$

That is, each kernel weight for the given observation *j* must be divided by the sum of the kernel weights in order for the matching weights *w(i, j)* to sum to one. Consequently, the kernel-based matching approach requires choices with regards to the type of kernel function (*K*) and bandwidth size (*h*), with the former typically proving less important than the latter (Caliendo and Kopeinig, 2008). I use the Epanechnikov kernel function. In comparison to the other most common kernel functions – the triangle and Gaussian kernels – the Epanechnikov kernel function strikes a balance between prioritizing non-treated individuals with similar propensities while still incorporating the results for a range of the non-treated (Reynolds and DesJardins, 2009). However, models are also estimated with the alternate kernel types in order to test their sensitivity to this choice. Results, which are compiled in Appendix 1.C, indicate that ATT results estimated with the Epanechnikov kernel are typically more conservative than those with the other kernel types.

Choice of bandwidth size is essentially a tradeoff between limiting variance, which increases as bandwidth increases, and limiting bias, which increases as bandwidth decreases. I estimate models with bandwidths between 0.1–0.11, meaning that for a given treated individual the comparison match is derived from untreated individuals whose propensity score fell within 0.05 and 0.055 on either side of the treated individual's score. I tried a range of bandwidths and this choice of bandwidth had the best covariate balance between the treatment and matched comparison group, which is an important prerequisite for the matching procedure to maintain validity (Guo and Fraser, 2010). Analyses were conducted using the PSMATCH2 program (Leuven and Sianesi, 2003) for the Stata statistical package.

Data

I use data from the British Department of Education's Longitudinal Study of Young People in England (Department for Education and National Centre for Social Research, 2012). The Longitudinal Study of Young People in England (LSYPE) used a two stage probability proportional to size sampling procedure. The primary sampling units were schools, of which 647 were surveyed during the first wave. Respondents were born between September 1, 1989 and August 31, 1990, and were interviewed annually between 2003, at which point they were in eighth grade, and 2010. The average number of respondents per school was 24 in the first wave, although with attrition this average had fallen to 13.2 by the seventh wave. Respondents' parents were interviewed for the first four years. Religious, ethnic and linguistic minorities were 'oversampled' in proportion to the English population. I use the survey's population weights in order to account for this 'oversampling,' differential response rates across waves, and the use of clustered sampling within schools.

However, as with most longitudinal surveys, the LSYPE is prone not only to sample attrition but also to missing components – in particular, non-participation from parents – and item non-response (Piesse and Kalton, 2009). The challenge of missing parent-interview data is mitigated somewhat by only using items from the interviews with the primary parent or caregiver, rather than also using corroborative items from interviews with the secondary parent or caregiver, thus avoiding the chance of losing an additional nine percent of respondents.

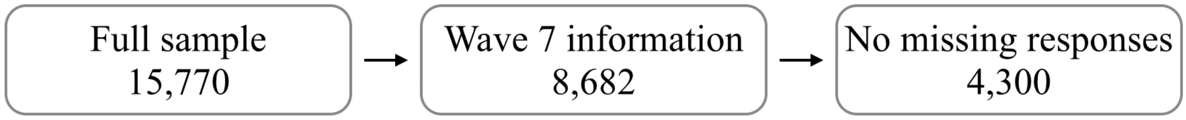
For the outcomes of interest, item non-response was lower than 1% for both the questions relating to A-level study (0.91 percent) and university attendance (0.21 percent), indicating that missing data for this variable almost entirely reflects attrition (Anders, 2012). Appendix 1.A provides information about non-response rates on other variables that have missing values, first for when A-level study is observed, and then for when university study is observed. For all variables, missing values were not fully nested within missing values for any other variable, with the exception of the geographic variables (region, neighborhood wealth, and urbanicity), all of which are generated from the same base variable (postal code). As Table 1.A1 indicates, non-response is especially high for the family-income and National Statistics Socio-economics Classification (NSSEC) variables, two characteristics that are highly correlated both with academic progression and to one another (Piesse and Kalton, 2009). Analyses indicated that missingness for either variable was strongly associated with progression to higher education and was therefore not occurring at random.

Since item non-response was particularly high for only two covariates, it might be tempting to create dummy variables to identify non-response for these respective covariates in order to hold the cases with missing data. However, while this approach would help to maintain sample size, any subsequent estimated models are likely to produce biased coefficients (Jones, 1996), and even original proponents of dummy non-response now reject this approach (Cohen and Cohen, 2003). Instead, I use listwise deletion, which yields approximately unbiased coefficient estimates even when data is not missing at random (Little, 1992). The main weakness of listwise deletion is the loss of sample. Yet, while the resulting loss of observations leads to larger standard errors, the estimated standard errors produced by models after using listwise deletion tend to provide accurate estimates of true standard errors, making listwise deletion an “honest” approach to managing item non-response (Allison, 2002).

Table 1.1. Descriptive statistics for sample

	% or Mean(SD)		% or Mean(SD)
Enroll in university	0.45	Predict child continues past 10th grade	0.83
Enroll in prestigious university	0.09	Parent Saving for university	0.27
Teacher encouragement	0.48	Home Internet	0.91
Region		*Religion*	
Northwest	0.14	Christian	0.58
York Humber	0.11	Muslim	0.03
East Midlands	0.09	Hindu	0.01
West Midlands	0.12	Sikh	0.01
East	0.11	Other	0.01
London	0.10	None	0.37
Southeast	0.16	Religiosity (1=high, 4=low)	3.1(1.0)
		Academic aptitude	
Southwest	0.11	5th grade English	27.4 (4.0)
Rural area	0.22	5th grade math	27.4 (4.6)
IDACI score	0.18 (0.16)	5th grade science	28.9 (3.4)
Household income	34,234 (29,738)	8th grade English	34.7 (5.7)
Parent education		8th grade math	37.4 (7.3)
No educational qualifications	0.07	8th grade science	34.8 (6.2)
Sub-10th grade qualification	0.05	Female	0.51
		Racial/ethnic identity	
10th grade qualification	0.25	Afro-Caribbean	0.04
12th grade qualification	0.20	South Asian	0.06
Post-high school, not degree	0.18	Other minority	0.01
University degree	0.19	English foreign language	0.03
Parent occupational status (NSSEC)			
High	0.47	British-born	0.96
Medium	0.28	Aware of EMA policy	0.64
Low	0.14	Free school meals	0.09
Parent long-term unemployed	0.11	Special educational needs	0.18
N	4,300		

Figure 1.1. *Defining the sample size*



Consequently, this analysis uses a partial sample of respondents from the LSYPE dataset. Figure 1.1 presents the stages in which this sample was restricted. The first wave of interviews sampled 15,770 respondents in 2003, when they were in eighth grade; by the seventh and final wave in 2010, the number of participants had fallen to 8,682. However, I only consider those respondents still participating by the seventh wave and who, with the exceptions mentioned, responded to all variables used in the analyses. This left a restricted sample of 4,300 respondents. Table 1.1 presents a full range of descriptive statistics for the restricted sample of 4,300 observations, after the requisite adjustment for population weights.

Given that I restrict my sample on the basis of missing responses, it is not possible to check the comparability of the restricted and full samples across a broad range of respondent characteristics. However, because the LSYPE is linked to the government's National Pupil Database, it is possible to compare attainment in national examinations between my restricted sample and the full LSYPE sample, as presented in Table 1.2. The restricted sample performed significantly better than the full sample in examinations at all three time points, indicating that the restricted sample is not directly representative of the full sample. Consequently, subsequent model estimates of the impact of encouragement should not be taken as accurate point estimates for the full LSYPE cohort, but rather as indicative of trends.

Table 1.2. Comparison of restricted and full samples

	Restricted sample		Full sample		
	Mean (SD)	*n*	Mean (SD)	*n*	Difference in means
National examinations at:					
Fifth grade (Key Stage 2 average)	27.9(3.6)	4,300	27.1(4.1)	14,559	0.8*
Eighth grade (Key Stage 3 average)	35.7(6.0)	4,300	33.8(6.8)	14,828	1.9***
Tenth grade (GCSE capped points)	326.5(83.8)	4,300	295.6(111)	15,329	30.9***

Notes: Asterisks indicate significant difference in means between samples: * p<0.1, ** p<0.05, *** p<0.01

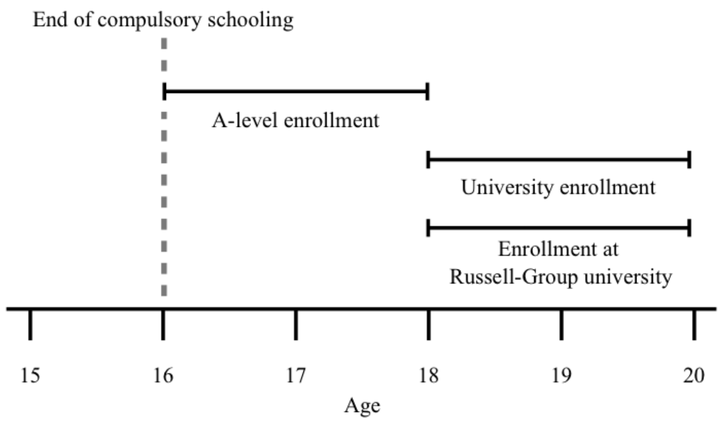
Model variables

Outcomes of interest

This analysis focuses on three outcomes, which are presented in relation to respondent ages in Figure 1.2. Each outcome measure is self-reported and operationalized by its own dichotomous dummy variable. The first is a given student's enrollment in A-levels, generally considered the more academic pathway in the final years of high school, within two years of completing compulsory education (i.e., tenth grade). The second is enrollment in a university-degree course at ages 18–20 (it is not possible to measure enrollment at a later age because respondents were 19 to 20 years old when the most recent survey wave was conducted). The third is enrollment in a university-degree course at a high-prestige institution, again at ages 18–20. I operationalize high prestige as Russell Group membership. The Russell Group consists of 24 of the United Kingdom's 119 universities. It is a self-selecting membership group, which arguably heightens its elitist identity. Although Russell Group membership is not an objective signifier of university quality, 22 of the 24 Russell group universities were ranked among the top 30 of the United Kingdom's 162 universities in the most recent Times'

(2015) league tables, and its member institutions tend to set higher academic undergraduate entry standards than other British universities.

Figure 1.2. *Outcomes of interest*



Key explanatory variable

The key explanatory variable of interest is whether a given student reported receiving encouragement from schoolteachers to continue to A-levels, the more academic pathway offered for the non-compulsory final two years of schooling in England. This measure is operationalized as a student's self-report during tenth grade (the final compulsory year of schooling) of whether they recall any teacher encouraging them to continue to A-levels. The variable is dichotomous, with students either reporting encouragement from one or more teachers, or not at all.

Model controls

Because my aim is not simply to establish an association but a potential cause and effect, it is important to account for other factors that are linked both to schoolteacher encouragement and progression to postsecondary education. In propensity-score matching, this is done through the explanatory variables used to generate each individual's propensity score, as per equation 2. Informed by the research literature on university participation in

England, the variables that I select can be grouped into four main categories: (1) finances, (2) academic achievement, (3) social class, and (4) elements of ethnic, religious, and gender identity. These groupings are not meant to define the role of each variable in teacher encouragement and university access, but simply to support their inclusion in the estimation of the propensity score. All of the variables that I list here were used in the estimation of students' propensity scores. All variables are measured during LSYPE's first two survey waves, which ensures that the explanatory variables are not dependent on teacher encouragement in tenth grade.

Finances. Multiple researchers have attributed participation gaps to wealth inequality. Typically, these researchers do not explicitly ground this attribution within a theoretical framework. Still, they tend to be consistent in how they characterize the mechanism by which wealth influences university access: poorer households perceive university to be unaffordable; even though government loans mean that this is not strictly the case, this perception hampers poorer students' intentions to apply to university. Students from poorer households are especially likely to be intimidated by the costs of attending university (Pennell and West, 2005); while they tend to be knowledge about fees, they are far less likely to know about the range of scholarships and bursaries available to offset these (Chowdry et al. 2012; Christie and Munro, 2003). This challenge is further demonstrated by Davies et al.'s (2008) survey data, in which two thirds of respondents not attending university said debt affected their decision "much" or "very much", twice the rate of those who did attend. Moreover, the majority of these respondents reported that they chose whether to attend prior to any knowledge of bursary options.

Financial controls consist of five items: household income, university saving, region, urbanicity, and neighborhood wealth. Family household income is measured as a continuous variable, and its impact is estimated with a logarithmic function. Parental saving is a dichotomous variable, noting whether any parent or caregiver in the given student's household reported having started saving for the student's costs of higher education by the time the student was in ninth grade. Region and urbanicity are included because the cost of living varies in England according to both factors. For region, dummy variables are used to represent the nine Government Office Regions. Urbanicity is a dichotomous variable indicating whether the student lived in a rural or an urban area. Neighborhood wealth is measured according to the percentage of children below age 16 in the local area who live in low income households, as provided by the government's Income Deprivation Affecting Children Index (IDACI). In addition, I account for internet access, on the basis that insufficient knowledge of financial support among poorer students is exacerbated by their relatively poor access to home internet (Gorard & Selwyn, 1999; Lee, 2008). Internet access is specified by a dichotomous variable indicating whether the student had access to the internet in their home at all while they were in ninth grade.

Academic achievement. An alternative body of research emphasizes disparities in educational attainment in accounting for the class gap in higher education participation. The theoretical justification for this emphasis can be traced back to Carneiro and Heckman's (2002, 2003) extension of human capital theory, in which they argue that childhood disparities in cognitive development, rather than financial resources, between poorer and wealthier students are the principal causes of divergences in future education and employment. When controlling for achievement, at age eighteen the gap in continuation to further and higher education by socio-economic class may be as low as one percent (Chowdry et al., 2008). Since the introduction of university fees in 1998, social class effects become insignificant with finer controlling of educational achievement (Galindo-Rueda et al., 2004). As a result, policies aimed at improving attainment and attendance among school students may have the greatest impact (Anders, 2012a).

Academic achievement controls consist of students' test scores in six national examinations. Students sat tests in English, science, and mathematics during the fifth grade and again during the eighth grade; each score is measured as a continuous variable. For models that try to distinguish whether the impact of encouragement may vary according to prior achievement, students are divided into tertiles according to their average score in these examinations. The number of attainment levels chosen is three because the high–mid–lower grouping is a recurrent theme in British education policy, both for how schools set by ability and how government policy briefings have tended to conceptualize student ability (Gillard, 2008; Ireson and Hallam, 1999).

Social class. Researchers who focus on the role of social class in causing educational disparities tend to derive their conceptual frameworks from Bourdieu's (1998b) theorization of habitus. As elaborated upon earlier, this theorization claims that working class students' experiences of stigmatization in formal education deter them from progressing to higher education. Jackson and Jonsson (2013) argue that too much emphasis is now put on academic performance, claiming that they "are able to point to differences in the choices made by members of different social groups, conditional on performance, as the crucial explanandum" (pp. 332-333). The class gap in attainment varies by educational stage: while it may be lessening in recent years at the transition between tenth-grade GCSE examinations and twelfth-grade A-level examinations, it has persisted between A-levels and higher education (Jackson, 2011). For university choice, student conceptions of the type of universities that they feel would provide a good 'fit', as well as those that they think worthy of even considering, is guided in large part by social class (Ball et al., 2002).

Social class variables consist of the main parent or caregiver's occupational class, education, educational expectations for the student, and home internet access. Occupational class is grouped as low, medium, high, and long-term unemployed; these are derived according to the National Statistics Socio-Economic Classification (NSSEC) system. Parental education is operationalized as a categorical variable, taking the highest qualification among parents and/or caregivers in the student's household. Parent/caregiver expectation is measured by a dichotomous variable stating whether the adult expected the student to continue to non-compulsory education (asked when the student was in ninth grade).

Racial and ethnic identity controls. Many researchers have expanded Bourdieu's (1998b) theorization of habitus to examine the role of racial, ethnic, and religious identity in forming students' beliefs about higher education (see, for example, Archer et al., 2001; Archer and Leathwood, 2003; Burke, 2010). It is worth noting though that access disparities in England differ from those in the United States with respect to ethnic minorities. In England, students from non-White backgrounds are proportionally 'over-represented' in universities, in the sense that a greater proportion of university students are non-White than the proportion of English citizens of the traditional university-going ages. However, this varies by minority group, and is particularly pronounced among students from Indian, Pakistani, Bangladeshi, Chinese, and mixed backgrounds, whereas rates are lower for Afro-Caribbean students (Chowdry et al., 2008). Further, minorities are twice as likely as White students to feel poorly informed about forms of financial relief to offset the costs of study (Adnett and Tlupova, 2007), and, on average, minority students do not fare as well as White students in achievement at university or subsequently in the labor market (Conor et al., 2004).

In addition, analyses of higher education access should not focus solely on aggregate attendance rates at the expense of analyses of which universities students are attending.

Minority students are more likely to attend ‘post-1992’ institutions, which have higher dropout rates and are typically regarded as inferior to older universities (Ball, 2008). Despite the fact that tuition costs have been uniform for decades and now vary only slightly, distance from home plays a particularly important role in choice of institution for low-income students, and Pakistani and Bangladeshi girls (Gibbons and Vignoles, 2012).

Ethnic identity is first operationalized as students’ self-designation into racial groups, which is categorized into four dummy variables: White, South Asian, Afro-Caribbean, and all other designations. In addition, a dummy is included for whether the student speaks a language other than English in their household. Religion is operationalized as a categorical variable, grouped as Christian, Muslim, Hindu, Sikh, other, and none. Each student’s main parent/caregiver was also asked to state the importance of religion to his or her life on a four-point scale. In addition, sex is specified by a dichotomous variable for whether the student is female.

Findings

Descriptive statistics

In Table 1.3, rates of reported teacher encouragement, enrollment in A-levels, enrollment in university, and enrollment in a prestigious university are presented, first for all students and then according to parental education and prior academic achievement. These figures suggest that the likelihood of receiving encouragement varies considerably across this range of observable characteristics: students with higher levels of parental education are far more likely than other students to receive teacher encouragement, and this is also true for those with higher test scores. These differences confirm the need for an analytic strategy that can account for confounding effects, as teacher encouragement is not occurring independently of these other characteristics. Omitted variable bias is highly probable, that is, factors existing prior to encouragement are likely to be driving both the occurrence of encouragement and university enrollment.

Table 1.3. Reported encouragement and enrollment by key characteristics

	Reports teacher encouragement (percent)	Enrolls in A-levels (percent)	Enrolls in any university (percent)	Enrolls in prestigious university (percent)
All	48	60	45	9
Academic achievement				
Lower	34	19	10	<1
Middle	44	49	32	1
Higher	57	87	71	19
Parental education				
No qualifications	40	40	26	3
10^{th} grade qualification	46	50	33	3
12^{th} grade qualification	46	56	41	7
University degree	53	77	63	16

Model Fit

Propensity score matching may help to account for confounding effects, but it is important to check the appropriateness of the matching model prior to reporting ATT estimates. In order to provide a robust estimation of the counterfactual, propensity-score matching analyses must adhere to three assumptions: conditional independence, common support, and covariate balance (Reynolds and DesJardins, 2009).

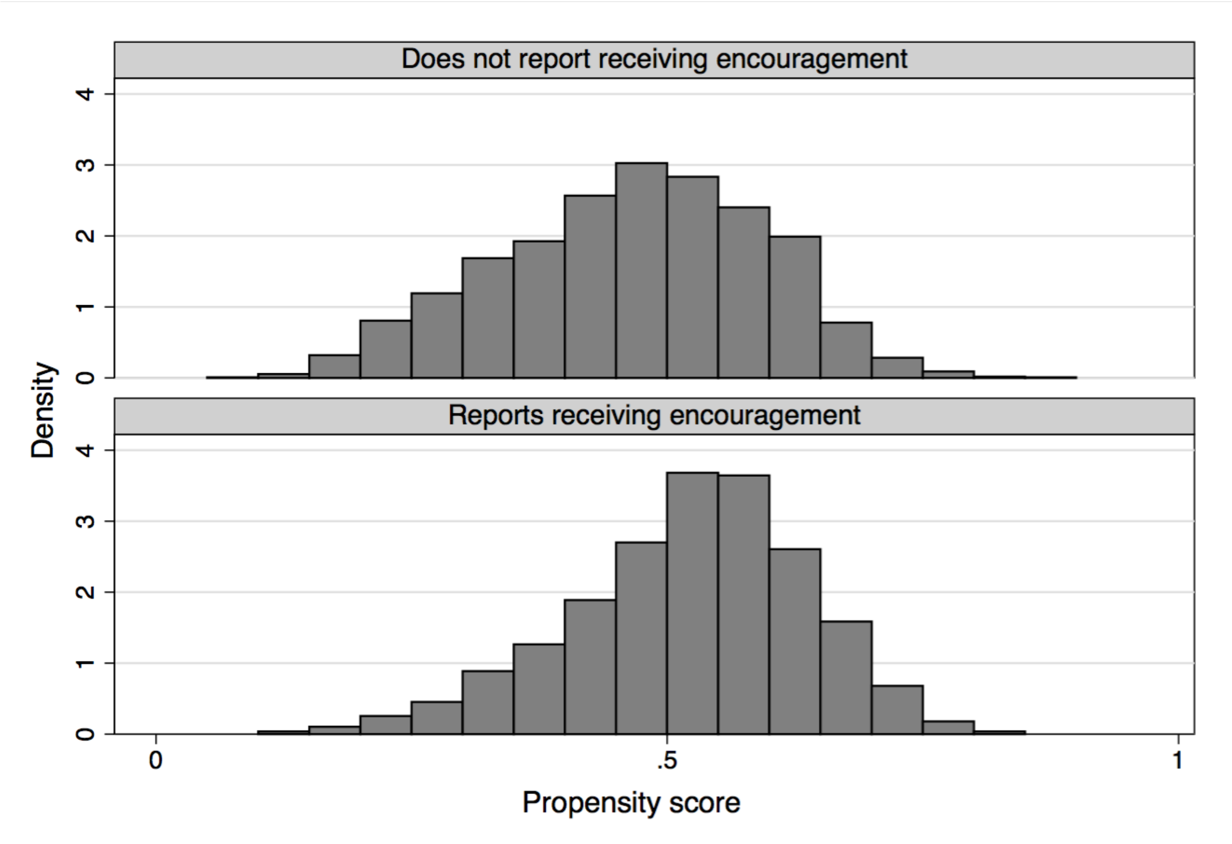
Conditional independence. While previous research comparing propensity-score matching to randomized allocation argues that the former may adequately emulate the latter (e.g. Dehejia, 2005; Dehejia and Wahba 1999, 2002), Smith and Todd (2005a; 2005b) have argued that the approach has considerable pitfalls. Among these is the assumption that, conditional on the observed covariates used in the model, the outcome of interest is independent of treatment assignment (Rosenbaum and Rubin, 1983). Therefore, model estimates are sensitive to accurate selection of the characteristics that lead to the calculation of the propensity scores. I aim to overcome this challenge by using past theory and literature to inform the variables that are used to generate the propensity scores. Fortunately, the LSYPE dataset has detailed information about respondents, and so the key themes in past literature on university access in England are represented. The fact that so many of the variables used in the propensity score estimation are significantly different between treatment and control groups also suggests that there is good reason to include them.

The greater challenge though is ascertaining which variables that are not included in the model would have further improved the propensity score estimations. Appendix 1.B of this paper provides information about Mantel-Haenszel tests (Mantel and Haenszel, 1959) for each model, which assess sensitivity to omitted confounding variables by calculating how large an independent impact that variable would need to have on the odds ratio for the outcome in a given model to become insignificantly different from zero.

Regarding what such potentially confounding omitted variables might actually be, perhaps the most compelling are data on school attended, teachers, and prior student effort. When estimating students' propensity scores (equation 2), I was not able to run a school fixed effects model due to the large number of schools in the dataset with very few sampled students. This is mitigated, albeit probably only moderately, by the inclusion of IDACI scores. Besides providing direct measure of neighborhood deprivation, the IDACI is highly correlated to school attended: 55% of the variation in IDACI is explained by which school each student attends. This is beneficial since the IDACI score can serve somewhat as a proxy for part of the variation in school attended.

Unfortunately, the LSYPE contains no information on which teacher provided a given student with encouragement, nor on any characteristics that might inform the likelihood of them doing so, such as experience, subject specialty, or attitude. While student effort is likely to correlate with a number of the variables used in this study's models, such as attainment or aspirations, some more direct proxy of effort would further improve the study's validity. The research literature in psychology has developed a number of constructs for measuring effort (see, for example, Marsh et al. 2003; Midgley et al., 2000; Pintrich et al., 1993); unfortunately, none of these were collected in the LSYPE surveys. Again though, the model's associated Mantel-Haenszel tests indicate that any such measure would need to have a strong impact, independent of the existing variables, on the prediction of teacher encouragement in order to render spurious this study's model estimates.

Figure 1.3. *Common support between treated and matched untreated groups*



Common support. The common support assumption requires that there are a sufficient number of treated and untreated observations with comparable propensity scores.

When treated individuals have propensity scores far higher than those of the untreated, the propensity score-matching approach is an inappropriate analytical procedure because the comparability of the groups is highly questionable (Blundell et al., 2005).

Figure 1.3 provides a visual representation of the extent to which treated and non-treated students have comparable observed likelihoods of receiving teacher encouragement. This suggests that most if not all treated students had a calculated propensity that was equivalent to that of some students in the matched control group. This was confirmed during the propensity matching estimation, which dropped just 1 of the 2,181 untreated students due to a lack of common support. This represents less than 0.1 percent of the sample, indicating that the paper's matching analysis meets this model assumption.

Covariate balance. The third core assumption is that observed characteristics are balanced between the treatment and comparison groups after matching. In other words, the matching process can only be successful when group means and standard deviations for each of the covariates do not differ significantly between the matched treatment and control groups. However, within the research literature no clear standard has been reached that would provide the criteria for what does and does not constitute significant differences between covariates (Reynolds and DesJardins, 2009). Multiple methods exist for testing covariate balance between treatment and control groups, the most common of which is likely to be standardized differences and associated percentage bias (Caliendo and Kopeinig, 2008; Smith and Todd, 2005a). However, the standard of a 20% bias limit, as established by Rosenbaum and Rubin (1985), seems somewhat arbitrary (Reynolds and DesJardins, 2009). For ease of analysis, I opt for a two-sample t-test of the difference in means across the groups for each of the covariates. Balance is achieved for any given covariate when it is not possible to reject the null hypothesis of no difference between the groups at the five percent significance level.

Table 1.4. Comparison of treated, all untreated and matched untreated groups

Variable	Treated Mean	All untreated Difference	Matched untreated Difference	Variable	Treated Mean	All untreated Difference	Matched untreated Difference
Region				*Religion*			
Northwest	.12	+.02	+.01	Christian	.57	+.01	+.01
York Humber	.10	+.03*	<.01	Muslim	.04	-.01	<.01
East Midlands	.10	-.02	<.01	Hindu	.01	<.01	+.01
West Midlands	.13	-.02	<.01	Sikh	.01	<.01	<.01
East	.11	.01	<.01	Other	.01	-.01**	-.01**
London	.10	-.01	<.01	None	.35	+.01	<.01
Southeast	.17	-.02	-.01	Religiosity (1=high, 4=low)	3.1	<.01	<.01
Southwest	.11	+0.1	<.01	*Test scores*			
Rural area	.23	-.01	<.01	5^{th} grade English	28.0	-1.3**	-.01
IDACI score	.18	.01	<.01	5^{th} grade math	28.1	-1.5**	-.01
Household income	35,480	-2,460	-472	5^{th} grade science	29.5	-1.0**	<.01
Parent education				8^{th} grade English	35.7	-1.9**	-.03
No educational qualifications	.06	+.02**	<.01	8^{th} grade math	38.8	-2.6**	-.02
Sub-10^{th} grade qualification	.04	+.02**	<.01	8^{th} grade science	35.9	-2.3**	-.03
10^{th} grade qualification	.24	+.02	<.01	Female	.52	-.01	+.01
12^{th} grade qualification	.19	+.02	+.01	*Racial/ethnic identity*			
Post-high school, not degree	.20	-.04*	+.01	White	.89	+.01	<.01
University degree	.22	-.07**	-.02	Afro-Caribbean	.04	+.01	+.01
Parent occupational status (NSSEC)				South Asian	.06	-.01	<.01
High	.51	-.09**	-.02	Other minority	.01	<.01	<.01
Medium	.29	+.02	+.01	English foreign language	.03	-.01	.01
Low	.11	+.05**	+.01	British-born	.96	<.01	+.01
Parent long-term unemployed	.09	+.03**	+.01	Free school meals	.08	+.02*	+.01
Predict child continues past 11^{th} grade	.88	-.10**	-.01	Special educational needs	.15	+.07**	+.01
Parent Saving for university	.31	-.10**	-.01				
Home Internet	.93	-.04**	-.01	*N* =	2,119	2,181	2,180

Notes: Asterisks indicate significant difference from treatment group: ** p<0.01, * p<0.05

In Table 1.4, model covariates are presented for the treated, all untreated, and matched untreated samples. When comparing treated students to all untreated students, 20 of the 46 covariates are significantly different at the five percent level. In contrast, just one of the covariates was significantly different between the treatment and matched sample groups, indicating that the analytical model meets the covariate balance assumption.

Model estimates

My first research question asks whether perceived teacher encouragement influences enrollment in A-levels. Results for this outcome are presented in Table 1.5. For all students who report encouragement, the ATT estimate is 7.7 percentage points, representing an increase in enrollment rates of 12%. While this is markedly smaller than the raw gap between all treated and untreated students (18 percentage points), it is still large enough to suggest that the ATT of teacher encouragement is significant at any conventional level.

Table 1.5. Average treatment effect on treated (ATT): A-level study

	Treated (percent)	Matched untreated (percent)	ATT (pp increase)	ATT (% increase)
All	73.9	66.2	7.7(1.5)**	12**
Parental education				
No qualifications	63.5	52.0	11.5(3.9)**	22**
10^{th} grade qualification	66.7	53.9	12.8(3.1)**	24**
12^{th} grade qualification	67.4	63.2	4.2(3.6)	7
University degree	85.6	80.1	5.5(2.1)*	7*
Academic achievement				
Lower	35.0	26.6	8.4(3.3)*	32*
Middle	64.0	52.3	11.7(2.6)**	22**
Upper	91.2	87.9	3.4(1.5)*	4*

Notes: * $p<0.05$; ** $p<0.01$

Further, the impact of teacher encouragement appears to be greater for students with lower levels of parental education. For those from the two lower groups (no qualifications and tenth-grade qualifications), A-level enrollment rates increase by 11.5 and 12.8 percentage points, respectively, both of which are significant at the .01 level. The effect of encouragement is relatively small for the other two groups (twelfth-grade qualifications and university degree holders), by 4.2 and 5.5 percentage points, respectively, only the latter of which is significant at the .05 level. The impact of encouragement is significant for all three of the achievement groupings, although the size of the impact is far greater for students in the lower and middle thirds.

Table 1.6. Average treatment effect on treated (ATT): enrollment in any university

	Treated (percent)	Matched untreated (percent)	ATT (pp increase)	ATT (% increase)
All	57.4	53.3	4.1(1.6)**	8**
Parental education				
No qualifications	49.4	39.1	10.3(3.9)**	26**
10th grade qualification	45.9	38.6	7.3(3.1)*	19*
12th grade qualification	50.9	47.8	3.1(3.7)	8
University degree	70.3	70.1	0.1(2.5)	0
Academic achievement				
Lower	21.7	17.1	4.6(2.8)	27
Middle	46.0	35.6	10.4(2.6)**	29**
Upper	74.9	76.5	–1.6(2.0)	-2

Notes: * $p<0.05$; ** $p<0.01$.

Table 1.6 presents ATT estimates for the impact of encouragement on enrollment in any university. Across all students, the difference between treated and matched students provides an ATT of 4.1 percentage points, which is significant at the .01 level. Once again, the impact of teacher encouragement varies by parental education, proving significant only at the lower levels (no qualifications or tenth grade qualifications). For these groups, there are significant increases in the likelihood of university enrollment, of 10.3 and 7.3 percentage points, respectively, representing an increase of about a fifth over the matched comparison group. When grouped by academic achievement, aggregate enrollment among the middle third of students increases significantly, by 10.4 percentage points, equivalent to an increase of around a quarter in the probability of university enrollment. However, no change is observable for the upper and lower thirds, indicating that teacher encouragement appears only to have an observable impact on those in the middle third.

Table 1.7. Average treatment effect on treated (ATT): enrollment in prestigious university

	Treated (percent)	Matched untreated (percent)	ATT (pp increase)	ATT (% increase)
All	12.9	10.1	2.8(0.9)**	28**
Parental education				
No qualifications	5.4	5.2	0.2(1.7)	4
10th grade qualification	5.6	3.2	2.4(1.2)*	75*
12th grade qualification	11.5	7.5	4.0(2.1)	41
University degree	20.7	17.0	3.7(1.9)	25
Academic achievement				
Lower	1.2	0.3	0.9(0.7)	300
Middle	3.0	1.6	1.4(0.8)	88
Upper	22.4	19.0	3.4(1.8)	18

Notes: * $p<0.05$; ** $p<0.01$. Standard errors for ATT estimates are shown in parentheses.

Table 1.7 presents the results of all matching models for which the outcome of interest is enrollment in a prestigious university. The impact of teacher encouragement on the aggregate of treated students is statistically significant, with an ATT estimate of 2.8 percentage points. Although this is smaller than the percentage point difference for attendance at all universities, it represents a larger percentage increase (28%) in the likelihood of the outcome than the aggregate ATT for enrollment at any university (8%).

Whereas encouragement led to an increase in likelihood of enrollment in a high-prestige university for all parent education groups, only the tenth grade qualification groups was significant at the .05 level, although it is worth noting that estimates for the two higher levels (twelfth grade and degree qualifications) are also close to this level of significance. Regarding academic achievement, almost all students who enrolled in a prestigious university were already in the highest third prior to tenth grade, although the impact teacher encouragement on the likelihood of these students attending a prestigious university was not quite significant at the .05 level. There was an increase for the middle achievement group, from 1.6% among the matched comparison group to 3.0% among those who received encouragement, but, as with the upper third of students, the impact of encouragement was close to but not significant at the .05 level. Almost no students in the bottom third went on to enroll in a prestigious university, regardless of whether they received encouragement.

Discussion

Limitations

One key limitation of the propensity-score matching approach is its reliance on observed variables. However perfect the match between treatment and control groups on these variables, there will always be some degree of uncertainty on the quality of matching on unobserved variables. I have sought to moderate this shortcoming, both through guiding my selection of variables via the past literature and testing for sensitivity to omitted variables with the Mantel-Haenszel test (see Appendix 1.B).

Another methodological limitation relates to measurement of the 'treatment' – a teacher encouraging a tenth grade student to stay on for the non-compulsory years of high school – since it relies entirely on student reports. There is no means of corroborating student assessments of encouragement with the perspective of their teachers. However, while this reliance on student self-reports may encourage readers to be cautious in accepting the findings presented here, it is useful in itself to learn from students' perceptions of receiving encouragement.

Although the LSYPE cohort completed compulsory high school in 2006, it is not possible to make a straightforward extrapolation of these findings onto school students in the present day. Since 2013, England's government has made it compulsory for youth to remain some form of education or training until age 17, and in 2015 the government extended this requirement until age 18. There is not yet empirical evidence of the policy's impact on youth enrollment behavior. It is important to note though that the government's definition of education or training is broad: apprenticeships, part-time education, or training while employed suffice. The fact that I operationalize post-compulsory enrollment as studying for A-levels, the most academically rigorous option for most students, leads me to think that it is unlikely my findings would now be rendered redundant by a policy change that enforces only limited levels of education or training.

Another contextual difference is that students now face higher financial barriers to continuing in education after tenth grade. LSYPE respondents had access to the Education Maintenance Allowance, which made weekly payments to low-income students who stayed in the non-compulsory years of high school. In 2011, the government replaced the Education Maintenance Allowance with the 16-19 Bursary Fund, which offers support to a far smaller proportion of students. In addition, the LSYPE cohort finished high school at a point where annual tuition fees for all English universities were set at approximately US$4,500. In contrast, this year's cohort of school leavers face varying fee levels, with an average around US$12,000. Given these financial disincentives, both for completing A-levels and attending university, recent policy changes in England lessen the extent to which these findings can represent conditions in the present day.

Contributions

In spite of these limitations, this study makes an important contribution to the research literature on this subject. My findings support Bourdieu's theorization of habitus, but also diverge from it in an important respect. By indicating that teacher encouragement is influential in students' progress to higher education, my findings support the notion that students' participation in formal education is at least partially dependent on the social cues they receive that legitimize their progress. In addition, my findings support the notion that students respond differently to such cues according to social class.

However, they differ from Bourdieu and Passeron's (1990) claim that teachers are a complicit, rather than a resistant, component of an education system that perpetuates socio-economic disparities. This is because my model estimates suggest that teacher encouragement has a positive impact on future enrollment, and that this impact is greatest for students with lower socio-economic status. As discussed earlier, research on the British education system is divided, with some emphasizing the role of teachers in perpetuating inequality (for example, Ball, 2003; Preston, 2003; Reay, 1998) and others the role of teachers in resisting inequality (for example, Hollingworth and Archer, 2009; Mills, 2008; Oliver and Kettley, 2010). My findings corroborate the latter group. Perhaps this should encourage greater optimism about the role that formal education is able to play in tackling inequality

To my knowledge, this research is the first to provide inferential analysis on the role of student-teacher interactions in university access. Since its methodology is unique among empirical applications of Bourdieu's theorization of habitus, my findings broaden the evidentiary base for proponents of this theoretical tradition. The propensity-score model estimates suggest a sustained benefit to students when they receive teacher encouragement to commit to further education. The outcome is separated from the reported encouragement by around three to four years, yet significant benefits can be observed for recipients of encouragement. These findings add to the research literature by providing more robust evidence that near-immediate effects of teacher encouragement may also enjoy a degree of longevity.

For policy researchers whose focus is university access, it should be informative to receive inferential evidence that relations between students and their high school teachers play a role in the decision to continue in postsecondary education. University access research ought not to conceptualize teachers solely as deliverers of course material and thus academic preparedness, envisaging family and counselors as the significant adults in influencing students' career planning and aspirations. Instead, future access research stands to benefit from accounting for the relational aspects of teacher-student interactions. From the perspective of many students, teachers are likely to represent the most immediate embodiment of the educational system through which they are navigating. Even in the early years of high school, encouragement and support from this source appears to influence the likelihood of students making the transition into the non-compulsory stages of the K-16 'pipeline'.

Policy researchers may also take interest in the evidence of heterogeneous effects not only according to socio-economic status but also to prior attainment. This evidence helps to identify which students are most likely to be influenced. I find that the impact of teacher encouragement on enrollment in A-levels and university degrees are greatest for students in the middle tercile of attainment; this suggests that teacher encouragement has the most impact on enrollment behavior among those on the margin of continuing to non-compulsory education. This finding corroborates past research emphasizing the importance of prior attainment to university access (for example, Anders, 2012a; Chowdry et al., 2013; Marcenaro-Gutierrez et al., 2007). It appears that students with middling attainment are often on the verge of continuing in education; they are more uncertain than those with higher attainment (who are sure they should continue) or those with lower attainment (who are sure they should not continue). Teacher encouragement seems to offer an effective means to improve enrollment rates among these marginal students.

There is not commensurate evidence though of a subset of marginal students for whom the impact encouragement on attending a prestigious university is especially high. Estimates from models with alternative kernel specifications tend to be significant at the .05 significance level (Appendix 1.C), but I prefer to rely upon the more conservative estimates when using the Epanechnikov kernel. Still, given that these conservative estimates are close to the .05 significance level, they are too ambiguous to definitively reject the presence of heterogeneous effects for attendance at a prestigious university. Heterogeneous effects for this outcome may not exist, but it is also plausible that they do but my current estimates are not able to detect them. This could be due to insufficient sample size, or my categorization of students into ability terciles may fail to identify effects among a more specific ability band, be it within or across the middle and upper terciles.

This study's findings should also be of interest to schoolteachers and policymakers. Its most straightforward, but arguably most important, implication is for teachers: if their encouragement is having an observable impact on future educational outcomes for students, this information should be shared with them. Many teachers are taking the initiative to encourage students to progress with education beyond the stages in which they will work directly with them, but it is unclear whether they know that this is having an impact. Still, it is worth reiterating that I focus on the average treatment effect on the treated, which only estimates the average impact of encouragement on those who received it, not on those who did not. Hence, it would be unreasonable to extrapolate this study's model estimates to all children or to conclude that teachers should start providing encouragement as widely as possible. Nonetheless, if this paper's findings confirm to teachers the importance of their encouragement to students, this may induce positive feedback loops in their behavior.

I began by discussing the importance of recognizing the dominance of some trends in policymaking, and the relative diminution of others. It is in this respect that this study might inform the work of policymakers. Britain's current coalition government has emphasized schooling as its main policy focus for improving educational opportunity and access to higher education. This is logical: as mentioned earlier, university access disparities by socio-economic group are high, but close to non-existent when only considering those who complete A-levels. However, within the broad topic of schooling policy, academic achievement has received almost the entirety of the government's attention. Further, policymakers predominantly discuss teachers' ability to raise academic achievement by improving pedagogy and discipline.

The mechanisms by which teachers can improve university access are thus bounded at two levels: what they can contribute (improving academic achievement), and how they can do this (pedagogy and discipline). While policy necessitates prioritization, this conception of teachers is reductionist. Course delivery and classroom management offer important but not definitive sources of focus; teachers also have the opportunity to develop a more relational role, providing a personal intermediary between student and educational system. My findings suggest that students are responsive to non-formal cues and legitimation; as the primary figureheads of formal education, teachers may have more forms of influencing inequality than we currently acknowledge.

References

Adnett, N. & Tlupova, D. (2007). *Variable tuition fees and widening participation: Issues and some early evidence*. Paper presented at the British Educational Research Association Annual Conference, London, UK. Retrieved from http://www.staffordshire.ac.uk/schools/business/iepr/docs/workingpaper37.pdf

Allen, K. (2013). ‘Blair's children’: young women as ‘aspirational subjects’ in the psychic landscape of class. *The Sociological Review*. doi: 10.1111/1467-954X.12113

Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage Publications

Anders, J. (2012). *Using the Longitudinal Study of Young People in England for research into higher education access* (No. 12-13). Department of Quantitative Social Science-Institute of Education, University of London.

Anders, J., & Micklewright, J. (2013). *Teenagers' expectations of applying to university: how do they change?* (No. 13-13). Department of Quantitative Social Science. London, UK: Institute of Education, University of London.

Archer, L., Hollingworth, S., & Halsall, A. (2007). 'University's not for me—I'm a Nike person': urban, working-class young people's negotiations of style, identity and educational engagement. *Sociology*, *41*(2), 219-237.

Archer, L., & Hutchings, M. (2000). 'Bettering yourself'? Discourses of risk, cost and benefit in ethnically diverse, young working-class non-participants' constructions of higher education. *British Journal of Sociology of Education*, *21*(4), 555-574.

Archer, L., Hutchings, M. & Ross, A. (2003). *Higher education and social class: issues of exclusion and inclusion.* London, UK: Routledge Falmer.

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2012). Science aspirations, capital, and family habitus: how families shape children’s engagement and identification with science. *American Educational Research Journal*, *49*(5), 881-908.

Archer, L., & Leathwood, C. (2003). Identities, inequalities and higher education. In L. Archer, M. Hutchings, & A. Ross (Eds.) *Higher education and social class: Issues of exclusion and inclusion* (pp. 171-191). London: Routledge Falmer.

Archer, L., Pratt, S. D., & Phillips, D. (2001). Working-class men's constructions of masculinity and negotiations of (non) participation in higher education. *Gender and Education*, *13*(4), 431-449.

Archer, M. S. (2000). *Being human: The problem of agency*. Cambridge: Cambridge University Press.

Baker, S., & Brown, B. J. (2008). Habitus and homeland: educational aspirations, family life and culture in autobiographical narratives of educational experience in rural Wales. *Sociologia Ruralis*, *48*(1), 57-72.

Ball, S. J. (2003). *Class strategies and the education market: The middle classes and social advantage*. London, UK: Routledge Falmer.

Ball, S. J. (2008). *The education debate: Policy and Politics in the Twenty-First Century.* Bristol, UK: The Policy Press.

Ball, S. J. (2012). *Politics and policy making in education: Explorations in policy sociology*. London, UK: Routledge Falmer.
Ball, S. J., Davies, J., David, M., & Reay, D. (2002). 'Classification' and 'judgement': social class and the 'cognitive structures' of choice of higher education. *British Journal of Sociology of Education*, *23*(1), 51–72.
Ball, S. J., Reay, D., & David, M. (2002). 'Ethnic choosing': minority ethnic students, social class and higher education choice. *Race, ethnicity and education*, *5*(4), 333-357.
Berzin, S. C. (2010). Educational aspirations among low-income youths: Examining multiple conceptual models. *Children & Schools*, *32*(2), 112-124.
Blanden, J. & Machin, S. (2004). Educational inequality and the expansion of UK higher education. *Scottish Journal of Political Economy*, *51*(2), 230–249.
Blundell, R., Dearden, L., & Sianesi, B. (2005). Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *168*(3), 473-512.
Boaler, J. (1997). Setting, social class and the survival of the quickest, *British Educational Research Journal*. *23*(5), 575–595.
Bourdieu, P. (1986). *Distinction: A social critique of the judgement of taste*. London: Routledge & Kegan Paul.
Bourdieu, P. (1990). *The logic of practice*. Cambridge: Polity Press.
Bourdieu, P. (1998a). *Practical reason: On the theory of action*. Stanford, CA: Stanford University Press.
Bourdieu, P. (1998b). *The state nobility: elite schools in the field of power*. Stanford, CA: Stanford University Press.
Bourdieu, P. (2005). *The social structures of the economy*. Cambridge: Polity Press.
Bourdieu, P. & Passeron, J. C. (1977). *Reproduction in Education, Society, and Culture*. Beverly Hills, CA: Sage.
Bourdieu, P., & Passeron, J. C. (1979). *The inheritors: French students and their relation to culture*. Chicago: University of Chicago Press.
Bourdieu, P., & Passeron, J. C. (1990). *Reproduction in education, society and culture*. London: Sage.
Bourdieu, P., & Wacquant, L. (1992). *An invitation to reflexive sociology*. Chicago, IL: University of Chicago Press.
Brooks, R., & Everett, G. (2008). The prevalence of "life planning": evidence from UK graduates. *British Journal of Sociology of Education. 29*(3) 325–37.
Burke, P. J. (2010). Masculinity, Subjectivity and Neoliberalism in Men's Accounts of Migration and Accessing Higher Education'. *Gender and Education*, *23*(2), 169-184.
Burke, P. J. (2012). *The right to higher education: Beyond widening participation*. Abingdon, UK: Routledge.
Byrom, T., & Lightfoot, N. (2012). Transformation or transgression? Institutional habitus and working class student identity. *Journal of Social Sciences, 8*(2), 126-134.
Caliendo, M. & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31–72.
Carneiro, P., & Heckman, J. J. (2002). The evidence on credit constraints in post−secondary schooling. *The Economic Journal, 112*(482), 705-734.

Carneiro, P. and Heckman, J. J. (2003) Human capital policy. In J.J. Heckman, A. Krueger, & B. Friedman (Eds.), *Inequality in America: What role for human capital policies?* (77–239). Cambridge, MA: MIT Press.

Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2008). *Widening participation in higher education: analysis using linked administrative data*. London, UK: Institute for Fiscal Studies.

Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2013). Widening participation in higher education: analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(2), 431-457.

Chowdry, H., Dearden, L., Hin, W., & Lloyd, B. (2012). *Fees and student support under the new higher education funding regime: what are different universities doing? IFS Briefing Note BN 134.* London, UK: Institute for Fiscal Studies.

Christie, H. & Munro, M. (2003). The Logic of Loans: students' perceptions of the costs and benefits of the student loan. *British Journal of Sociology of Education*, *24*(5), 621-636.

Cochrane, M. (2007). *Spoilt for choice? Pupil perceptions of the options process at Year 9*. Paper presented at the British Educational Research Association Annual Conference, London, 5-8 September. Retrieved from: http://www.leeds.ac.uk/educol/documents/165852.htm

Cochrane, M. (2011). Children's university aspirations and the effects of cultural and social capital. In J. Adams, M. Cochrane, & L. Dunne (Eds.), *Applying theory to educational research: An introductory approach with case studies* (pp. 95-107). Hoboken: John Wiley & Sons.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, N.J.: L. Erlbaum Associates.

Connor, H., Tyers, C., Modood, T. & Hillage, J. (2004). *Why the difference?: a closer look at higher education minority ethnic students and graduates*. Nottingham, UK: DfES Publications.

Cox, C. B., & Dyson, A. E. (1969). *The crisis in education* (Black Paper 2). London: Critical Quarterly Society.

Crossley, N. (2001). The phenomenological habitus and its construction. *Theory and Society*, *30*(1), 81-120.

Davies, P., Slack, K., Hughes, A., Mangan, J. & Vigurs, K. (2008). *Knowing where to study? Fees, bursaries and fair access*. London, UK: The Sutton Trust.

Dehejia, R. (2005). Practical propensity score matching: A reply to Smith and Todd. *Journal of Econometrics, 125(*1–2), 355–364.

Dehejia, R. H. & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, *94*(448), 1053–1062.

Dehejia, R. H. & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and Statistics, 84*(1), 151–161.

Department for Education and National Centre for Social Research (2012). *Longitudinal Study of Young People in England: Waves One to Seven, 2004-2010. 12th Edition.* Colchester, UK: UK Data Archive. SN: 5545.

Durkheim, E. (1973). *Moral education*. London: Collier Macmillan. . (Original work published 1925.)
Eisenhart, M. A. (2009). Generalizing from educational research. In K. Ercikan & W. M. Roth (Eds.), *Generalizing from educational research: Beyond qualitative and quantitative polarization* (pp. 51–66). New York, NY: Routledge.
Franklin, M. (1995). The effects of differential college environments on academic learning and student perceptions of cognitive development. *Research in Higher Education*, *36*(2), 127-153.
Frolich, M. (2004). Finite-sample properties of propensity-score matching and weighting estimators. *Review of Economics and Statistics, 86*(1), 77–90.
Galindo-Rueda, F., Marcenaro-Gutierrez, O., & Vignoles, A. (2004). The widening socio-economic gap in UK higher education. *National Institute Economic Review*, *190*(1), 75-88.
Gibbons, S. & Vignoles, A. (2012). Geography, choice and participation in higher education in England. *Regional science and urban economics*, *42*(1), 98-113.
Giddens, A. (1991). *Modernity and self-identity: Self and society in the late modern age*. Stanford, CA: Stanford University Press.
Gillard, D. (2008). *Us and Them: a history of pupil grouping policies in England's schools.* Retrieved from www.educationengland.org.uk/articles/27grouping.html
Glennerster, H. (2001). *United Kingdom education, 1997-2001*. London, UK: London School of Economics and Political Science.
Gorard, S., & Selwyn, N. (1999). Switching on the learning society?-questioning the role of technology in widening participation in lifelong learning. *Journal of Education Policy*, *14*(5), 523-534.
Gunn, S. (2005). Translating Bourdieu: cultural capital and the English middle class in historical perspective. *The British journal of sociology*, *56*(1), 49-64.
Guo, S. & Fraser M. W. (2010). *Propensity Score Analysis: Statistical Methods and Applications.* Thousand Oaks, CA: Sage.
Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, *65*(2), 261-294.
Hernandez−Martinez, P., & Williams, J. (2013). Against the odds: resilience in mathematics students in transition. *British Educational Research Journal*,*39*(1), 45-59.
Holdsworth, C. (2006). 'Don't you think you're missing out, living at home?' Student experiences and residential transitions. *The Sociological Review*, *54*(3), 495-519.
Holland, P. W. (1986). Statistics and causal inference. *Journal of the American statistical Association*, *81*(396), 945-960.
Hollingworth, S., & Archer, L. (2009). Urban schools as urban places: school reputation, children's identities and engagement with education in London. *Urban Studies, 47*(3), 584–603.
Hollway, W. & Jefferson, T. (2000). *Doing qualitative research differently*. London: Sage.
Ireson, J., & Hallam, S. (1999). Raising standards: is ability grouping the answer?. *Oxford review of education*, *25*(3), 343-358.
Jackson, M. (2011). *Social Background and Educational Transitions in England*. Paper presented at the annual meeting of the American Sociological Association Annual Meeting, Caesar's Palace, Las Vegas, NV.

Jackson, M. & Jonsson, J. O. (2013). Why does Inequality of Educational Opportunity Vary Across Countries? Primary and Secondary Effects in Comparative Context. In M. Jackson (Ed.), *Determined to Succeed? Performance versus Choice in Educational Attainment* (pp. 306-338). Stanford, CA: Stanford University Press.

Jenkins, R. (1982). Pierre Bourdieu and the reproduction of determinism. *Sociology*, *16*(2), 270-281.

Jenkins, R. (1992). *Pierre Bourdieu*. London: Routledge.

Jones, M. P. (1996). Indicator and stratification methods for missing explanatory variables in multiple linear regression. *Journal of the American Statistical Association*, *91*(433), 222-230.

Lareau, A. (1997) Social Class Differences in Family-School Relationships: The Importance of Cultural Capital. In A.H. Halsey, H. Lauder, P. Brown, A.S. Wells (Eds.) *Education: Culture, Economy, Society* (pp. 713–717). Oxford: Oxford University Press.

Lau, R. W. (2004). Habitus and the practical logic of practice: an interpretation. *Sociology*, *38*(2), 369-387.

Lee, L. (2008). The impact of young people's internet use on class boundaries and life trajectories. *Sociology*, *42*(1), 137-153.

Leuven, E. & Sianesi, B. (2003). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*. Retrieved from http://ideas.repec.org/c/boc/bocode/s432001.html. Version 4.0.6

Little, R. J. (1992). Regression with missing X's: a review. *Journal of the American Statistical Association*, *87*(420), 1227-1237.Anders, J. (2012a). The Link between Household Income, University Applications and University Attendance. *Fiscal Studies*, *33*(2), 185–210.

Lizardo, O. (2004). The cognitive origins of Bourdieu's habitus. *Journal for the Theory of Social Behaviour, 34*(4), 375-401.

Machin, S. & Vignoles, A. (2004). Educational inequality: the widening socio-economic gap. *Fiscal Studies*, *25*(2), 107–128.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719-48.

Marcenaro-Gutierrez, O., Galindo-Rueda, F., & Vignoles, A. (2007). Who actually goes to university? *The Economics of Education and Training, 32*(2), 333–357.

Marsh, H. W., Craven, R. G., Hinkley, J. W., & Debus, R. L. (2003). Evaluation of the Big-Two-Factor Theory of academic motivation orientations: An evaluation of jingle-jangle fallacies. *Multivariate Behavioral Research*, *38*(2), 189-224.

McHarg, J., Mattick, K., & Knight, L. V. (2007). Why people apply to medical school: Implications for widening participation activities. *Medical Education*, *41*(8), 815-821.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., & Urdan, T. (2000). Manual for the Patterns of Adaptive Learning Scales (PALS). Ann Arbor: University of Michigan.

Mills, C. (2008). Reproduction and transformation of inequalities in schooling: the transformative potential of the theoretical constructs of Bourdieu. *British Journal of Sociology of Education*, *29*(1), 79-89.

Moogan, Y. J. (2011). An analysis of school pupils'(with low social economic status) perceptions of university, regarding programmes of study. *Educational Studies*, *37*(1), 1-14.
Mouzelis N. (1995). *Sociological theory: What went wrong?* London: Routledge.
Nash, R. (1999). Bourdieu, 'habitus', and educational research: is it all worth the candle? *British Journal of Sociology of Education*, *20*(2), 175-187.
Nash, R (2005). Cognitive *habitus* and collective intelligence: concepts for the explanation of inequality of educational opportunity. *Journal of Education Policy*, *20*(1), 3-21.
Oliver, C., & Kettley, N. (2010). Gatekeepers or facilitators: The influence of teacher habitus on students' applications to elite universities. *British Journal of Sociology of Education*, *31*(6), 737-753.
Ovenden, K. (2000). The politics of protest. *Socialist Review*, 242. Retrieved from http://pubs.socialistreviewindex.org.uk/sr242/contents.htm
Pennell, H. & West, A. (2005). The Impact of Increased Fees on Participation in Higher Education in England. *Higher Education Quarterly*, *59*(2), 127–137.
Piesse, A., & Kalton, G. (2009). *A strategy for handling missing data in the Longitudinal Study of Young People in England (LSYPE)*. DCSF-RW086. London, UK: Department for children, schools and families.
Pimlott-Wilson, H. (2011). The role of familial habitus in shaping children's views of their future employment. *Children's Geographies*, *9*(1), 111-118.
Pintrich, P. R., Smith, D. A., García, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and psychological measurement*, *53*(3), 801-813.
Preston, J. (2003). White trash vocationalism? Formations of class and race in an Essex further education college. *Widening participation and lifelong learning*, *5*(2), 6-17.
Pugsley, L. (1998). Throwing your brains at it: higher education, markets and choice. *International Studies in Sociology of Education*, *8*(1), 71-92.
Reay, D. (1995). 'They employ cleaners to do that': habitus in the primary classroom. *British journal of sociology of education*, *16*(3), 353-371.
Reay, D. (1998). 'Always knowing' and "never being sure": familial and institutional habituses and higher education choice. *Journal of Education Policy*, *13*(4), 519–529.
Reay, D., (2001). Finding or losing yourself?: Working-class relationships to education. *Journal of Education Policy*, 16, 333-346.
Reay, D. (2004). "It's all becoming a habitus": beyond the habitual use of Pierre Bourdieu's concept of habitus in educational research. *British Journal of Sociology of Education, 25*(4), 431–44.
Reay, D. (2005). Beyond consciousness? The psychic landscape of social class. *Sociology*, *39*(5), 911-928.
Reay, D. (2006). The zombie stalking English schools: social class and educational inequality. *British journal of educational studies*, *54*(3), 288-307.
Reay, D. (2013). Social mobility, a panacea for austere times: tales of emperors, frogs, and tadpoles. *British Journal of Sociology of Education*, *34*(5-6), 660-677.Sweetman, P. (2003). Twenty−first century dis−ease? Habitual reflexivity or the reflexive habitus. *The sociological review*, *51*(4), 528-549.Reay, D., Crozier,

G., & Clayton, J. (2009). 'Strangers in paradise'? Working-class students in elite universities. *Sociology*, *43*(6), 1103-1121.

Reay, D., David, M. E., & Ball, S. J. (2001). Making a difference?: Institutional habituses and higher education choice. *Sociological Research Online, 5*(4).

Reay, D., David, M. E., & Ball, S. J. (2005). *Degrees of choice: social class, race, gender and higher education.* Stoke on Trent, UK: Trentham Books.

Redmond, P. (2006). Outcasts on the inside: graduates, employability and widening participation. *Tertiary Education and Management*, *12*(2), 119-135.

Reynolds, C. L. & DesJardins, S. L. (2009). The use of matching methods in higher education research: Answering whether attendance at a 2-year institution results in differences in educational attainment. *Higher education: Handbook of theory and research*, *24*, 47-97.

Rosenbaum, P. R., & Rubin, D. B. (1983). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society Series B, 45*(2), 212–218.

Rosenbaum, P. R. & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33–38.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, *66*(5), 688.

Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine, 2*(127), 757–763

Sayer, A. (2005). Class, moral worth and recognition. *Sociology*, *39*(5), 947-963.

Shanghai Ranking Consultancy (2013). *Academic ranking of world universities 2013. Shangai, China: Shanghai Jiao Tong University*. Retrieved from http://www.shanghairanking.com/ARWU2013.html

Simeoni, D. (1998). The pivotal status of the translator's habitus. *Target*, *10*(1), 1-39.

Smith, J. A. & Todd, P. E. (2005a). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, *125*(1–2), 305–353.

Smith, J. A. & Todd, P. E. (2005b). Does matching overcome LaLonde's critique of nonexperimental estimators? Rejoinder. *Journal of Econometrics*, *125*(1–2), 365–375.

Steedman, C. (1988). The mother made conscious: the historical development of primary school pedagogy. In M. Woodhead & A. McGrath (Eds.), *Family, school & society* (pp. 82-95). London: The Open University.

Stevenson, H., & Wood, P. (2014). Markets, managerialism and teachers' work: the invisible hand of high stakes testing in England. *International Education Journal: Comparative Perspectives*, *12*(2).

Thomas, L., Bland, D., & Duckworth, V. (2012). Teachers as advocates for widening participation. *Widening Participation and Lifelong Learning*, *14*(2), 40-58.

Thornbury, R. (1978). *The changing urban school*. London: Methuen.

Thrupp, M. (1999). *Schools making a difference: Let's be realistic! School mix, school effectiveness and the social limits of reform*. Buckingham: Open University Press.

Times Higher Education (2012). *World University Rankings 2012-2013*. Retrieved February 24, 2013, from http://www.timeshighereducation.co.uk/world-university-rankings/2012-13/world-ranking

U.S. News and World Reports (2012). *World's Best Universities; Top 400*

Universities in the World. Retrieved from http://www.usnews.com/education/worlds-best-universities-rankings/top-400-universities-in-the-world

Vignoles, A. (2013). Widening participation and social mobility. In C. Callendar & P. Scott (Eds.), *Browne and beyond: Modernizing English higher education* (pp. 112–129). London: Institute of Education Press.

Vincent, C., & Martin, J. (2002). Class, culture and agency: researching parental voice. *Discourse*, *23*(1), 108-127.

Wacquant, L. (2005). Shadowboxing with ethnographic ghosts: a rejoinder. *Symbolic Interaction*, *28*(3), 441-447.

Appendix 1.A: Missing data

For each row variable, Table 1.A1 presents the percentage of respondents with missing information for a given variable when model-outcome variables were observed. Table 1.A2 presents the correlation of missingness between variables when enrollment in A-levels was observed. Table 1.A3 presents the correlation of missingness between variables for when enrollment in any university degree and a university degree at a prestigious institution were observed.

Table 1.A1. Percentage of observations missing each variable, by outcome.

Variable	A Levels n=11,823	University n=8,664
Encouragement	4.1	2.8
Government Office Region	3.2	2.3
IDACI score	3.2	2.3
Urbaninicity	3.2	2.3
Family income	25.7	24.6
Parental education	5.2	4.7
Parental NSSEC	18.9	9.6
Parental aspirations	4.1	3.3
Saving for HE	3.6	2.5
Internet access	2.1	1.1
Religion	4.1	3.0
KS2 English	6.4	5.5
KS2 Math	6.2	5.4
KS2 Science	6.4	5.6
KS3 English	7.1	6.7
KS3 Math	5.3	5.0
KS3 Science	6.3	6.0
Gender	0.1	0.0
Race/ethnicity	4.2	3.1
Gender	4.8	4.9
Special educational needs	4.3	3.4

Table 1.A2. Correlation of missing variables for A-level models.

	SEN	Encouragement	Region	IDACI	Urbanicity	Income	Parental education	NSSEC	Parent aspirations	Parent saving	Internet	Religion	KS2 Eng.	KS2 Math	KS2 Sci.	KS3 Eng.	KS3 Math	KS3 Sci.	Gender	Ethnicity	FSM
SEN	1																				
Encouragement	0.00	1																			
Region	-0.01	-0.04	1																		
IDACI	0.06	0.03	-0.14	1																	
Urbanicity	-0.02	0.00	0.09	-0.24	1																
Income	-0.02	-0.03	0.08	-0.40	0.11	1															
Parental education	0.04	0.04	-0.08	0.40	-0.12	-0.43	1														
NSSEC	0.08	0.04	-0.06	0.44	-0.11	-0.47	0.54	1													
Parent aspirations	0.00	0.00	0.01	0.00	0.01	-0.03	0.08	0.01	1												
Parent saving	0.06	-0.01	0.06	-0.07	-0.01	0.13	-0.20	-0.13	-0.15	1											
Internet	0.04	0.01	-0.05	0.24	-0.05	-0.25	0.24	0.25	0.01	-0.16	1										
Religion	0.04	0.02	-0.03	0.17	-0.09	-0.12	0.18	0.19	-0.05	0.16	0.03	1									
KS2 English	-0.06	-0.02	0.06	-0.25	0.05	0.26	-0.35	-0.31	-0.05	0.37	-0.23	-0.06	1								
KS2 Math	-0.08	-0.04	0.03	-0.23	0.05	0.25	-0.32	-0.28	-0.04	0.35	-0.22	-0.06	0.75	1							
KS2 Science	-0.08	-0.03	0.04	-0.25	0.07	0.26	-0.34	-0.29	-0.03	0.30	-0.20	-0.12	0.75	0.79	1						
KS3 English	-0.07	-0.01	0.10	-0.30	0.07	0.28	-0.38	-0.34	-0.10	0.43	-0.25	-0.05	0.81	0.68	0.66	1					
KS3 Math	-0.10	-0.04	0.08	-0.31	0.08	0.29	-0.37	-0.34	-0.06	0.40	-0.26	-0.06	0.75	0.88	0.74	0.77	1				
KS3 Science	-0.11	-0.03	0.08	-0.33	0.10	0.30	-0.40	-0.35	-0.05	0.40	-0.26	-0.09	0.76	0.77	0.77	0.80	0.88	1			
Gender	0.01	0.03	0.02	0.01	0.00	-0.03	0.01	0.02	-0.03	0.12	0.03	0.03	0.15	-0.05	-0.01	0.17	-0.02	0.00	1		
Ethnicity	0.09	0.00	0.06	0.38	-0.17	-0.22	0.21	0.26	-0.05	0.27	0.06	0.39	-0.10	-0.11	-0.17	-0.10	-0.13	-0.17	0.02		
Free school meals	0.12	0.01	-0.04	0.40	-0.10	-0.42	0.36	0.51	0.00	-0.02	0.21	0.16	-0.22	-0.20	-0.21	-0.25	-0.25	-0.27	0.02	0.26	1

Table 1.A3. Correlation of missing variables for university models.

	SEN	Encouragement	Region	IDACI	Urbanicity	Income	Parental education	NSSEC	Parent aspirations	Parent saving	Internet	Religion	KS2 Eng.	KS2 Math	KS2 Sci.	KS3 Eng.	KS3 Math	KS3 Sci.	Gender	Ethnicity	FSM
SEN	1																				
Encouragement	0.00	1																			
Region	-0.01	-0.05	1																		
IDACI	0.07	0.04	-0.12	1																	
Urbanicity	-0.03	0.01	0.09	-0.24	1																
Income	-0.02	-0.04	0.07	-0.40	0.10	1															
Parental education	0.04	0.05	-0.08	0.40	-0.12	-0.43	1														
NSSEC	0.09	0.04	-0.05	0.43	-0.10	-0.47	0.55	1													
Parent aspirations	0.00	-0.01	0.00	-0.02	0.02	-0.03	0.08	0.00	1												
Parent saving	0.05	-0.02	0.04	-0.06	-0.01	0.12	-0.20	-0.13	-0.13	1											
Internet	0.04	0.02	-0.03	0.22	-0.04	-0.24	0.24	0.25	0.02	-0.14	1										
Religion	0.03	0.02	-0.04	0.18	-0.09	-0.13	0.19	0.19	-0.06	0.14	0.04	1									
KS2 English	-0.06	-0.04	0.05	-0.24	0.04	0.25	-0.35	-0.32	-0.09	0.37	-0.22	-0.07	1								
KS2 Math	-0.08	-0.06	0.03	-0.23	0.04	0.24	-0.32	-0.29	-0.07	0.35	-0.21	-0.07	0.75	1							
KS2 Science	-0.07	-0.05	0.03	-0.24	0.06	0.25	-0.33	-0.30	-0.06	0.31	-0.20	-0.13	0.76	0.79	1						
KS3 English	-0.07	-0.02	0.09	-0.30	0.06	0.26	-0.38	-0.34	-0.08	0.42	-0.24	-0.07	0.81	0.68	0.67	1					
KS3 Math	-0.08	-0.06	0.07	-0.30	0.08	0.29	-0.37	-0.34	-0.07	0.40	-0.25	-0.06	0.75	0.89	0.75	0.77	1				
KS3 Science	-0.09	-0.05	0.07	-0.32	0.09	0.29	-0.40	-0.36	-0.07	0.40	-0.25	-0.10	0.76	0.78	0.78	0.80	0.88	1			
Gender	0.01	0.04	0.02	0.02	0.00	-0.05	0.02	0.03	-0.04	0.10	0.04	0.04	0.13	-0.07	-0.03	0.14	-0.05	-0.03	1		
Ethnicity	0.09	0.01	0.05	0.39	-0.18	-0.24	0.23	0.28	-0.07	0.24	0.08	0.38	-0.11	-0.12	-0.18	-0.12	-0.14	-0.19	0.04	1	
Free school meals	0.11	0.00	-0.03	0.40	-0.09	-0.42	0.36	0.51	0.00	-0.02	0.21	0.16	-0.22	-0.20	-0.20	-0.25	-0.25	-0.26	0.03	0.28	1

Appendix 1.B: Sensitivity to omitted variables

This appendix provides information about the bandwidth and sensitivity tests used in each matching model. All matching models were produced using kernel matching with the Epanechnikov kernel, and Table 1.B1 presents information about the bandwidth used for each matching model (for a more detailed discussion of kernel and bandwidth choice, see DesJardins and Reynolds, 2009).

Stata's *mhbounds* program (Becker and Caliendo, 2007) was used to calculate the Mantel-Haenszel test statistic (Mantel and Haenszel, 1959) for each model. This statistic tests a model's sensitivity to confounding factors, i.e., unobserved variables that might influence both assignment to encouragement and the likelihood of attending university. The Mantel-Haenszel test determines how strongly the influence of an unobserved dichotomous variable would need to be in order to undermine the given model. In the following table, each model's sensitivity is reported in terms of the critical value at which the Mantel-Haenszel test statistic significance level exceeds the 0.05 level. For example, the first model's critical test statistic value is 1.6. This indicates that, in order for the 95% confidence interval of the model's ATT to include zero, an unobserved dichotomous variable would need to cause the odds ratio of treatment assignment to differ between the treatment and comparison groups by a factor of 1.6 (Becker and Caliendo, 2007). To provide a reference point, in the logistic regression model used to generate the propensity scores, the greatest odds ratio for any dichotomous variable was 1.5, for the variable asking whether the student's parent expected them to continue beyond GCSEs.

[Table begins on following page]

Table 1.B1. Mantel-Haenszel test results for each model.

Outcome	Subsample	Bandwidth	Sensitivity
A-level study	All	0.11	1.6
	Parent/caregiver holds		
A-level study	No qualifications	0.11	1.55
A-level study	10th grade qualification	0.11	1.4
A-level study	12th grade qualification	0.11	1.35
A-level study	University degree	0.11	1.25
	Academic achievement		
A-level study	Lower	0.11	1.3
A-level study	Middle	0.11	1.6
A-level study	Upper	0.11	1
Attends any university	All	0.1	1.6
	Parent/caregiver holds		
Attends any university	No qualifications	0.1	1.55
Attends any university	10th grade qualification	0.1	1.55
Attends any university	12th grade qualification	0.1	1.35
Attends any university	University degree	0.1	1.25
	Academic achievement		
Attends any university	Lower	0.1	1.1
Attends any university	Middle	0.1	1.6
Attends any university	Upper	0.1	1
Attends prestigious university	All	0.11	1.65
	Parent/caregiver holds		
Attends prestigious university	No qualifications	0.11	1
Attends prestigious university	10th grade qualification	0.11	1.4
Attends prestigious university	12th grade qualification	0.11	1.35
Attends prestigious university	University degree	0.11	1.35
	Academic achievement		
Attends prestigious university	Lower	0.1	1
Attends prestigious university	Middle	0.1	1.2
Attends prestigious university	Upper	0.1	1.1

Notes: All models use kernel matching with the Epanechnikov kernel. The "Sensitivity" column presents the critical value at which the Mantel-Haenszel (Mantel and Haenszel, 1959) test statistic's significance level exceeds the 0.05 level.

Appendix 1.C: Alternative model specifications

Table 1.C1. ATT results for models when re-run with alternate kernel specifications.

		Kernel type			
Outcome	Subsample	Epan.	Uniform	Normal	Biweight
A-levels	All	.077 (.015)	.084 (.014)	.106 (.014)	.074 (.014)
	Parent/caregiver holds				
A-levels	No qualifications	.115 (.039)	.110 (.037)	.130 (.037)	.101 (.037)
A-levels	10th grade qualification	.128 (.031)	.133 (.030)	.158 (.030)	.119 (.030)
A-levels	12th grade qualification	.042 (.036)	.060 (.034)	.082 (.034)	.050 (.034)
A-levels	University degree	.055 (.021)	.056 (.021)	.073 (.021)	.051 (.021)
	Academic achievement				
A-levels	Lower	.084 (.033)	.075 (.030)	.084 (.030)	.070 (.030)
A-levels	Middle	.117 (.026)	.132 (.025)	.145 (.025)	.125 (.025)
A-levels	Upper	.034 (.015)	.036 (.015)	.041 (.015)	.033 (.015)
University	All	.041 (.016)	.048 (.015)	.072 (.015)	.036 (.015)
	Parent/caregiver holds				
University	No qualifications	.103 (.039)	.093 (.036)	.116 (.036)	.082 (.036)
University	10th grade qualification	.073 (.031)	.086 (.029)	.108 (.029)	.071 (.029)
University	12th grade qualification	.031 (.037)	.043 (.035)	.065 (.034)	.033 (.034)
University	University degree	.001(.025)	.007 (.024)	.025 (.024)	-.002(.024)
	Academic achievement				
University	Lower	.046 (.028)	.034 (.026)	.043 (.026)	.027 (.026)
University	Middle	.104 (.026)	.121 (.024)	.132 (.024)	.114 (.024)
University	Upper	-.016(.020)	-.015(.020)	-.009(.020)	-.019(.020)
Prestigious university	All	.028 (.009)	.032 (.009)	.041 (.009)	.027 (.009)
	Parent/caregiver holds				
Prestigious university	No qualifications	.002 (.017)	.008 (.015)	.013 (.015)	.006 (.015)
Prestigious university	10th grade qualification	.024 (.012)	.026 (.011)	.029 (.011)	.023 (.011)
Prestigious university	12th grade qualification	.040 (.021)	.034 (.019)	.039 (.019)	.032 (.020)
Prestigious university	University degree	.037 (.019)	.046 (.018)	.057 (.018)	.039 (.018)
	Academic achievement				
Prestigious university	Lower	.009 (.007)	.010 (.007)	.010 (.007)	.009 (.007)
Prestigious university	Middle	.014 (.008)	.014 (.007)	.015 (.007)	.013 (.007)
Prestigious university	Upper	.034 (.018)	.038 (.018)	.042 (.018)	.035 (.018)

Chapter 2: Do Discrete Grade Boundaries Shape Students' Future Outcomes?

Summary

The research literature in England has highlighted the importance of educational achievement throughout school in predicting subsequent progression to higher education. However, in contrast to existing theorizations of this relationship, exam results may not only demonstrate students' prior academic achievement but also influence their future achievement. I examine the impact of discrete exam grades, independent of attainment, on students' progress to post-compulsory education. Employing a regression discontinuity design, I compare students who have received different grades on national examinations in eighth grade despite performing almost identically. Further, I investigate whether the impact of grades differs according to socio-economic status. Results indicate that both the English and mathematics assessments influence performance in school-leaving exams and enrollment in post-compulsory schooling. For lower socio-economic students, this impact is higher than for other students and extends to university enrollment.

Introduction

Like many other countries, England's recent governments have promoted higher education as a means to national economic growth. Enrollment rates have increased accordingly: by 2012, student enrollment in the United Kingdom stood at over 2.3 million, representing a six-fold increase since the 1960s (Greenaway and Haynes, 2003; Bermingham, 2014). The proportion of citizens aged 18–22 who were enrolled in a degree course rose from 5% in 1960 to 40% in 2013 (Callender, 2006; Vignoles, 2013).

However, this expansion has disproportionately served students from the higher socio-economic classes (Blanden and Machin, 2004; Galindo-Rueda, Marcenaro-Gutierrez, and Vignoles, 2008; Glennerster, 2002; Machin and Vignoles, 2004), and so undergraduate enrollment remains highly stratified (Anders, 2012a; Archer, Hutchings, and Ross, 2003; Chowdry et al., 2013). By 2009, students from the most advantaged quintile of households were six times more likely to attend university than those from the least advantaged quintile (Vignoles and Powdthavee, 2009). Thus, although England's three major political parties all support increasing undergraduate enrollment among poorer students (Conservatives, 2013; Labour, 2010; Liberal Democrats, 2013), concerns about access persist.

While researchers have offered a range of theorizations for this disparity, the strongest inferential studies to date indicate that achievement in the early and middle years of schooling is crucial. Such research demonstrates that achievement measures in national examinations explain a great deal of the gap in undergraduate enrollment across socio-economic classes (see, for example, Anders, 2012a; Chowdry et al., 2013; Marcenaro-Gutierrez et al., 2007; Vignoles and Powdthavee, 2009). However, the mechanisms by which earlier achievement drives future enrollment behavior have not been fully clarified.

In this study, I contribute to the research literature on educational access in England by further developing our understanding of these mechanisms. I do this by testing a key assumption in the education policy literature, namely that national summative assessments reflect students' academic ability in a transparent manner. I test an alternative hypothesis: that the exam results that students receive shape their subsequent educational outcomes, and, furthermore, that this impact may differ according to students' socio-economic class. I do this by separating the effect of information that students receive about their academic achievement from the underlying achievement itself.

More specifically, I aim to distinguish the impact upon a student's educational progress of receiving a lower or higher grade, independent of achievement. For example, do students who receive a higher grade in mathematics attain at a higher level in the future than those who performed indistinguishably but received a lower grade? Through a regression discontinuity design, I compare the impact of receiving the average grade, versus receiving the grade below, on the transition to post-compulsory education. In addition, I test whether impacts differ according to students' socio-economic status, thus exploring whether this factor may help to explain disparities in university enrollment.

The exams used are the Key Stage 3 national assessments in English and mathematics that, until 2009, students took at the end of eighth grade.[1] The key grade boundary of interest lies between the average grade and the grade just below. Grades are defined according to cut-points in the underlying, continuous test score; students, parents, and teachers receive information on the grades but not on the continuous test scores. Examining authorities do not decide the cut-points until after students have sat the exams, and these cut-points change every year. The outcomes of interest are three of the criteria by which academic progress is often judged in English policy debates: (1) achievement in tenth-grade national examinations, (2), enrollment in the non-compulsory grades of high school, and (3) enrollment in a university degree course.

Background and literature

One key feature of England's national curriculum is that students begin to specialize in subjects from an early age. At age 14, students select about 10 subjects – from approximately

1 These assessments are now commonly replaced by an alternative, Cognitive Assessment Tests, the implications of which I discuss later in this chapter.

40 options – on which they will be tested at age 16 in national General Certificates of Secondary Education (GCSE) examinations. Schooling ceases to be compulsory after these exams; those who do not pass five or more GCSEs – around 40% of students (House of Commons Education Committee, 2013) – tend to leave the standard high school system in favor of more vocational courses or employment.

Those that do pass at least five GCSEs are able to continue to the final two years of high school to take A-levels. Students tend to choose three or four A-level subjects, can only study those for which they took GCSEs, and must select specific subjects in order to study a particular subject for university. For example, students hoping to study medicine at university are advised to take advanced science options at GCSEs and have to take chemistry at A-level. In their university applications, students must specify which subject they plan to study. During their undergraduate degree programs, students do not take introductory classes across a range of subjects; instead, they only study courses in either a single- or dual-subject program from the outset.

All undergraduate applications are managed by a single organization: the Universities & Colleges Admissions Service (henceforth UCAS). The application process is largely uniform across institutions: universities have access to candidates' personal statements, anticipated A-level results (as predicted by schoolteachers), and GCSE results. Only a minority of institutions use interviews to further screen applicants.

Students hoping to progress beyond the compulsory stages of education are thus required to choose appropriate GCSE subjects at age 13–14 and perform well in these subjects at age 15–16. GCSEs provide a strong predictor of future university attendance (Chowdry et al., 2013), operating as a "symbolic and material currency in terms of future educational progression" (Davey and Fuller 2013, 3.1). UCAS has been in place for 20 years, GCSE exams for 28 years, and A-levels for over 60 years. For two decades then, the national school curriculum and undergraduate admissions process for English universities have followed a consistent pattern with uniform processes. However, this system also puts pressure on students to envisage coherent academic trajectories and perform well in examinations from mid-adolescence.

Although GCSE results are the first examinations to formally influence students' subsequent opportunities, national assessments begin earlier in the lifespan. In fifth grade, students complete Key Stage 2 assessments in English, science, and mathematics. Until 2009, students would repeat these subjects in Key Stage 3 assessments during eighth grade. In 2009, the UK government abolished testing in Key Stage 3 exams. A standardized alternative, Cognitive Abilities Tests (CATs), have become increasingly widespread in recent years; the producer of these tests claims that they are now used by two thirds of schools (GL assessment, 2015). While the UK government has not mandated CATs, it promotes their use (Department for Education, 2014). In the discussion section of this chapter I will return to the implications of the use of these CATs instead in place of the Key Stage 3 assessments; for all empirical analysis, though, I focus on conditions in which students took the Key Stage 3 assessments.

For the Key Stage 3 assessment in each subject, a given student, their parent(s), and teachers received level scores, equivalent to discrete letter-grade boundaries. Consequently, this grading system allowed for direct comparisons among students. In theory, all national examinations prior to Year 11 are low stakes for students in the sense that their primary role is to assess the performance of schools (Daugherty, 1995), rather than to provide students with qualifications that influence their subsequent study and employment opportunities. Yet, Conner (1999) argued that Key Stage 3 exams had attained a de facto high-stakes status, and the motivational impact on students of knowing how one has been ranked in such assessments should not be overlooked (Broadfoot, 1999).

As noted, the strongest inferential studies to date have established that strong predictors of postsecondary enrollment occur from childhood and early adolescence. Among these predictors, prior attainment in national examinations is the most powerful (Vignoles, 2013). However, while such studies have established the predictive power of attainment measures, the mechanisms by which attainment measures might influence enrollment have not yet been confirmed through quantitative analyses. One starting point is to divide explanations between those for which pre-existing attributes define and are reflected in attainment measures – such as academic aptitude, self-discipline, and confidence – and those for which assessment feedback has some impact in itself, independent of underlying causes of a given level of attainment. In this study, I focus on the latter.

Among policymakers, a common rationale for summative testing is that it raises academic standards (Harlen and Deakin Crick, 2003; Kellaghan and Greaney, 2001), but this perspective may take too little account of the complexity of factors relating to motivation (Kellaghan et al., 1996). There is countervailing evidence suggesting that summative testing is detrimental to students who achieve lower grades, i.e., instead of general uplift, testing leads to greater polarization (Bourdieu, 1998; Paris et al., 1991; Pollard et al., 2000). Correlational studies between students' self-esteem and achievement find an increase in this relationship for cohorts who have taken more national examinations (Davies and Brember, 1998, 1999). Ethnographic studies find that classroom interactions changed in the wake of students receiving exam results, with the self-esteem of lower performers dropping (Leonard and Davey, 2001; Reay and William, 1999).

Prior research across the social sciences offers a range of potential theorizations for why the Key Stage 3 assessments might have a detrimental influence on those receiving lower grades. Among these, perhaps the most pervasive in educational research is attribution theory. Coming from psychology, attribution theory is a phenomenological system that focuses on an individual's judgment of the cause underlying a negative experience. The three key dimensions of an individual's reasoning are whether an event was due to factors that the individual believed to be 1) internal to themselves, 2) controllable, and 3) stable (Weiner, 2010). For example, if a student attributes a disappointing exam result to insufficient effort, this cause is internal, unstable, and controllable, so they will experience guilt and regret, which are positive motivators. In contrast, a student who receives identical negative feedback but attributes this to insufficient aptitude, a cause that is internal, uncontrollable, and stable, will experience hopelessness, shame, and humiliation, and so will see little point in studying for future assessments. In this theorization, summative examinations such as those at Key Stage 3 are detrimental to motivation and achievement (Graham, 1990).This is because they can consolidate what Dweck (2006) labels a "fixed mindset": in response to this form of feedback, children are likely to frame exam results as evidence of permanent ability traits, lessening their belief that ability can be developed through work and so deterring future effort (Dweck 1986; 2000).

Human capital theory, the most prevalent theory among education economists, proposes similar student reactions to exam feedback to those of attribution theory, albeit with an emphasis on different motivating forces. Stemming from the work of Becker (1967, 1980) and Schultz (1961), human capital theory frames education as an investment that individuals choose to undertake primarily in order to increase their future earnings. Different students will encounter different rates of return from education, and exam grades are likely to influence a given student's judgment of the rate of return to undertaking further education.

For example, students receiving the highest grades should, on average, be those most disposed to anticipate that future effort is likely to lead to greater attainment and associated career benefits.

Students who perform poorly on examinations relative to others can still benefit from further investment in education. In their decision-making, however, such students can lack sufficient information to accurately predict returns, or they may face psychic costs, i.e., stressors, that offset the benefit of investing in further education (Becker, 1976). While Becker did not expound greatly on the nature of psychic costs, claiming that "quantitative estimates of psychic gains are never directly available" (Becker, 1980, p. 198), attribution theory is just one conceptual approach that specifies why receiving low grades can stress students and deter future effort. These psychic costs may be exacerbated by discount rates: for many students, short-term costs may weigh more heavily in their reasoning than benefits that will only bear fruition in the longer term.

From the field of sociology, Bean's (1980, 1981, 1983) adaption of Price's (1977) worker turnover theory treats exam grades as an extrinsic reward. Price theorized that pay was a key determinant of an employee's decision to leave an employer, with increased pay serving to improve job satisfaction, organizational commitment, and limit search behavior, thus reducing the likelihood of an employee leaving. In Bean's adaption to undergraduate behavior, grades serve as a proxy for payment, providing students with an extrinsic reward for their effort that then incentivizes greater future effort. Those who receive higher grades are thus more likely to remain in education, as this reward increases their positive attitudes towards their educational experience and provides a resource that enables greater future opportunity.

Whereas attribution, human capital, and worker turnover theory all operate primarily with the student interpreting their exam results independently of others, intellectual competence is in part a function of social interactions (Aronson and Steele, 2005). For example, research from developmental psychology indicates that early adolescence is a period in which peer comparisons are particularly important to self-esteem. At this age, students typically form stronger assessments of their psychological capacities, whereas at earlier ages physical and behavioral attributes are more important; these self-assessments are based largely on comparisons with peers (Wigfield and Karpathian, 1991). This shifting focus is accompanied by a growing belief that academic ability is a stable trait that is less related to effort (Dweck and Leggett,1988; Nicholls, 1989), serving as a deterrent to put more effort into studying because students who fail under such circumstances are stigmatized more than those who fail without putting in so much effort (Covington, 1992). Consequently, students who receive below-average exam grades may focus on present failure at the expense of future possibilities, with the result that they avoid challenging tasks, revise their academic ambitions downwards, and do not pursue post-compulsory education (Anderman and Maehr, 1994; Black and William, 1998).

It is also feasible that grades influence the behavior not only of students but also, for example, of parents, teachers, and school leaders. Within economics, signaling theory and bounded rationality have both supplemented human capital theory by relaxing the strict assumptions in Becker's (1980) treatise that it is straightforward for individuals to make judgments about ability since they can both access and interpret all relevant information with ease. In contrast to human capital theory, the key conceptual underpinning of bounded rationality is that the cognitive cost of processing makes access to too much information counterproductive (Conlisk, 1996). As a consequence, students, parents, and teachers may respond to the cruder measure of ability that discrete grades represents rather than try to make more fine-grained analyses of the causes of the attainment that underlies these groupings (Papay et al., 2010).

Labor economists have theorized that grades provide an important source of information by which employers sort potential candidates in the absence of the opportunity to observe their competency in the specific job. Here, sorting comprises both signaling, wherein students use exam grades as a means to send positive signals about themselves to others, and screening, in which grades provide a sorting function by which employers use them as a proxy to assess ability (Bills, 2003). Signaling theory is most commonly attributed to Spence (1973, 1981), who developed this approach to explain employers' assessments of job candidates under conditions of uncertainty. In the case of Key Stage 3 exams, employer behavior serves as an analogy: future employers do not have access to a student's results in these exams. Nonetheless, it is tenable that key adults in a student's educational development could use signals as a tool to help them assess the student's educational development long before employers in the student's adulthood. Becker's (1980) original treatise on human capital acknowledges that parents are likely to play some advisory role in helping adolescents to reach an optimal equilibrium position in how much more education they should commit to, and, presuming that they do not hold perfect knowledge of ability, grades provide an important means to signal ability (Rosenbaum, 1986).

In sum, there are numerous theorizations of how grade labels might have important repercussions for students as they progress into post-compulsory education and early adulthood. Some of these theorizations may be more pertinent than others; for example, some focus on adult behavior and so might not translate so well to adolescents. Focusing first on the individual-level theorizations, this concern is pertinent to weighing the respective merits of attribution theory, worker-turnover theory, and human-capital theory. Of the three, attribution theory has the largest literary base indicating its applicability to the thought processes of 13–14 year-olds, especially given the body of developmental psychology focusing on the development of children's thought processes in relation to assessments at this specific age. Furthermore, Key Stage exams exhibit key characteristics of the type of assessments on which attribution theory focuses: as a result of their official and formal nature, students tend to perceive them as valid, reliable measures of ability, reducing the likelihood of students attributing lower grades to chance, preparation, or other causes that would provide hope of future improvement (Black and William, 2006; Pollard et al., 2000; Vispoel and Austin 1995).

In contrast, while Bean adapts worker turnover theory to a group typically of the same age as younger working adults, transferring Bean's theories to behavior during early adolescence is more tenuous: we currently lack research on whether children at ages 13–14 explicitly relate exams to an educational system for which they have an avowed commitment, or lack thereof. However, if one were to delimit the borrowing from worker turnover theory focus solely on the claim that exam grades serve as extrinsic rewards, this theory then becomes very similar to attribution theory: both focus on the student's response to the assessment at the individual level and its impact on effort towards future academic assessments. The two theories differ though in that attribution theory emphasizes the role of negative experiences in demotivating lower achievers, while worker turnover theory instead emphasizes the role of positive experiences in motivating higher achievers. Observational studies from the UK suggest that the former holds greater sway: researchers noted that the frequency and magnitude of negative, stressful experiences that students reported in relation to Key Stage exams far outweigh more positive experiences (Leonard and Davey, 2001; Pollard et al., 2000; Reay and William, 1999).

Human capital theory requires a further stage in the reasoning of students. Whereas attribution and the adapted worker-turnover theory both relate student experiences with exams to their subsequent judgment of similar experiences within the education system, human capital theory requires that students anticipate future employment opportunities and link their experiences and decision-making in school to these opportunities. There is evidence that students in the UK link school exams to adult life, and that students interpret low attainment in exams as a harbinger of failure in adult life. For example, on the basis of their observations in a London school, Reay and William (1999) depict a classroom climate where Key Stage examinations serve as a criterion by which students judge themselves and one another. When asked what a high Key Stage grade would say about a classmate, one student responds "that he's heading for a good job and a good life and it shows he's not gonna be living on the streets," whereas her own expectation of a low grade would say that "I might not have a good life in front of me and I might grow up and do something naughty" (pp. 346–347).

Yet, although students may link exam performance in school to employment in adulthood, there is not yet commensurate evidence that students are altering their behavior in response to summative assessments. While developmental psychology indicates that students can interpret their development in relation to others by the age (either 13 or 14) at which they sat the Key Stage 3 exams, there is not yet compelling qualitative evidence that they link these judgments to career opportunities at this stage and adjust their behavior according to anticipated future career opportunities. When challenged on the extent to which school children in the United States truly altered their behavior in education so as to align their attainment with feasible goals, Becker (1980) argued that parental input may influence students' confidence, aspirations, and knowledge about academic requirements. This consideration is all the more important in the UK where, as discussed earlier, the school system requires students to specialize by subject from early adolescence, meaning that students have to make potentially crucial decisions at ages when they are all the more likely to rely upon family and school teachers.

This contextual factor adds credibility to the theorizations – signaling theory and bounded rationality – that expand their focus beyond student decision-making occurring at the individual level, independently of others. This is particularly true with regards to teachers, as policy demands for accountability may pressure teachers into allocating more resources to some students than others. In recent years, for example, policymakers have treated the number of students achieving five GCSEs (including English and mathematics) at grade C as a key criterion in judging school performance (Vignoles, 2013), and this metric figures prominently in governmental school inspections (see, for example Taylor, 2012; Wilshaw, 2013). Teachers and school leaders may allocate resources accordingly, with less focus placed on those students projected to perform well above or below this cutoff (Ball, Maguire, and Braun, 2012; Gilborn and Youdell, 2000).

As mentioned, England's school system is characterized by early subject specialization. Within schools, ability streams may influence student achievement (Ireson, Hallam, and Hurley, 2005), and only some students may be able to take higher-tier GCSE examinations, in which they can attain the higher grades, or more advanced GCSE subjects, such as triple science. Given the high correlation between students' earlier and later performance in national assessments (McCall and Alcott, 2013), Key Stage 3 grade levels are likely to present an important tool for teacher and school leaders to help make such decisions. Teachers receive the exam results in terms of discrete grades rather than the continuous underlying scores, and may prefer this heuristic over some form of assessment marked on a continuous scale that they would have to generate in addition to the national examinations.

While these theories differ with regards to how they theorize the primary link between exam results and subsequent student behavior, they lead to similar hypotheses about the effects of grade labels. Attribution theory, human capital theory, and Bean's (1980, 1981) modified worker turnover theory argue, respectively, that results mediate students' affective sense of self, assessments of future earning potential, and sense of satisfaction with education. However, while differing in this respect, each draws the same conclusion about the effect of exam grades on students' future study behavior. While attribution theory focuses on the deterrent effect of lower grades, human capital theory and worker turnover theory focus on higher grades acting as an incentive. Signaling theory and bounded rationality both suggest that teachers and parents would use exam grades as a sorting mechanism to differentiate between students. Implicitly then, the hypothesis that would follow from each theory is that increasing polarization is likely to occur between those who attain higher grades and those who attain lower grades on earlier examinations.

In addition, none of these theorizations of grade label effects would lead, in and of themselves, to the anticipation of heterogeneous effects according to the subject of assessment or the outcome of interest. For example, none of the theories emphasize the relative importance of one subject matter over the others, no matter the mechanism by which they conceptualize a labeling effect occurring. Given that GCSEs, A-levels, and university enrollment serve as sequential points in the education system, there is no clear reason to think that the effect of grade labels may be more noticeable in relation to one of these outcomes than another, especially given that none of the theories anticipate either a dissipation or compounding of effects over time.

The theories discussed thus far provide a basis for understanding how grade labels during schooling may influence students' trajectories towards future markers of progress through the education system. What these theories do not offer, though, is a means to theorize why these labels may serve to exacerbate socio-economic disparities in education. For this, it is worth turning to one of the most prevalent theorizations of the formal school system in the UK research literature: Bourdieu's reconceptualization of habitus.

Bourdieu (1973; Bourdieu and Passeron, 1977) theorizes that behavior is socially situated, with the key influences in decision-making being the anticipated acceptance and stigmatization of others. For Bourdieu, it is crucial to recognize both that humans operate within social structures (Bourdieu 1986) and that human agency makes it possible for individuals to resist such structures (Bourdieu and Wacquant, 1992; Ovenden 2000). Bourdieu maintained this individual-collective balance by embedding habitus within a troika of analytical concepts, of which the other two components were field and doxa (Davey and Fuller, 2013).

Within this theorization, habitus is both the attitudes and tendencies that one develops as a result of interactions with others (Nash, 2005) as well as one's ensuing actions (Reay, 2004). Cultivated over time, an individual's habitus forms a range of dispositions and behaviors indicating what she considers to be appropriate, desirable and possible. An individual's decision-making is not fixed across the lifespan, but it is bounded at a given time according to prior opportunities and constraints. Field denotes the social space in which individuals interact, and it represents a more objective interpretation of a given situation than does habitus (Grenfell and James, 1998): "social reality exists, so to speak, twice, in things and in minds, in fields and in habitus" (Bourdieu and Wacquant, 1992, p. 127). Doxa refers to the beliefs that are taken for granted by individuals within any one field. This embeds the individual within the collective, as multiple habituses and their component notions of what is and is not desirable and feasible are entwined, shifting an individual's self-restricting response from 'not for the likes of me' to 'not for the likes of us' (Archer et al., 2007; Davey and Fuller, 2013).

The existence of hierarchy – and conformity to it – provides a crucial point of contrast from more individualistic theorizations such as human capital, since variation across individuals' habituses suggest variation in their beliefs about the benefits and drawbacks of higher education. Two students can approach the same enrollment decision with ex-ante similar propensities to benefit from choosing to enroll, but they may react differently to this decision according to their differing beliefs about what constitutes a desirable and/or feasible outcome. In addition, two students of differing habituses who undertake the same educational experience are also likely to benefit to differing extents since "the rate of return on educational capital is a function of the social and economic capital that can be used to exploit it" (Bourdieu and Passeron, 1979, p. 79). Bourdieu's body of work thus represents a critical analysis of the extent to which social mobility exists in contemporary Western societies. Education was of primary interest to him, but he was skeptical of its capacity to drive mobility, instead viewing it as an assertion of cultural power (Bourdieu, 1973). In one of their seminal works on education, Bourdieu and Passeron (1977) argue that "all pedagogic action is, objectively, symbolic violence insofar as it is the imposition of a cultural arbitrary by an arbitrary power" (p. 5).

While acknowledging that the aspects of identity that shape an individual's habitus are numerous, Bourdieu and co-authors' (Bourdieu, 1998, 2005; Bourdieu and Passeron, 1977, Bourdieu and Wacquant, 1992) work on educational structures emphasized social class. Considering elite universities in France, Bourdieu and Wacquant (1992) depict university as a field, wherein undergraduates attempting to access and navigate this space are taking part in a competitive game for which there are rules (doxa). Students' knowledge of these rules is unequal, as are their dispositions (habituses) to take on study, with the more disadvantaged belonging to lower social classes.

Disadvantage occurs, according to this line of argument, because the educational system legitimizes the cultural practices and preferences more typical of the upper and middle classes over those more typical of the lower classes. In contrast to these students, those from the working classes face a greater onus to "engage in rational computation in order to reach the goals that best suit their interests" (Bourdieu, 1990, p. 108). These students' study experiences are also vulnerable to stigmatization should they behave in line with their own habituses because "taste classifies, and it classifies the classifier" (Bourdieu, 1986, p. 56), and in this field the prevailing doxa holds working-class tastes to be vulgar (Bourdieu, 1986; Bourdieu and Passeron, 1979).

Studies of pre-elementary and elementary education in England note that both mechanisms are in operation even by these stages (see, for example, Reay, 1995, 1998; Vincent, Braun, and Ball, 2008; Walker and Clark, 2010). From the earliest years of schooling, teachers label students as intelligent, average, or slow, and stream them accordingly; these designations are often predicted by social class and have a lasting impact on students' academic confidence (Steedman, 1988; Thomas, Bland, and Duckworth, 2012). In interviews, Reay (1995) found that middle-class parents were quick to raise concerns about group-reading activities for fear that these were aiding other children's progress at the expense of their own.

Such class disparities widen further when differential parental knowledge and confidence lead to asymmetrical competition to gain entry for their children to specific middle schools and, by extension, GCSE courses (Ball, Bowe, and Gewirtz, 1996). As noted already, England's school system makes the consequences of subject choice and performance in middle school for later opportunities considerable. As a result, early manifestations of the constraining factors identified within habitus are compounded by the time students reach the later years of schooling, well before their course choices and exam performance have easily visible implications for university applications (Vincent, 2001). Hence, while many studies focus on students around the final years of the compulsory school system, this associated research indicates that the institutionalized constraints that students face are established from the earliest years of schooling.

Within this theorization of formal education, exams serve a key factor, as across schooling they serve to consecrate the advantages of more privileged students, serving as both cause and effect of greater separation (Bourdieu, 1998, p. 104). Researchers of education in England have used Bourdieu's theories to argue not only that better aggregate performance on exams serves to separate social classes, but that differences in parental behavior according to social class further exacerbate disparities. This is because parents from lower social classes have less confidence in contesting practices within a school, such as placing their child in a lower ability stream or demanding more extensive feedback on classwork (Cochrane, 2007, 2011; Giddens, 1991; Pugsley, 1998).

Consequently, working-class families may be especially prone to accepting summative judgments about their children's progress and ability levels. National exams are a key case in point, as parents from lower social classes are less likely to challenge their validity or encourage their children to accept them as incontestable judgments of ability. Following from attribution theory, learners who attribute success to effort, and who perceive ability to be changeable and controllable, are likely to deal with failure constructively and so persevere with future learning tasks (Schunk, 1991). Hence, differential responses to exam results by class are important because it is plausible that this may serve to exacerbate pre-existing attainment gaps.

A major strength of the habitus-based research literature is that it provides clear empirical evidence for its theorizations of working-class students' perceptions. These are given a central role in most research papers through sustained, uninterrupted quotations that make clear that students vary considerably in the role they perceive education to play in their own lives, as well as the factors that impinge on their decisions to pursue non-compulsory education. Typically, the analyses that accompany interview and focus group data are cogent and nuanced, synthesizing students' views in a manner that establishes themes but is also careful to recognize discordant viewpoints. Such research approaches strengthen this literature's ability to "make sense of the world from the perspective of participants" (Eisenhart, 1988, p. 103).

Researchers employing habitus have also presented persuasive evidence to support their claim that English schools and universities are rooted in a normative education system that values upper- and middle-class traits. Although they only represent a small portion of the habitus-based literature, Ball's (2012) and Reay's (2001) documentary analyses provide compelling evidence that policymakers have pathologized working-class students' language and behavior throughout the history of compulsory education in England. Further, research focusing on parental interventions with schooling belies Becker's (1980) assumption from human capital that parents might compensate for shortcomings in students' capacities to navigate the education system, suggesting instead that the enthusiasm and capacity to make favorable interventions is typically concentrated among those parents whose children face the fewest structural barriers.

Aims of the present work

While I am not able to test the respective merits of these theorizations, I make an important contribution to the research literature. This is because my empirical approach enables an inferential verification of the presence, or lack thereof, of a labeling effect from Key Stage 3 exams, and so can initiate the type of generalizable analyses on this phenomena that are currently lacking for the English context. To date, those studies that are more critical of summative assessment in England have relied predominantly on qualitative research methods to discern the impact of examinations on student behavior (see, for example, Leonard and Davey, 2001; Reay and William, 1999). The major exceptions, by Davies and Brember (1998, 1999), provide quantitative representations of the impact of Key Stage 1 and Key Stage 2 assessments on self-esteem, but without providing the type of valid counterfactual scenarios necessary to make inferential claims.

Quasi-experimental methods have not yet been applied to estimate the impact, if any, of grades received during these assessments on students' progress to post-compulsory education. Such methods have already identified effects of performance feedback in different educational settings. In Swedish elementary schools, Sjögren (2010) found mixed impacts of the use of grading according to gender and parental education levels. Positive effects of feedback on subsequent performance have been found for high school students in Spain (Azmat and Iriberri, 2010) and the United States (Papay, Murnane, and Willet, 2010). In the UK, Sartarelli (2011) has explored the impact of Key Stage scores on student behavior, finding no impact on most outcomes with the exception of bullying.

I aim to transfer such quasi-experimental methods to the study of the impact of summative assessments in England on students' subsequent educational outcomes. Specifically, I attempt to identify the impact of receiving a low or high exam grade, independent of achievement. In addition, I investigate whether effects differ for students of different socio-economic strata. Using a regression discontinuity design, I compare the impact of receiving the average exam level (versus below average) on the following outcomes:

1. GCSE performance at age 16.
2. Enrollment in A-levels by age 19.
3. Enrollment in a university degree course by age 20.

Data and summary statistics

I use data from the Longitudinal Study of Young People in England (LSYPE). The Longitudinal Study of Young People in England (LSYPE) used a two stage probability proportional to size sampling procedure, with the primary sampling units being schools. LSYPE respondents were born between September 1, 1989 and August 31, 1990. Interviews were conducted annually between the spring of 2004, when the youths were in eighth grade, and 2010, providing seven waves of data. The survey's first wave sampled 15,770 youths, but attrition in later waves means that the sample sizes are reduced to 14,947, 11,186, and 8,233 according to whether the outcome of interest related to GCSEs, A-levels, or university enrollment, respectively.

This attrition is likely to have lessened the extent to which LSYPE is representative of the national cohort that it was intended to represent. Whereas the proportion of English domiciled young people aged 17–19 enrolled in university was 33% in 2008/09 and 34% in 2009/2010 (Department for Business Innovation and Skills, 2012), enrollment among the LSYPE cohort averaged 43% across these years, even when using survey weights, suggesting attrition according to characteristics for which LSYPE's sample weights cannot control (Anders, 2012b). Consequently, although my findings are likely to be indicative of national trends, subsequent model estimates should not be presumed to be nationally representative.

Another consequence of survey attrition is that the samples that I use for my three outcomes of interest differ slightly from one another. As presented in Table 2.1, it appears that respondents from higher socio-economic backgrounds and higher Key-Stage-3 attainment were more likely to continue participating in the later LSYPE waves, during which the questions on A-levels and university enrollment were asked. This adds a further complicating factor because apparent changes in the impact of encouragement over the educational stages (for example, evidence of a dissipating impact or a compounding impact) may be partially attributable to differences between the sub-samples analyzed at each stage.

Table 2.1: Description of the data

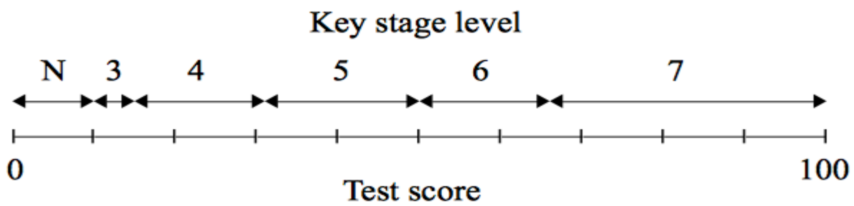
	When the following outcome is observed:		
	GCSE	*A-levels*	*University*
Observations	14,947	11,186	8,233
NSSEC Level (%)			
Lower	27	27	25
Intermediate	30	28	28
Higher	43	45	47
Mean Key Stage 3 Score (Standard Deviation)			
English	33.1(6.3)	33.7 (6.1)	34.4 (6.0)
Mathematics	35.5 (8.1)	36.3 (8.0)	37.2 (7.8)

Source: Longitudinal Study of Young People in England.

The key predictor variables in my analyses are respondents' Key Stage 3 test results in English and mathematics. I do not analyze respondents' Key Stage 3 test results in science for two reasons. First, unlike English and mathematics, it is not possible to assess students' subsequent progress in science because students do not follow a core curriculum with a common exam in this subject. Second, the predominance of English and math performance in school league tables could mean that science results factor less into teachers' planning for student interventions and ability streaming. Consequently, any impact of test results in science on the more general outcomes of enrollment in A-levels and university study may differ from the respective impacts of mathematics and English. Hence, in subsequent analyses I focus on the English and mathematics exams.

LSYPE is linked to the UK Government's National Pupil Database, which holds administrative data on the test scores that students obtained in their Key Stage 3 subject tests. In each subject test, students could receive a grade of 3, 4, 5, 6, or 7 (as well as 8 for mathematics only) depending on the marks that they obtained, from within a range 0–100 for English and 0–150 for mathematics. Figure 2.1 provides an example of how continuous score mapped onto grade levels in one of these assessments.

Figure 2.1: An example of how test score maps to key stage level



Note: Example comes from 2003 English assessment, which LSYPE respondents took. Fewer than one percent of respondents score below Level 3.

Data on the first outcome, passing GCSE examinations, is obtained through the LSYPE's link to the National Pupil Database. This outcome is defined in two respects: subject-specific and general performance. Subject-specific performance compares students' Key Stage math grades to whether they receive a grade of C or higher in GCSE math, and the equivalent for Key Stage English to GCSE English. General performance is assessed through the capped GCSE score, which comprises a continuous point score based on the eight best exam results for a given student. The capped GCSE score is a common measure in education policy in England (Vignoles, 2013); it is based on students' performance in just eight, rather than all, exams to account for the tendency for students at fee-paying schools to take more GCSE examinations than do students in the state-provided system. The second and third outcomes, enrollment in A-levels and enrollment in university, rely on students' self-reports during the survey's fourth and subsequent waves.

To analyze heterogeneous effects according to socio-economic status, I use the three-class grouping – higher, intermediate, and lower – of the National Statistics Socio-economic Classification (henceforth NSSEC) offered by the UK government's Office for National Statistics (2014), with the caveat that the long-term unemployed are also included in the "lower" grouping due to insufficient sample sizes. Students for whom NSSEC data is missing are included in models of general effects but are not included in models of differential impacts according to socio-economic status.

Table 2.2: Test scores and outcomes for full sample and according to NSSEC level

		By NSSEC level:		
	All	*Lower*	*Intermediate*	*Higher*
English Level 6 or above	0.34	0.26	0.35	0.55
Maths Level 6 or above	0.53	0.44	0.59	0.74
Mean capped GCSE point score	303	287	317	353
Attain C or higher in GCSE English	0.59	0.51	0.65	0.81
Attain C or higher in GCSE mathematics	0.54	0.46	0.60	0.76
Enroll in A-levels	0.56	0.51	0.47	0.73
Enroll in University degree	0.49	0.41	0.46	0.62

Note: Results for the full sample ("All") include students with missing NSSEC data. Source: Longitudinal Study of Young People in England.

Table 2.2 provides descriptive statistics on the label thresholds and outcome measures, first for all students and then according to socio-economic status. There is a clear socio-economic disparity for each of the label thresholds and outcome measures, with higher NSSEC students outperforming intermediate NSSEC students, who in turn outperform lower

NSSEC students. Table 2.3 presents average outcomes for all sampled students according to whether they scored above or below the Level 6 threshold in each subject. For each subject, those scoring above the threshold realize each of the outcomes at higher rates than those who scored below the threshold, and always by at least a factor of two. All differences are statistically significant at any conventional level.

Table 2.3: Sample averages for outcomes by Level 6 threshold

	Below Level 6	*Level 6 and above*	*p-Value*
English			
Capped GCSE point score	265	376	<0.001
Attain C or higher in GCSE English	0.41	0.96	<0.001
Enroll in A-levels	0.40	0.84	<0.001
Enroll in university degree	0.32	0.73	<0.001
Mathematics			
Capped GCSE point score	239	359	<0.001
Attain C or higher in GCSE mathematics	0.15	0.89	<0.001
Enroll in A-levels	0.32	0.75	<0.001
Enroll in university degree	0.25	0.64	<0.001

Source: Longitudinal Study of Young People in England.

The empirical strategy

The use of regression-discontinuity designs (RDD) in social science research dates back half a century (see Thistlethwaite and Campbell, 1960; Campbell and Stanley, 1963). Although the approach received limited attention in subsequent decades (Cook, 2008), it has become increasingly popular in recent years (Lee and Lemieux, 2010; McCall and Bielby, 2012). In the field of education, studies have used the RDD approach to investigate a range of influences such as financial aid (Van der Klaauw, 2002), scholarships (DesJardins and McCall, 2014), remedial education programs (Jacob and Lefgren, 2004), and pre-school interventions (Ludwig and Miller, 2007).

RDD's popularity is likely due to its intuitive appeal as the most valid estimation strategy under specific allocation mechanisms, absent random allocation (Cook, 2008; Lee and Lemieux, 2010). The specific mechanisms are those in which treatment is assigned via a cutoff point on an observed continuous variable. So long as individuals are unable to manipulate whether they fall on one side of the cutoff point or the other, the variation in treatment around this point is essentially random (Lee and Lemieux, 2010). The present study is well suited to an RDD approach because the assessment feedback device – discrete exam grades – varies at distinct cutoff points on the underlying continuous scores that students achieve in each test. Students are unable to manipulate which side of the grade cutoff their scores fall because grade boundaries are not defined until after each year's test. In addition, students, parents, and teachers are provided with discrete grades after the test but not the exact value on the underlying continuous mark scale. This is important because knowledge of the continuous mark might mitigate the response of any of these groups in relation to the grade; for example, teachers might see students who scored just below the threshold as more comparable to those scoring just above it than to those scoring many points below the threshold.

In order to get a higher grade on the Key Stage test, students must meet or exceed a given test score. Thus, a given student's receipt of a higher grade (G) depends on her test score (S). More specifically, the grade depends on their score in relation to a grade cutoff (C), whereby those students for whom $S \geq C$ receive grade G. Taking receipt of a higher grade as a treatment, for student i its relation to a future outcome (Y) can be denoted as

$$Y_i = \begin{cases} Y_i^1 \, if \; G_i = 1 \\ Y_i^0 \, if \; G_i = 0 \end{cases} \qquad (1)$$

or

$$Y_i = Y_i^0 + G_i\left(Y_i^1 - Y_i^0\right) \qquad (2)$$

However, the inferential challenge is that it is not possible to observe both Y_i^1 and Y_i^0 because student *i* cannot both receive and not receive the treatment *G*. Random allocation overcomes this challenge since, taking α as a vector of all variables that could influence Y_i prior to receipt of the treatment, in the equation

$$Y_i = \alpha + \tau G_i + \varepsilon_i \qquad (4)$$

G_i and the error term ε_i are independent. Thus, the estimate of the treatment effect τ is obtained by subtracting the average of *Y* for all untreated students from the average of *Y* for all treated students.

The properties of the Key Stage grade boundaries make it possible to partially emulate this ideal randomized scenario through a "sharp" RDD design. In a "sharp" design, the cutoff variable perfectly predicts allocation of the treatment (Imbens and Lemieux, 2008), so that

$$\begin{gathered} G_i = 1\left\{S_i \geq c\right\} \\ G_i = 0\left\{S_i < c\right\} \end{gathered} \qquad (5)$$

The "sharp" design is possible because all students whose scores fall below the boundary receive the lower discrete grade, while all those whose scores fall at or above the boundary receive the higher grade.

Returning to equation 4, accepting that the expected value of the error term ε_i varies with S_i, but assuming that it does so as a continuous function of *S*, Hahn et al. (2001) have shown that

$$\tau = \lim_{x \downarrow c} E\left[Y_i | S_i = s\right] - \lim_{x \uparrow c} E\left[Y_i | S_i = s\right] \qquad (6)$$

provides the treatment effect of *G*. Note though that this only uses observations from the arbitrarily selected range $[c - x, c + x]$ around the cutoff point (McCall and Bielby, 2012), whereas all observations may be used under randomized allocation. Thus, the RDD provides an estimate of the local average treatment effect, but it does not allow for extrapolations to observations further from the threshold (DiNardo and Lee, 2011; Shadish et al., 2002).

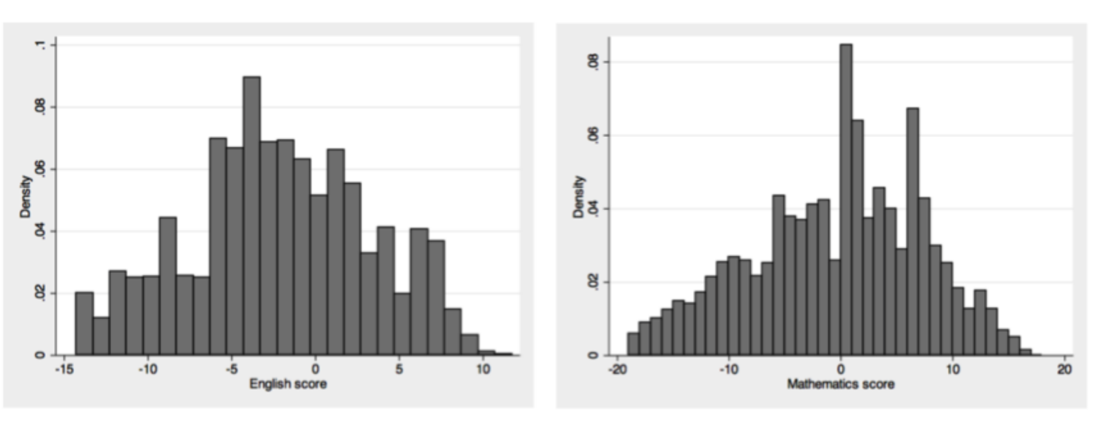
In order to limit bias, kernel-based polynomials provide a popular means to predict τ under the RDD design (Hahn et al., 2001; Lee and Lemieux, 2010). Consequently, researchers face choices regarding three factors that have implications for a given model's validity: the kernel, the bandwidth, and the use of polynomial terms in the regression model (Mealli and Rampichini, 2012). Among these, the choice of kernel is of lesser importance because model estimates are not so sensitive to different kernels as they are to different bandwidths and polynomial terms (Imbens and Lemieux, 2008; McCall and Bielby, 2012). I follow McCall and Bielby (2012), who use a Gaussian kernel, although all models were re-estimated with alternative kernels in order to check whether estimates are highly sensitive to this choice. Estimates using these alternative kernels were not substantively different to those using the Gaussian kernel.

As Mealli and Rampichini (2012) note, choice of bandwidth offers a tradeoff between precision and bias. In order to allow for bandwidths larger than those most commonly used in prior research (Calonico et al., 2013a), I use the formulation offered by Calonico et al. (2013b) to select data-driven bandwidths for each model in order to minimize the mean square error. In their formulation, Calonico et al. (2013b) base their confidence intervals on a bias-corrected discontinuity estimator to account for the impact of large bandwidth choices, but depart from the prior literature by using an alternative formulation of standard errors in order to account for the greater variability in the calculation of a given t-statistic caused by estimated bias-correction.

In addition, it is important that students cannot manipulate their positioning around the threshold. In recent years, McCrary's (2008) formal test of the distribution of the running variable has become perhaps the most common means for examining potential manipulations around the cutoff. As Figure 2.2 indicates, the running variables – i.e., the test score in each Key Stage 3 subject assessment – would not pass McCrary's test.

However, this can be explained by contextual factors. First, the examining authority does not define grade boundaries until after the test, and the central examination body already holds examination scores by the time that these boundaries are announced. Second, one key practice that examiners use for the Key Stage 3 assessments is "borderlining," which is the practice of re-marking those papers that fall three marks below the discrete Level boundaries (QCA, 2004; Quinlan and Scharaschkin, 1999). Since this is done for those papers falling just below the boundary but not for those just above, and given the variation in examiners' scores (House of Commons Children, Schools and Families Committee, 2008), it is natural that the frequency of scores will drop just before the boundary and rise just after. While "borderlining" proved controversial and was eventually stopped (House of Commons Children, Schools and Families Committee, 2008; National Audit Office, 2008), it was still in use in 2004, the year in which LSYPE respondents sat these assessments.

Figure 2.2: Histograms of test scores by subject



Source: Longitudinal Study of Young People in England. On the x axis, scores are standardized so that –6, 0, and 6 represent the minimum scores for Levels 5, 6, and 7, respectively.

In addition, I conduct an alternative sensitivity check for potential manipulation, presented in Appendix 2.A, in which I regress whether students fall just above or just below each threshold on the key background characteristics of NSSEC grouping and Key Stage 2 performance. Of the 18 models that I estimate in this sensitivity check, just one has a significant regressor, providing limited evidence of any systematic manipulation of positioning around the Level 6 threshold in any of the exams.

Model estimates

For all students

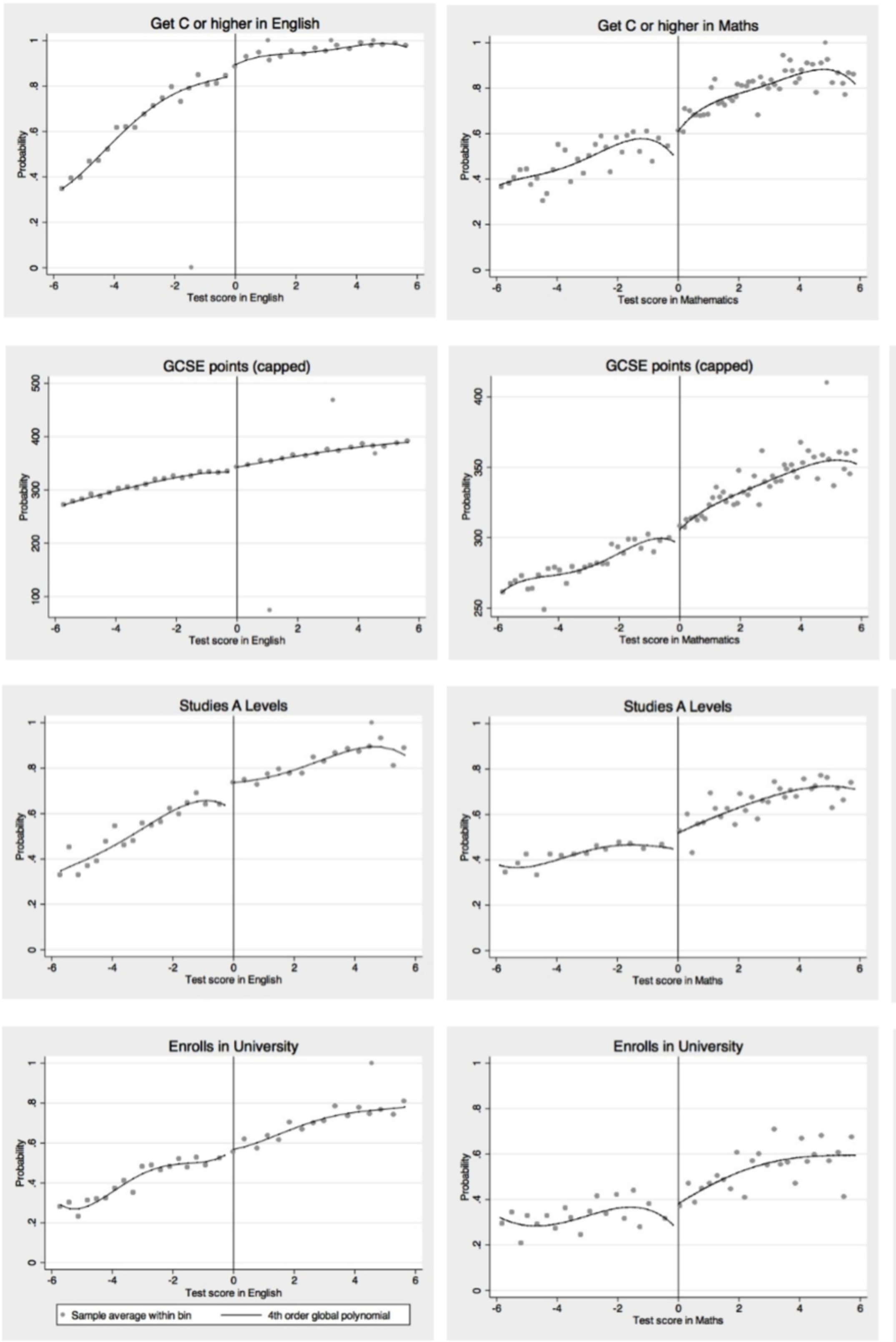
Table 2.4: Estimated impact of Level 6 label

	Subject	
Outcome	(1) *English*	(2) *Mathematics*
Capped GCSE point score	3.741*	4.788
	(2.738)	(3.852)
Attain C or higher in GCSE subject	0.028	0.017
	(0.025)	(0.039)
Enroll in A-levels	0.094**	0.037
	(0.041)	(0.026)
Enroll in university degree	0.039	0.061
	(0.033)	(0.038)

Note: Coefficients for the first row represent the expected change in GCSE points score associated with the Level 6 label. Coefficients for the second, third, and fourth rows represent the percentage-point change in probability, associated with the Level 6 label, of a student achieving the given outcome. Results include students with missing NSSEC data. Standard errors are reported in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Source: Longitudinal Study of Young People in England.

Model estimates for the impact of the Level 6 label are presented in Table 2.4. An accompanying visual depiction is presented in Figure 2.3, where scores are standardized so that the Level 6 threshold is represented by zero. As noted in Section 2, estimates are based on non-parametric models that use Calonico et al.'s (2013b) estimation procedure in order to

Figure 2.3: Regression discontinuity estimates: all students



Source: Longitudinal Study of Young People in England. On the x axis, scores are standardized so that –6, 0, and 6 represent the minimum scores for Levels 5, 6, and 7, respectively.

minimize standard errors. Consequently, bandwidth size varies across models; information about the bandwidth and sample size used for each model is presented in Appendix 2.B.

These findings provide moderate evidence of a labeling effect from the English examination, but essentially no evidence of the mathematics label having an impact. The Level 6 label in the English assessment is associated with a 3.7 point increase in capped GCSE point score. It is worth noting though that this is only significant at the 0.1 level, and that this coefficient equates to less than a one-grade increase (e.g., from a C to a B) on a single GCSE exam, which would be represented by a six-point increase. In addition, the Level 6 label in English is associated with a 9.4 percentage-point increase in enrollment in A-levels. This may seem a large increase given that 67% of the full sample enroll in A-Levels, although it is worth noting that the lower bound for this 9.4 percentage-point increase is just 1.4 percentage points. The Level 6 label in English does not appear to have a statistically significant impact on a given student's likelihood of attaining a C or higher in GCSE English. In the mathematics examination, the Level 6 label is not significant in relation to any of the outcomes. For university enrollment, although the label from mathematics is close, neither it nor the English label is significant at the 10% level.

By socio-economic status

Estimates of the Level 6 label according to socio-economic status are presented in Table 2.5. These indicate that labeling effects differ according to socio-economic status: while there is limited evidence of an impact for students from the higher and mid-NSSEC groupings, substantial labeling effects are visible for the lower NSSEC group.

For high-NSSEC students, three of the 11 models provide significant non-zero estimates, albeit each of these is at the 10% level. The English label is linked to improved performance in GCSE English, with a 4.6 percentage point increase in the likelihood of attaining a C or

Table 2.5: Estimated impact of Level 6 label, by NSSEC level

	Subject	
	(1) *English*	(2) *Mathematics*
NSSEC: lower		
Capped GCSE point score	20.11**	28.21***
	(8.40)	(7.57)
Attain C or higher in GCSE subject	0.207***	0.057
	(0.062)	(0.066)
Enroll in A-levels	0.110**	0.082
	(0.054)	(0.056)
Enroll in university degree	0.205**	0.008
	(0.092)	(0.066)
NSSEC: intermediate		
Capped GCSE point score	-1.189	5.438
	(8.47)	(5.862)
Attain C or higher in GCSE subject	-0.020	0.055
	(0.035)	(0.071)
Enroll in A-levels	0.111	0.061
	(0.075)	(0.065)
Enroll in university degree	0.015	0.144**
	(0.083)	(0.071)
NSSEC: higher		
Capped GCSE point score	-1.062	9.494
	(4.200)	(5.160)
Attain C or higher in GCSE subject	0.046*	0.108**
	(0.037)	(0.054)
Enroll in A-levels	-0.068	0.094*
	(0.045)	(0.055)
Enroll in university degree	-0.030	0.039
	(0.052)	(0.048)

Note: Standard errors are reported in parentheses. * $p<0.1$; ** $p<0.05$; *** $p<0.01$.
Source: Longitudinal Study of Young People in England.

higher at GCSE. The Level 6 label in mathematics is linked to a 10.8 percentage point increase in the likelihood of getting a C or above in GCSE mathematics and a 9.4 percentage point increase in the likelihood of enrollment in A-levels. For mid-NSSEC students, there is evidence of a labeling effect for just a single outcome: the mathematics exam is positively associated with an increase of 14 percentage points in the likelihood of university enrollment by age 20, which is significant at the five percent level.

In contrast to the other two groupings, there is considerable evidence of a labeling impact on lower-NSSEC students. The English examination appears to have a positive impact on all four outcomes, with coefficients of 20 points for capped GCSE point score (equivalent to a boost of more than three grade levels), 21 percentage points for attaining at least a C in GCSE English, 11 percentage points for A-level enrollment, and 21 percentage points for university enrollment. The mathematics exam has a positive impact on capped GCSE point score of 22 percentage points, but no discernible impact on pass rates in GCSE mathematics or enrollment in either A-levels or a university degree.

Discussion and future directions for research

At Key Stage 3, it appears that the Level 6 grade label had a polarizing effect on otherwise similar students at the aggregate level. This finding supports the notion, as hypothesized by a range of theoretical backgrounds, that feedback from summative testing influences student behavior. It corroborates analyses conducted in non-English contexts (e.g., Papay et al., 2010) or on non-academic outcomes (e.g., Sartarelli, 2011), and substantiates the claims made by a body of literature (e.g., Black and William, 1998, 2006; Harlen and Deakin Crick, 2003; Reay and William, 1999) on the English context that has provided compelling phenomenological evidence but lacks the inferential studies necessary to establish a plausible counterfactual.

Moreover, my findings add weight to the claim that students' pathways through formal education depend, at least partially, on social class, thus expanding the evidentiary base for researchers using Bourdieu's theorization of habitus. Specifically, labeling effects appear to be have been greatest for students from lower socio-economic classes. However, my analytical approach does not make it possible to evaluate the respective merits of the mechanisms that these researchers have offered, such as whether such tests serve to confirm teacher prejudice or working-class families are especially prone to accepting summative judgments about their children's progress and ability levels.

It is worth noting though that my estimates of the size of Key Stage 3 labeling effects, especially when considering the lower bounds of their associated confidence intervals, cannot explain the majority of the link between school achievement and university enrollment that has been identified in the literature (e.g., Chowdry et al., 2013; Anders, 2012). Nonetheless, it seems that grade labels play some part both in the achievement-enrollment relationship and socio-economic disparities in enrollment. Since the purpose of testing at Key Stage 3 was not even to award qualifications to students for their performance (and neither, implicitly, to punish those with relatively poor performance), these findings indicate that it may be worth at least reassessing current feedback procedures.

Potential adjustments could lie on a spectrum at which the more extreme end would be dropping assessments entirely, should they provide little information about school quality. A less extreme change would be to not share results with students, or, should it be the case that teachers or school leaders are the source of labeling effects from the assessments, these tests might be anonymized. A more moderate approach would be to provide feedback scores on a continuous scale. Even if this were to complicate the interpretation of scores for children or parents, this would likely be compensated for by the benefits of moving away from the present mode of over-simplification. Perhaps most moderate of all, greater care should be taken over students' and parents' interpretation of the grades so that lower performers do not see their grade levels as definitive judgments of academic ability, but rather as something that is malleable with greater effort in future.

Currently, Key Stage 2 assessments in English and mathematics follow a similar format to the assessments taken by the LSYPE cohort. CATs, the common replacement for Key Stage 3 assessments, differ from their predecessor more drastically. As noted, even though the government does not require schools to use CATs, a majority of schools do so. Given this absence of central enforcement, the implementation of CATs varies substantially across schools, and our knowledge of this variation is limited. For instance, CATs provide schools with each student's underlying continuous score and band groupings, and schools are also able to define their own band groupings. Some schools use these tests to group children into ability bands, set achievement targets for both students and teachers in the subsequent tenth-grade exams, and share results with students' parents, either in the format of the continuous score, achievement bands, or a mixture of the two.

This uneven implementation of CATs constrains future inferential statistical research. Even if data were to become more widely available, endogeneity would present a problem: there are likely to be unobservable factors influencing why certain schools implement CATs in a particular way that will also be linked to future educational outcomes. However, applications of the regression discontinuity design could still be used for Key Stage 2 assessments to identify whether whether labeling effects are visible from this earlier age too, and if so, whether their impact may be greater or lesser than those occurring in adolescence.

In addition, the UK government has more secure datasets than the LSYPE that would enable more detailed analyses of any relationship between Key Stage 2 assessments and student outcomes in A-levels and higher education. For example, it would be possible to examine performance in A-levels and the type of subjects studied. Similarly, for higher education, it would be informative to analyze differences in institutional quality and subject major. Among other benefits, such research would enable more nuanced analyses of the relationship between Key-Stage grade labels in a particular subject, e.g. mathematics, and subject-specific outcomes, such as performance in GCSE mathematics, or the study of STEM subjects at university.

Qualitative research could evaluate the competing and interacting nature of the mechanisms underlying summative testing in a manner that is not possible with the type of inferential analysis I have used. Important work has already been produced on the impact of Key Stage assessments on student esteem (Leonard and Davey, 2001; Reay and William, 1999). Alterations in behavior may also occur for students, peers, parents, teachers, or some other interested party. Further studies that use qualitative methods, such as ethnographic observations or interviews, could shed greater light on who else besides students is responding to grade labels, and the mechanisms by which their responses have an impact.

In contrast to the more uniform practices of the Key Stage 2 exams, the uneven implementation of CATs could be a boon for such research. The variety in conditions could enable researchers to tease out the impact of small differences in how schools choose to share and act upon test results. For example, more phenomenological analyses could investigate whether students respond differently to grade labels when they have direct consequences, such as defining ability streams, or the more general form of judgment that stems from being placed on an ability scale. Similarly, it would be useful to know how parents might interpret CATs scores, and whether they are reliant on discrete boundaries when presented alongside the underlying continuous scores. Thus, the very source of difficulty that CATs present for inferential research could prove fruitful for qualitative research, providing us with more nuanced insights into the role of summative assessment in education.

References

Aaronson, J., & Steele, C. (2005). Stereotypes and the fragility of academic competence, motivation, and self-concept. In A.J. Elliot and C.S. Dweck, (Eds.), *Handbook of competence and motivation* (pp. 436-456). New York, NY: Guilford Press.

Anderman, E. M., & Maehr, M. L. (1994). Motivation and schooling in the middle grades. *Review of educational Research*, *64*(2), 287-309.

Anders, J. (2012a). The Link between Household Income, University Applications and University Attendance. *Fiscal Studies*, *33*(2), 185–210.

Anders, J. (2012b). *Using the Longitudinal Study of Young People in England for research into higher education access.* Department of Quantitative Social Science-Institute of Education report number 12–13, University of London. Retrieved from http://ideas.repec.org/p/qss/dqsswp/1213.html

Archer, L., Hollingworth, S., & Halsall, A. (2007). 'University's not for me—I'm a Nike person': urban, working-class young people's negotiations of style, identity and educational engagement. *Sociology*, *41*(2), 219-237.

Archer, L., Hutchings, M., & Ross, A. (2003). *Higher education and social class: issues of exclusion and inclusion*. London: Routledge Falmer.

Azmat, G., & 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), 435-452.

Ball, S. J. (2012). *Politics and policy making in education: Explorations in policy sociology*. London: Routledge Falmer.

Ball, S. J., Bowe, R., & Gewirtz, S. (1996). School choice, social class and distinction: the realization of social advantage in education. *Journal of Education Policy*, *11*(1), 89-112.

Ball, S. J., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary schools*. Abingdon, UK: Routledge.

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in higher education*, *12*(2), 155-187.

Bean, J. P. (1981). *The synthesis of a theoretical model of student attrition*. Paper presented at the American Education Research Association's annual meeting, Los Angeles, CA.

Bean, J. P. (1983). The Application of a Model of Turnover in Work Organizations to the Student Attrition Process. *Review of Higher Education*,*6*(2), 129-48.

Becker, G. S. (1967). *Human capital and the personal distribution of income*. Ann Arbor, MI: Institute of Public Administration.

Becker, G. S. (1976). *The economic approach to human behavior*. Chicago, IL: University of Chicago press.

Becker, G. S. (1980). *Human capital: A theoretical and empirical analysis, with special reference to education* (2nd ed.). Chicago, IL: The University of Chicago Press.

Bermingham, J. (2014). *Statistical first release 197*. Cheltenham, UK: Higher Education Statistics Agency.

Bills, D. B. (2003). Credentials, signals, and screens: Explaining the relationship between schooling and job assignment. *Review of educational research*, *73*(4), 441-449.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through

classroom assessment. *Phi Delta Kappa*, 1–13
Black, P., & Wiliam, D. (2006). The reliability of assessments. In J. Gardner (Ed.) *Assessment and learning* (pp. 119-131). London: SAGE.
Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, *21*(1), 5-31.
Blanden, J., & Machin, S. (2004). Educational inequality and the expansion of UK higher education. *Scottish Journal of Political Economy*, *51*(2), 230–249.
Bourdieu, P. (1973). Cultural reproduction and social reproduction. In R. Brown (Ed.), *Knowledge, education, and cultural change: Papers in the sociology of education* (pp.71-112). London: Tavistock.
Bourdieu, P. (1986). *Distinction: A social critique of the judgement of taste*. London: Routledge & Kegan Paul.
Bourdieu, P. (1990). *The logic of practice*. Cambridge: Polity Press.
Bourdieu, P. (1998). *The state nobility: Elite schools in the field of power*. Stanford, CA: Stanford University Press.
Bourdieu, P. (2005). *The social structures of the economy*. Cambridge: Polity Press.
Bourdieu, P., & Passeron, J. C. (1977). *Reproduction in education, society and culture*. London: Sage.
Bourdieu, P., & Passeron, J. C. (1979). *The inheritors: French students and their relation to culture*. Chicago: University of Chicago Press.
Bourdieu, P., & Wacquant, L. (1992). *An invitation to reflexive sociology*. Chicago, IL: University of Chicago Press.
Broadfoot, P. (1999). *Empowerment or performativity? English assessment policy in the late twentieth century*. Paper presented at the British Educational Research Association Annual Conference, University of Sussex, 2–5 September.
Callender, C. (2006). The impact of tuition fees and financial assistance. In D. B. Johnstone, M. J. Rosa, H. Vossensteyn, & P. N. Teixeira, (Eds.), *Cost-sharing and accessibility in higher education: A fairer deal?* (pp. 105–132). Dordrecht: Springer.
Calonico, S., Cattaneo, M. D., & Titiunik, R. (2013a). Robust data-driven inference in the regression-discontinuity design. The Stata Journal, vv, ii, 1–34.
Calonico, S., Cattaneo, M. D., & Titiunik, R. (2013b). *Robust nonparametric confidence intervals for regression-discontinuity designs.* Retrieved from http://www.iza.org/conference_files/PolicyEval_2013/cattaneo_m9171.pdf
Campbell, D. T., Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In Gage, N.L. (Ed.), *Handbook of research on teaching* (pp. 171–246). Chicago, IL: Rand McNally.
Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2013). Widening participation in higher education: analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(2), 431-457.
Cochrane, M. (2007). *Spoilt for choice? Pupil perceptions of the options process at Year 9*. Paper presented at the British Educational Research Association Annual Conference, London, 5-8 September. Retrieved from: http://www.leeds.ac.uk/educol/documents/165852.htm
Cochrane, M. (2011). Children's university aspirations and the effects of cultural and social capital. In J. Adams, M. Cochrane, & L. Dunne (Eds.), *Applying theory to*

educational research: An introductory approach with case studies (pp. 95-107). Hoboken: John Wiley & Sons.

Conlisk, J. (1996). Why bounded rationality?. *Journal of economic literature*, 34, 669-700.

Conner, C. (1999). Through a glass darkly: Assessing the future? *Education 3-13*, *27*(3), 32-37.

Conservatives (2013). *Where We Stand – Universities and Skills*. Retrieved from http://www.conservatives.com/Policy/Where_we_stand/Universities_and_Skills.aspx

Cook, T. D. (2008). "Waiting for life to arrive": a history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics*, *142*(2), 636-654.

Covington, M. V. (1992). *Making the grade: A self-worth perspective on motivation and school reform*. Cambridge, MA: Cambridge University Press.

Daugherty, R. (1995). *National Curriculum assessment: A review of policy 1987-1994.* London: Falmer Press.

Davey, G., & Fuller, A. (2013). Hybrid qualifications, institutional expectations and youth transitions: a case of swimming with or against the tide. *Sociological Research Online, 18*(1), 2.

Davies, J., & Brember, I. (1998). National Curriculum Testing and Self−esteem in Year 2—the first 5 years: a cross−sectional study. *Educational Psychology*, *18*(4), 365-375.

Davies, J., & Brember, I. (1999). Reading and mathematics attainments and self-esteem in years 2 and 6-an eight-year cross-sectional study. *Educational Studies*, *25*(2), 145-157.

Department for Business Innovation & Skills (2012). *Participation Rates in Higher Education: Academic Years 2006/2007–2010/2011 (Provisional)*. London: Department for Business, Innovation & Skills. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/16203/12-p140-participation-rates-in-he-2010-11.pdf

Department for Education (2014). *Schools win funds to develop and share new ways of assessing pupils*. Retrieved from https://www.gov.uk/government/news/schools-win-funds-to-develop-and-share-new-ways-of-assessing-pupils

DesJardins, S. L., & McCall, B. P. (2014). The impact of the Gates Millennium Scholars Program on college and post-college related choices of high ability, low-income minority students. *Economics of Education Review*, *38*, 124-138.

DiNardo, J., & Lee, D. S. (2011). Program evaluation and research designs. *Handbook of labor economics*, *4*, 463-536.

Dweck, C. S. (1986). Motivational processes affecting learning. *American psychologist*, *41*(10), 1040-1048.

Dweck, C. S. (2000). *Self theories: Their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.

Dweck, C. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological review*, *95*(2), 256.

Eisenhart, M. A. (1988). The ethnographic research tradition and mathematics

education research. *Journal for Research in Mathematics Education*, 99-114.
Fan, J., & Gijbels, I. (1996). *Local polynomial modelling and its applications: monographs on statistics and applied probability*. New York, NY: Chapman and Hall.
Galindo-Rueda, F., Marcenaro-Gutierrez, O., & Vignoles, A. (2008). The widening socio- economic gap in UK higher education. In S. Gorard (Ed.), *Quantitative Research in Education* (Vol. 1, pp. 75–88). London: SAGE.
Giddens, A. (1991). *Modernity and self-identity: Self and society in the late modern age*. Stanford: Stanford University Press.
Gillborn, D., & Youdell, D. (2000). *Rationing education: Policy, practice, reform, and equity*. Buckingham, UK: Open University Press.
GL assessments (2015). *Life after levels*. Retrieved from http://www.gl-assessment.co.uk/focus/life-after-levels
Glennerster, H. (2002). United Kingdom education 1997–2001. *Oxford Review of Economic Policy*, *18*(2), 120-136.
Graham, S. (1990). Communicating low ability in the classroom: bad things good teachers sometimes do. In S. Graham & V. S. Folkes (Eds.), *Attribution theory: Applications to achievement, mental health, and interpersonal conflict* (pp. 17-36). Hillsdale, NJ: Lawrence Earlbaum.
Greenaway, D., & Haynes, M. (2003). Funding higher education in the UK: the role of fees and loans. *The Economic Journal*, *113*(485), F150-F166.
Grenfell, M. & James, D. (1998). *Bourdieu and education: Acts of practical theory.* London: Falmer Press.
Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*,*69*(1), 201-209.
Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice*, *10*(2), 169-207.
House of Commons Children, Schools and Families Committee (2008). *Testing and assessment: Third report of session 2007–2008* (Vol. 2). London: The Stationery Office Limited.
House of Commons Education Committee (2013). *From GCSEs to EBCs: the Government's proposals for reform.* London: The Stationery Office Limited. Retrieved from http://www.publications.parliament.uk/pa/cm201213/cmselect/cmeduc/808/80802.htm
Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: a guide to practice. *Journal of Econometrics*, *142*(2), 615-635.
Ireson, J., Hallam, S., & Hurley, C. (2005). What are the effects of ability grouping on GCSE attainment?. *British Educational Research Journal*, *31*(4), 443-458.
Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: a regression-discontinuity analysis. *Review of economics and statistics*, *86*(1), 226-244.
Kellaghan, T., Madaus, G. & Raczek, A. (1996). *The use of external examinations to improve student motivation.* Washington, DC: AERA.
Kellaghan, T., & Greaney, V. (2001). *Using assessment to improve the quality of education*. Unesco, International Institute for Educational Planning.
Labour (2010). *The Labour Party Manifesto*. Retrieved from

http://www.labour.org.uk/labours-manifesto-for-a-future-fair-for-all

Lee, D. S., & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, *142*(2), 655-674.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355.

Leonard, M., & Davey, C. (2001). *Thoughts on the 11 Plus.* Belfast: Save the Children.

Liberal Democrats (2013). *What We Stand For – Education*. Retrieved from http://www.libdems.org.uk/education.aspx

Ludwig, J., & Miller, D. L. (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics*, *122*(1), 159-208.

Machin, S., & Vignoles, A. (2004). Educational inequality: the widening socio-economic gap. *Fiscal Studies*, *25*(2), 107–128.

Marcenaro-Gutierrez, O., Galindo-Rueda, F., & Vignoles, A. (2007). Who actually goes to university? *The Economics of Education and Training, 32*(2), 333–357.

McCall, B. P., & Alcott, B. M. (2013). *University Access and Its Determinants: Assessing the Key Points of Divergence*. Paper presented at the annual conference for the British Educational Research Association, Brighton, UK.

McCall, B. P., & Bielby, R. M. (2012). Regression discontinuity design: recent developments and a guide to practice for researchers in higher education. In Smart, J.C, & Paulsen, M. B. (Eds.), *Higher education: Handbook of theory and research* (pp. 249-290). Dordrecht: Springer.

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

Mealli, F., & Rampichini, C. (2012). Evaluating the effects of university grants by using regression discontinuity designs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *175*(3), 775-798.

Nash, R (2005). Cognitive *habitus* and collective intelligence: concepts for the explanation of inequality of educational opportunity. *Journal of Education Policy*, *20*(1), 3-21.

National Audit Office (2008). Third validation compendium report (Vol. 2). London: The Stationery Office Limited.

Nicholls, J. G. (1989). *The competitive ethos and democratic education*. Cambridge, MA: Harvard University Press.

Office for National Statistics (2014). *The National Statistics Socio-economic Classification (NS-SEC rebased on the SOC2010)*. Retrieved from http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html#5

Ovenden, K. (2000). The politics of protest. *Socialist Review*, 242. Retrieved from http://pubs.socialistreviewindex.org.uk/sr242/contents.htm

Papay, J. P., Murnane, R. J., & Willett, J. B. (2010). The consequences of high school exit examinations for low-performing urban students: Evidence from Massachusetts. *Educational Evaluation and Policy Analysis*, *32*(1), 5-23.

Paris, S. G., Lawton, T. A., Turner, J. C., & Roth, J. L. (1991). A developmental perspective on standardized achievement testing. *Educational Researcher*, *20*(5), 12-20.

Pollard, A., Triggs, P., Broadfoot, P., Osborn, M., & McNess, E. (2000). *What pupils say: Changing policy and practice in primary education*. London: Continuum.
Price, J. L. (1977). *The study of turnover*. Ames, IA: University of Iowa Press.
Pugsley, L. (1998). Throwing your brains at it: higher education, markets and choice. *International Studies in Sociology of Education*, 8(1), 71-92.
QCA (2004). *Report on Key Stage 3 English Review of Service Delivery Failure 2003–2004 to QCA Board*. Retrieved from dera.ioe.ac.uk/8228/1/10343_ks3_en_report_04.pdf
Quinlan, M., & Scharaschkin, A. (1999). *National curriculum testing: problems and practicalities. Paper presented at the British Educational Research Association Annual Conference.* Retrieved from www.assessnet.org.uk/e-learning/file.php/1/.../060709_Paper.pdf
Reay, D. (1995). 'They employ cleaners to do that': habitus in the primary classroom. *British journal of sociology of education*, *16*(3), 353-371.
Reay, D. (1998). 'Always knowing' and "never being sure": familial and institutional habituses and higher education choice. *Journal of Education Policy*, *13*(4), 519–529.
Reay, D., (2001). Finding or losing yourself?: Working-class relationships to education. *Journal of Education Policy*, 16, 333-346.
Reay, D. (2004). "It's all becoming a habitus": beyond the habitual use of Pierre Bourdieu's concept of habitus in educational research. *British Journal of Sociology of Education, 25*(4), 431–44.
Reay, D., & William, D. (1999). *'I'll be a nothing': structure, agency and the construction of identity through assessment*. British Educational Research Journal, *25*(3), 343–354.
Rosenbaum, J. E. (1986). Institutional career structures and the social construction of ability. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 139-171). New York, NY: Greenwood.
Sartarelli, M. (2011). *Do performance targets affect behaviour? Evidence from discontinuities in test scores in England. DoQSS Working Paper No. 11-02*. London.
Schultz, T. (1961). Investment in human capital. *The American Economic Review* *51*(1), 1–17.
Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational psychologist*, *26*(3-4), 207-231.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
Sjögren, A. (2010). *Graded children–Evidence of longrun consequences of school grades form a nationwide reform*. IFN Working Paper No. 839.
Spence, M. A. (1973). Job market signalling. *The quarterly journal of Economics*, 355-374.
Spence, M. A. (1981). Signalling, screening, and information. In S. Rosen (Ed.), *Studies in labor markets* (pp. 319-357). Chicago: University of Chicago Press.
Steedman, C. (1988). The mother made conscious: the historical development of primary school pedagogy. In M. Woodhead & A. McGrath (Eds.), *Family, school & society* (pp. 82-95). London: The Open University.
Taylor, C. (2012). *Improving attendance at schoo*l. London; Department for

Education.
Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, *51*(6), 309.
Thomas, L., Bland, D., & Duckworth, V. (2012). Teachers as advocates for widening participation. *Widening Participation and Lifelong Learning*, *14*(2), 40-58.
Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: a regression–discontinuity approach. *International Economic Review*, *43*(4), 1249-1287.
Vignoles, A. (2013). Widening participation and social mobility. In C. Callendar & P. Scott (Eds.), *Browne and beyond: Modernizing English higher education* (pp. 112–129). London: Institute of Education Press.
Vignoles, A., & Powdthavee, N. (2009). The socio-economic Gap in University Dropouts. *The B.E. Journal of Economic Analysis & Policy*, *9*(1).
Vincent, C. (2001). Social class and parental agency. *Journal of Education Policy*, *16*(4), 347-364.
Vincent, C., Braun, A., & Ball, S. J. (2008). Childcare, choice and social class: caring for young children in the UK. *Critical Social Policy*, *28*(1), 5-26.
Vispoel, W. P., & Austin, J. R. (1995). Success and failure in junior high school: A critical incident approach to understanding students' attributional beliefs. *American Educational Research Journal*, *32*(2), 377-412.
Walker, M., & Clark, G. (2010). Parental choice and the rural primary school: lifestyle, locality and loyalty. *Journal of Rural Studies*, *26*(3), 241-249.
Weiner, B. (2010). The development of an attribution-based theory of motivation: A history of ideas. *Educational Psychologist*, *45*(1), 28-36.
Wigfield, A., & Karpathian, M. (1991). Who am I and what can I do? Children's self-concepts and motivation in achievement situations. *Educational Psychologist*, *26*(3-4), 233-261.
Wilshaw, M. (2013). *The Annual Report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2012/2013*. London: The Stationery Office.

Appendix 2.A: Sensitivity tests around the grade thresholds

This appendix tests the suitability of the Key Stage 3 data for the RDD approach by testing for discontinuities in prior student characteristics around the Level 6 cut-off point in each assessment. Tables 2.A1–2.A3 test for discontinuities in NSSEC status, and Tables 2.A4–2.A6 test for discontinuities in performance at Key Stage 2, as measured by continuous test scores in those assessments.

Judging the models as a group, there is little evidence that students were able to systematically manipulate their position around the cut-off points. Of the 18 models, just one provides an estimate that is significantly different from zero. The significant estimate comes from the model that regresses students' positions around the Level 6 threshold of the Key Stage 3 mathematics assessment on their test scores in Key Stage 2 mathematics for students (column (1) of Table 2.A5). This finding seems to be counterintuitive, as it indicates that students scoring just above the Level 6 threshold did significantly worse in Key Stage 2 mathematics than those who scored just below this threshold.

All models are created using Calonico et al.'s (2012a) bandwidth selection approach and construction of standard errors. Each model uses a uniform kernel and local-polynomial of order one, i.e., linear. For all tables, standard errors are provided in parentheses. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.1

Table 2.A1: Test for discontinuity in rate of low-NSSEC children, by assessment.

	(1) English	(2) Mathematics	(3) Science
RD Estimate	-0.00808 (0.0172)	-0.0163 (0.0135)	-0.0110 (0.0158)
Observations	4,261	6,070	3,994

Table 2.A2: Test for discontinuity in rate of mid-NSSEC children, by assessment.

	(1) English	(2) Mathematics	(3) Science
RD Estimate	-0.0268 (0.0273)	0.0103 (0.0181)	-0.00504 (0.0227)
Observations	3,405	6,033	4,260

Table 2.A3: Test for discontinuity in rate of high-NSSEC children, by assessment.

	(1) English	(2) Mathematics	(3) Science
RD Estimate	0.0502 (0.0381)	-0.0244 (0.0254)	0.0138 (0.0273)
Observations	2,212	4,051	3,653

Table 2.A4: Test for discontinuity in Key Stage 2 English score, by assessment.

	(1) English	(2) Mathematics	(3) Science
RD Estimate	-0.156 (0.144)	0.248 (0.165)	-0.0547 (0.188)
Observations	4,147	2,744	3,698

Table 2.A5: Test for discontinuity in Key Stage 2 math score, by assessment.

	(1) English	(2) Mathematics	(3) Science
RD Estimate	-0.780*** (0.227)	-0.153 (0.170)	-0.193 (0.156)
Observations	2,395	5,481	3,278

Table 2.A6: Test for discontinuity in Key Stage 2 science score, by assessment.

	(1) English	(2) Mathematics	(3) Science
RD Estimate	-0.138 (0.152)	0.108 (0.156)	0.0143 (0.110)
Observations	2,498	3,312	5,540

Appendix 2.B: Specifications for discontinuity models

In this appendix, the following tables provide more detailed information about the RDD models used in Section 6.

Table 2.B1: Specification for models presented in Table 2.4, column (1)

	(1) Capped GCSE point score	(2) Attain C or higher in GCSE English	(3) Enrollment in A-levels	(4) Enrollment in university degree
RD Estimate	3.741	0.0280	0.0938	0.0389
Observations	5342	3405	2235	3276
Conventional S.E.	2.738	0.0250	0.0405	0.0328
Conventional p value	0.172	0.262	0.0207	0.236
Robust 95% CI	[-.79 ; 11.96]	[-.01 ; .1]	[.02 ; .2]	[-.03 ; .13]
Robust p-value	0.0858	0.137	0.0128	0.197
Order Loc. Poly.	1	1	1	1
Order Bias	2	2	2	2
BW Loc. Poly.	3.078	1.868	1.715	3.173
BW Bias	5.832	3.607	3.654	5.530
Kernel Type	Uniform	Uniform	Uniform	Uniform

Table 2.B2: Specification for models presented in Table 2.4, column (2)

	(1) Capped GCSE point score	(2) Attain C or higher in GCSE math	(3) Enrollment in A-levels	(4) Enrollment in university degree
RD Estimate	4.788	0.0170	0.0369	0.0610
Observations	3803	2751	4707	2368
Conventional S.E.	3.852	0.0388	0.0261	0.0375
Conventional p value	0.214	0.661	0.158	0.104
Robust 95% CI	[-5.21 ; 12.44]	[-.08 ; .08]	[-.01 ; .11]	[-.03 ; .14]
Robust p-value	0.422	0.968	0.134	0.208
Order Loc. Poly.	1	1	1	1
Order Bias	2	2	2	2
BW Loc. Poly.	2.511	1.698	4.611	2.904
BW Bias	4.825	4.089	8.610	5.588
Kernel Type	Uniform	Uniform	Uniform	Uniform

Table 2.B3: Specification for models presented in Table 2.4, column (3)

	(1) Capped GCSE point score	(2) Enrollment in A-levels	(3) Enrollment in university degree
RD Estimate	-2.149	-0.00974	-0.0147
Observations	3643	3364	2601
Conventional S.E.	2.976	0.0295	0.0358
Conventional p value	0.470	0.741	0.680
Robust 95% CI	[-8.39 ; 5.54]	[-.09 ; .04]	[-.11 ; .06]
Robust p-value	0.688	0.505	0.558
Order Loc. Poly.	1	1	1
Order Bias	2	2	2
BW Loc. Poly.	2.315	2.780	2.788
BW Bias	4.302	5.923	5.068
Kernel Type	Uniform	Uniform	Uniform

Table 2.B4: Specification for models presented in Table 2.5, column (1) for low-NSSEC students

	(1) Capped GCSE point score	(2) Attain C or higher in GCSE English	(3) Enrollment in A-levels	(4) Enrollment in university degree
RD Estimate	20.11	0.207	0.110	0.205
Observations	711	681	949	462
Conventional S.E.	8.405	0.0619	0.0538	0.0925
Conventional p value	0.0167	0.000842	0.0414	0.0267
Robust 95% CI	[3.03 ; 42]	[.11 ; .38]	[.01 ; .26]	[.02 ; .47]
Robust p-value	0.0235	0.000405	0.0411	0.0311
Order Loc. Poly.	1	1	1	1
Order Bias	2	2	2	2
BW Loc. Poly.	2.387	2.257	3.120	1.831
BW Bias	3.917	4.813	5.661	3.243
Kernel Type	Uniform	Uniform	Uniform	Uniform

Table 2.B5: Specification for models presented in Table 2.5, column (2) for low-NSSEC students

	(1) Capped GCSE point score	(2) Attain C or higher in GCSE math	(3) Enrollment in A-levels	(4) Enrollment in university degree
RD Estimate	28.21	0.0569	0.0819	0.00816
Observations	965	831	1073	786
Conventional S.E.	7.570	0.0657	0.0556	0.0656
Conventional p value	0.000195	0.387	0.141	0.901
Robust 95% CI	[13.31 ; 48.06]	[-.11 ; .18]	[-.03 ; .23]	[-.16 ; .14]
Robust p-value	0.000536	0.655	0.144	0.857
Order Loc. Poly.	1	1	1	1
Order Bias	2	2	2	2
BW Loc. Poly.	3.450	2.872	3.923	3.609
BW Bias	6.947	6.311	7.162	6.951
Kernel Type	Uniform	Uniform	Uniform	Uniform

Table 2.B6: Specification for models presented in Table 2.5, column (3) for low-NSSEC students

	(1) Capped GCSE point score	(2) Enrollment in A-levels	(3) Enrollment in university degree
RD Estimate	-0.424	-0.0169	0.0154
Observations	650	917	595
Conventional S.E.	6.854	0.0508	0.0719
Conventional p value	0.951	0.739	0.831
Robust 95% CI	[-16.32 ; 15.81]	[-.16 ; .07]	[-.17 ; .14]
Robust p-value	0.975	0.463	0.839
Order Loc. Poly.	1	1	1
Order Bias	2	2	2
BW Loc. Poly.	2.437	3.398	2.790
BW Bias	4.410	6.611	5.893
Kernel Type	Uniform	Uniform	Uniform

Table 2.B7: Specification for models presented in Table 2.5, column (1) for mid-NSSEC students

	(1) Capped GCSE point score	(2) Attain C or higher in GCSE English	(3) Enrollment in A-levels	(4) Enrollment in university degree
RD Estimate	-1.189	-0.0202	0.111	0.0152
Observations	454	1059	632	579
Conventional S.E.	8.467	0.0347	0.0752	0.0832
Conventional p value	0.888	0.561	0.141	0.855
Robust 95% CI	[-24.81 ; 14.87]	[-.09 ; .07]	[-.05 ; .31]	[-.2 ; .2]
Robust p-value	0.624	0.892	0.154	0.974
Order Loc. Poly.	1	1	1	1
Order Bias	2	2	2	2
BW Loc. Poly.	1.437	3.069	1.850	2.007
BW Bias	2.637	6.282	3.422	3.369
Kernel Type	Uniform	Uniform	Uniform	Uniform

Table 2.B8: Specification for models presented in Table 2.5, column (2) for mid-NSSEC students

	(1) Capped GCSE point score	(2) Attain C or higher in GCSE math	(3) Enrollment in A-levels	(4) Enrollment in university degree
RD Estimate	5.438	0.0553	0.0605	0.144
Observations	910	729	842	600
Conventional S.E.	5.873	0.0706	0.0652	0.0707
Conventional p value	0.354	0.434	0.354	0.0420
Robust 95% CI	[-8.07 ; 19.35]	[-.12 ; .19]	[-.1 ; .2]	[.01 ; .33]
Robust p-value	0.420	0.680	0.530	0.0427
Order Loc. Poly.	1	1	1	1
Order Bias	2	2	2	2
BW Loc. Poly.	3.425	2.625	3.054	2.611
BW Bias	6.034	5.327	5.833	4.888
Kernel Type	Uniform	Uniform	Uniform	Uniform

Table 2.B9: Specification for models presented in Table 2.5, column (3) for mid-NSSEC students

	(1) Capped GCSE point score	(2) Enrollment in A-levels	(3) Enrollment in university degree
RD Estimate	-7.654	-0.0447	-0.0390
Observations	767	582	640
Conventional S.E.	4.686	0.0718	0.0714
Conventional p value	0.102	0.534	0.585
Robust 95% CI	[-18.49 ; 3.04]	[-.21 ; .12]	[-.23 ; .09]
Robust p-value	0.160	0.613	0.404
Order Loc. Poly.	1	1	1
Order Bias	2	2	2
BW Loc. Poly.	2.405	1.761	2.321
BW Bias	4.537	3.514	4.565
Kernel Type	Uniform	Uniform	Uniform

Table 2.B10: Specification for models presented in Table 2.5, column (1) for high-NSSEC students

	(1) Capped GCSE point score	(2) Attain C or higher in GCSE English	(3) Enrollment in A-levels	(4) Enrollment in university degree
RD Estimate	-1.062	0.0459	-0.0680	-0.0301
Observations	1667	1121	1356	1311
Conventional S.E.	4.200	0.0370	0.0448	0.0519
Conventional p value	0.800	0.215	0.129	0.562
Robust 95% CI	[-11.86 ; 8.14]	[-.01 ; .15]	[-.16 ; .05]	[-.14 ; .1]
Robust p-value	0.716	0.101	0.306	0.747
Order Loc. Poly.	1	1	1	1
Order Bias	2	2	2	2
BW Loc. Poly.	3.375	2.400	2.869	3.147
BW Bias	6.153	5.651	4.972	5.426
Kernel Type	Uniform	Uniform	Uniform	Uniform

Table 2.B11: Specification for models presented in Table 2.5, column (2) for high-NSSEC students

	(1) Capped GCSE point score	(2) Attain C or higher in GCSE math	(3) Enrollment in A-levels	(4) Enrollment in university degree
RD Estimate	9.494	0.108	0.0935	0.0394
Observations	1140	1142	1150	1342
Conventional S.E.	5.163	0.0539	0.0554	0.0477
Conventional p value	0.0659	0.0453	0.0914	0.409
Robust 95% CI	[-.47 ; 22.52]	[-.03 ; .21]	[-.01 ; .24]	[-.08 ; .14]
Robust p-value	0.0602	0.131	0.0833	0.585
Order Loc. Poly.	1	1	1	1
Order Bias	2	2	2	2
BW Loc. Poly.	3.396	3.463	3.485	5.225
BW Bias	7.046	6.500	6.673	9.630
Kernel Type	Uniform	Uniform	Uniform	Uniform

Table 2.B12: Specification for models presented in Table 2.5, column (3) for high-NSSEC students

	(1) Capped GCSE point score	(2) Enrollment in A-levels	(3) Enrollment in university degree
RD Estimate	3.997	0.0504	0.0120
Observations	1221	1223	1313
Conventional S.E.	4.452	0.0494	0.0506
Conventional p value	0.369	0.307	0.813
Robust 95% CI	[-6.76 ; 14.39]	[-.08 ; .15]	[-.1 ; .14]
Robust p-value	0.480	0.513	0.739
Order Loc. Poly.	1	1	1
Order Bias	2	2	2
BW Loc. Poly.	2.812	2.876	3.659
BW Bias	5.111	5.378	6.321
Kernel Type	Uniform	Uniform	Uniform

Chapter 3: Did the Education Maintenance Allowance Increase Enrollment in Post-16 Education?

Co-authored with Brian McCall

Summary

From 2004 to 2011, England's government provided a conditional cash transfer – the Education Maintenance Allowance – for students below a household-income threshold who remained in post-compulsory education. While studies of the policy's pilot program demonstrated an impact on student retention and progression to higher education, comparable research on the program after its national rollout is lacking. We attempt to rectify this gap in the research literature. Using data from the Family Resources Survey, we use regression discontinuity analyses to ascertain the program's impact over the seven years that it was in operation. Further, we investigates whether the program's impact differed according to student gender and age. Results indicate that the policy did increase enrollment in post-compulsory education, although there is insufficient evidence to suggest that the policy had differential effects according to the aforementioned student characteristics.

Introduction

In the United Kingdom, as in the United States, there has been mounting pressure to increase education levels among citizens. Typically, this is predicated on evidence that additional education generates greater financial returns (Blundell et al., 2005; Walker and Zhu, 2011), leading to the claim that policymakers should promote post-secondary education in order for the British economy to remain globally competitive. Further, such policies should focus on socio-economically underprivileged students, given that the recent growth in further and higher education enrollment has occurred at a disproportionately high rate among the more privileged (Blanden and Machin, 2004; Vignoles and Powdthavee, 2009). In fact, there are concerns that some existing policy initiatives aimed at increasing the number of university graduates may widen the already significant educational gap between high and low income youth, as the former tend to be more responsive than the latter to policies aimed at increasing attendance rates (Archer et al., 2003; Ball et al., 2002).

One means of mitigating this problem is to implement policies aimed directly at youth from low income groups. The Education Maintenance Allowance (EMA) program is one of the key initiatives by UK governments to do this. Initiated in 2004, the EMA is a conditional cash transfer (CCT) offering youths from low-income households up to £30 (approximately US$45) per week if they remain in education between ages 16–19. Payments are made directly to youths, and they are free to use the money as they choose. In 2011, 45% of 16 to 18 year-olds in full-time education received the EMA (Department for Education, 2011). Research suggests that recipients were more likely to use EMA payments for costs relating to transport, school equipment, and their households than for social expenses (Ashworth et al., 2001; Legard et al., 2001).

However, while the governments of Scotland, Wales, and Northern Ireland have maintained their EMA programs, England's incumbent coalition government abolished its EMA program. This is in keeping with the English government's broader divergence in policymaking from the other UK governments, whereby the English government has increasingly sought to deregulate and reduce its expenditure. For the K-12 system, such policies include the introduction of free schools (which, much like charter schools, receive government funds but are independent of school districts), a reduction in funding for teacher training, and the removal of the requirement for teachers to hold a teaching qualification. For the financing of higher education, students have increasingly had to directly bear the costs of study. In 2010, England's government trebled tuition fee caps for undergraduate courses to £9,000 (approximately $13,000).

For the years between compulsory schooling and higher education, i.e., student ages 16–18, two major policy changes have occurred. In the first, the coalition government has implemented the preceding Labour government's plan to increase the school-leaving age to 18; until 2013, students could leave education at the end of the school year in which they turned 16. However, this policy does not require students to remain in full-time education after the age of 16: apprenticeships and part-time study, or training alongside employment or volunteering, suffice. Consequently, students must now remain in some form of education or training until age 18, but the range of permissible activities is broad. There is not yet empirical research on the impact of this policy change.

In the second change, the government went against Labour's plan to maintain the EMA as a means of supporting students from poorer households aged 16–19. In 2011, the government replaced the EMA, on which the government spent £560 million in its final year, with the 16–19 bursary fund, on which the government currently spends £180 million per year. Whereas the EMA was paid directly to students under a set criteria, bursary funds are allocated to schools, who then decide how to distribute the funds among their student bodies. Thus, the EMA's replacement is of lower total value, and its conditions are likely to be more opaque to many students. Again, there is not yet empirical research on the impact of this policy change.

The English government's rationale for abolishing the EMA has been contentious, arguing that the costs – £560 million in 2010 – were not reflected by sufficient policy impact (Department for Education, 2011). Citing research from the National Foundation for Educational Research (2011), the government claimed that 90% of 16 to 18 year-olds who were in education would still have been so in the absence of EMA payments. The key research, by Spielhofer et al. (2010), used survey data from 838 EMA recipients, of whom 12% agreed with the statement "they would not have participated in the courses they are doing if they had not received an EMA" (National Foundation for Educational Research, 2011, 2).

However, competing evidence exists of the EMA having a positive effect on enrollment and attainment. The most methodologically robust research is based on data collected during the EMA's pilot stage. This pilot stage began in 1999, when the EMA was implemented in 15 of England's 150 school districts. Four variants of EMA payments were trialled; these variants differed with regards to their maximum payment (£30 or £40), retention payments, and achievement bonuses. In addition, one made payments to students' parents rather than directly to students. During the pilot stage of the EMA, researchers identified a matched control group of school districts according to measures of neighborhood deprivation, school performance, and enrollment rates in post-compulsory education. In both pilot and control districts, data was collected on students' backgrounds in order to enable analysis with matching methods.

Research using this pilot stage data indicates that the EMA increased enrollment in post-compulsory education. Using kernel based matching models, Dearden et al. (2009) estimate that the EMA increased enrollment by approximately 4.5 percentage points in the first year after compulsory education and 6.7 percentage points in the second year, suggesting that the EMA increased not only enrollment but retention. The authors estimate that approximately two thirds of these students were drawn from unemployment rather than paid work. The authors also found that the EMA had a greater impact on enrollment among children whose families rented their housing than children whose family owned their housing, although this difference was not significant at the .10 level.

The EMA's impact on achievement is not quite as clear. In their analysis of the pilot stage data with propensity-score matching methods, Middleton et al.'s (2003) estimates of the EMA's impact on achievement were largely positive but not statistically significant, leading the authors to speculate that a larger sample would probably yield significant positive estimates. Subsequently, Chowdry et al. (2007) estimated linear matching models of the EMA's impact on achievement, but increased their sample size by using all of England's remaining school districts in their control sample. The authors found the EMA to have no positive impact on achievement among female students, but that average performance in A-level exams (taken at age 18) increased by 8.9% among male students.

However, no research exists to date on the policy's national rollout that uses the type of quasi-experimental methods necessary for robust causal inferences of the study's impact. There is reason to believe that conditions may be different since the national rollout; for example, the pilot policy was introduced after targeted students completed national tenth grade examinations (GCSEs). Since subsequent study opportunities depend upon GCSE performance, it is plausible that the impact of the EMA differs when students know of its existence prior to these exams. In addition, the value of EMA payments was fixed in absolute terms during the program's duration, so the inflation adjusted value of these payments decreased. To provide a point of reference, the Retail Price Index increased by 12.9% between the pilot study in 1999 and the EMA's national rollout in 2004 (Chowdry et al. 2007). Conversely, the household income threshold for the maximum EMA payment was

£13,000 during the pilot and £20,817 during the national rollout, meaning that a greater proportion were able to receive this maximum payment.

The present study resolves this gap in the research literature. Using nationally representative governmental datasets for each of the years of the EMA's implementation, we provide robust estimates of the policy's impact on enrollment levels among the target population. We do this by using a regression discontinuity design because, as discussed in later sections, this approach is appropriate for the EMA's eligibility criteria. Our primary research question in this study is as follows: what impact did the EMA program have among eligible 16- to 18 year-olds on the probability of enrollment in full-time education? In addition, we also test for heterogeneous effects of the EMA on enrollment according to region, youth, gender and age.

Theoretical Framework

Human capital theory postulates that anticipated financial returns are the primary motivating factor in students' enrollment decisions. The core argument underlying this claim is that skill development can lead to greater wealth accumulation in the future; hence, education is a financial investment. Prior to human capital, conventional economic theory held that tertiary education offered a form of consumption: students choose to take part for recreation and status (Machin, 2008). Gary Becker and Theodore Schultz, to whom the development of human capital theory is most commonly attributed (Dearden et al., 2011), claimed instead that individuals choose to undertake education so long as they think that the income premium resulting from their education will offset the costs they bear – primarily foregone income and study fees – in order to study.

Becker (1980) and Schultz (1961) acknowledge Adam Smith, an eighteenth-century economist, for initiating the notion of human capital. In *The Wealth of Nations*, Smith

(1776/2007) accounted for the skills of citizens when assessing the wealth of a given country. Further, he claimed that differences among people "arise not so much from nature, as from habit, custom, and education" (p. 15). Smith thus theorized that humans' skills influence their ability to generate wealth, and that these skills are fashioned, at least in part, by education. Becker and Schultz then sought to build on Smith's arguments by quantifying the economic returns to education. Focusing on white males in the United States in the first half of the twentieth century, both declared higher education to be a profitable investment. In this context, Becker (1980) argued that the average monetary return to college graduates was between 11–13%. Schultz (1961) posited that education had outpaced other forms of investments, claiming that its rate of return on a dollar invested had doubled in relation to alternatives.

Researchers argue that such investments in education are similarly beneficial in the United Kingdom, even when accounting for heterogeneous returns (Blanden and Machin, 2004; Blundell et al., 2005; Dearden, 1999; Dearden et al., 2002; Dolton and Vignoles, 2000). Further, more recent research claims that the degree premium will endure despite the fee increases. Barr (2011) argues that the incoming fee system is actually unnecessarily favorable to students and the government could charge graduates greater repayments in their early careers without threatening the economic value of a degree. Walker and Zhu's (2011) analyses, which allowed for, and found, greater variation in returns to degree completion than previously acknowledged in the English research literature, still led them to claim that the new tuition fees are "dwarfed by the scale of lifecycle earnings differentials" (p. 1184).

Human capital theory treats students' motivations and behaviors as though they consistently reflect its own description of the role of education. As a consequence, students are presumed to decide whether to enroll in education on the basis of anticipated future earnings. In addition, students are expected to be adept at making this judgment. Becker claimed that students operate on the basis of the "well-known equilibrium condition" (1976, p. 123) that they should take on tertiary education to the point where present costs outweigh expected future returns. Although Mincer (1974), Card (1999, 2001) and Heckman et al. (2006) have made major subsequent adjustments to Becker's (1967) calculations of what constitutes and how to calculate costs and returns, each has retained the underlying principle of equilibrium.

As a consequence, this framework theorizes that students both try and are successful in projecting future earning opportunities for graduates and non-graduates. When challenged on this assumption, Becker (1980) responded:

> Children have their decisions guided, however, as well as partly financed, by their parents, and as long as parents recover some monetary or psychic benefits from an increase in their children's economic well-being, parents have an incentive to help children make wise decisions. (p. 105)

Although this reply acknowledges that students might not make enrollment decisions entirely independently, it does not concede that students, even with the support of others, might make inaccurate forecasts about future returns from education.

This conviction – that students make sound financial judgments – underlies one explanation that human capital theory offers for differing degrees of enrollment according to student characteristics: that students vary in their decisions to enroll because their future returns also vary. This is characterized in the literature as heterogeneous returns. Becker (1980) allowed for a limited range of variations in student circumstances, comparing White males to White females and African American males, arguing that lower enrollment among the latter two groups in comparison to White males was explained by relatively lower income premiums from their college degrees.

Subsequent research utilizing human capital theory has developed more sophisticated means for theorizing and measuring heterogeneous returns to education, such as accounting more accurately for the challenge of endogeneity wherein students' educational choices are driven by pre-existing characteristics that may be correlated to future earnings even in the absence of additional education (see, for example, Blundell et al., 2000; Card, 1999, 2001; Heckman et al., 2006). Nonetheless, as they become increasingly intricate, these studies still adhere to the principle that students' enrollment decisions are based primarily on accurate financial forecasts that account for their specific circumstances.

Human capital theory has also identified factors that constrain the opportunities of some student groups to enroll in university even when they stand to gain from additional education. Given Becker's (1980) aforementioned claims about the predominance of financial considerations and the capacity for students to choose the most profitable opportunities available to them, it is perhaps unsurprising that he and other early exponents of human capital held that "the most important cause of difference in students' opportunities is the availability of funds" (Becker, 1980, p. 107), a factor termed within this literature as credit constraints. Since parents are presumed to assist their children in accessing higher education, limits in parental income should lead to more restricted opportunities for those children to manage the short-term financial burdens of university study (Becker and Tomes, 1986). This is exacerbated by the fact that human capital is not entirely analogous to more classical forms of capital because, barring a scenario in which human beings can be treated as commodities, it cannot be bought, sold, or offered as collateral for investments (Becker, 1980; Checchi, 2006).

Since 1998, England has followed a policy, proposed initially by Friedman (1955), wherein the government offers income-contingent loans for undergraduate students; repayments are then taken as a proportion of the borrower's subsequent earnings and any outstanding debt is forgiven after 20 years. This policy is congruous with human capital theory's explications of enrollment disparities because it defers the cost of study until a future time when poorer students would not face credit constraints (Barr, 2009, 2010; Glennerster, 1991). Further, enrollment rates have continued to grow in spite of changing fee and loan systems (Universities and Colleges Admissions Service, 2013).

Consequently, the research literature to date on educational access in the United Kingdom has focused instead on academic preparation, one of the longer-term constraints identified by Carneiro and Heckman (2002, 2003). Attainment through the compulsory school years is a key indicator of progression to higher education (Chowdry et al., 2013), and it serves to explain the great majority of the disparity across socio-economic classes for those who do progress through the final years of schooling (Marcenaro-Gutierrez et al., 2007). However, this caveat – those who progress to the final years of schooling – is crucial. In order to enroll in university, a student should complete A-level examinations that are studied for and taken at age 18, but compulsory schooling in England ends before youths reach this age. In 2011, about 25% of 17 years olds were not in school, and this fraction is considerably larger for 17 year-olds from low-income households. As such, increasing the number of youth who remain in education up to age 18 is an important policy issue for university access.

Policy background

One policy mechanism aimed at reducing attrition is the conditional cash transfer (CCT). Besides the United Kingdom's EMA policy, CCTs aimed at tackling educational attrition have been implemented in Colombia (Attanasio et al., 2010), Mexico (Attanasio et al., 2012), New York (Morris et al., 2012), and Australia (Dearden and Heath, 1996). CCTs provide recipients with a direct cash reward for fulfilling a given criteria – for education policy, the criteria are typically enrollment and attendance in an education program. These rewards tend to offer partial, rather than comprehensive, cover for the financial costs of remaining in education. For instance, in 2005 the EMA's maximum weekly payment, £30, was equivalent to wages for ten hours of work at the minimum wage for 16–17 year olds, or six hours of work at the minimum wage for 18 year olds. Nonetheless, CCTs embody the key principles of human capital theory: the marginal cost of enrolling in education is reduced for the individual, but the marginal benefit of additional education is unaffected. Thus, such policy tools should provide governments with an effective means to increase enrollment in scenarios in which they deem it to be sub-optimally low.

The English government provided the EMA between 2004–2011. The policy's target recipients were those aged 16 to 19 who had finished compulsory education (typically completed in England at age 16) who were enrolled in some form of full-time education. EMA was means-tested: the allowance was £30 (approximately US$45) per week for students whose annual household income was less than £20,817; £20 per week for those whose annual household income was between £20,818 and £25,521; and £10 per week for those whose annual household income was between £25,522 and £30,810. These thresholds exclude maintenance paid by non-resident parents and any income earned by the student. Until 2008, students could also receive two additional payments of £100 each year for completing the autumn and summer school terms. Students could receive EMA for up to three years, with eligibility being assessed annually, and receipt of EMA did not affect eligibility for any other government benefits. Money was paid directly into students' bank accounts.

The English government ended the EMA in 2011, although enrolled students were able to receive it for an additional year. The following year, the government replaced EMA with the 16 to 19 Bursary Fund, which offers up to £1,200 per year to a more restricted number of students. To date, the governments of Scotland, Wales, and Northern Ireland have kept the EMA.

Evaluation methodology

This paper investigates whether eligibility for the EMA influenced enrollment rates for eligible students. To do so, it employs the Neyman-Rubin counterfactual framework (Neyman, 1923; Rubin, 1974). Letting Y_i represent enrollment for student i, and denoting eligibility for EMA as $T = 1$,

$$Y_i = \begin{cases} Y_i^1 \; if \; T_i = 1 \\ Y_i^0 \; if \; T_i = 0 \end{cases} \tag{1}$$

and the causal impact of EMA on enrollment for student i is $(Y_i^1 - Y_i^0)$. However, since it is not possible to directly observe both Y_i^0 and Y_i^1, it is necessary to use an evaluation strategy that emulates the counterfactual, i.e., the condition that did not occur in real-world conditions. The Neyman-Rubin counterfactual framework does this by comparing the average outcome for treated participants (Y_i^1) to the average outcome for untreated participants (Y_i^0). Thus, from a sufficiently large sample, it is possible to estimate $(Y_i^1 - Y_i^0)$ by

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} T_i + \varepsilon_i . \tag{2}$$

However this estimate (β) of the true treatment effect (β) relies on the ignorable treatment assignment assumption (Rosenbaum and Rubin, 1983), which can be expressed as

$$(Y_0, Y_1) \perp T . \tag{3}$$

In other words, if $Cov(T_i, \varepsilon_i) \neq 0$, $E(\beta) \neq \beta$, and so the estimated program impact of EMA (β) will be biased. Random allocation of T would circumvent this challenge, since T_i then occurs independently of ε_i. However, given the means-tested nature of EMA, alternatives are necessary to assess its impact during its national rollout.

Absent random allocation, this paper uses the quasi-experimental method known as regression discontinuity design (RDD). The structure of the EMA lends itself to RDD because household income "cut-off" levels defined eligibility for EMA payments. RDD is based on the assumption that the running variable (in this case, annual household income) has a continuous impact on the outcome variable (enrollment in post-compulsory education), but that the treatment (receipt of an EMA payment) changes discontinuously as a function of the running variable (Imbens and Lemieux, 2008; McCall and Bielby, 2012). This structure allows the causal impact of the treatment on the outcome to be identified at the cut-off points.

In this paper, we focus on the effect of EMA eligibility on school enrollment. Since eligibility, rather than receipt, of EMA serves as the treatment, the EMA program satisfies a "sharp" RDD because the cut-off variable defines allocation of the treatment (Imbens and Lemieux, 2008). This is because there are clear boundaries according to the running variable: as noted, households with income above £30,810 are not eligible for an allowance whereas those in households with income between £25,522 and £30,810, £20,818 and £25,521, or £20,817 or less, are eligible for a payment of £10 per week, £20 per week, or £30 per week, respectively.

Intuitively, whether students' household incomes fall just above or just below a cut-off point is approximately randomly distributed, but whether household income falls above or below the cut-off point has a discrete effect on their eligibility for EMA. When this scenario holds, variation in treatment is essentially randomized around the cut-off (Lee and Lemieux, 2010). It is important to note though that any estimate of treatment effects around the cut-off are local (i.e., pertaining only to observations around the cut-off) rather than general treatment effects (DiNardo and Lee, 2011; Shadish et al., 2002).

Let λ_k represent the local average treatment effect of increasing weekly allowance amount from 10 × (*k*-1) to 10 × *k*, *k=1,2,3*. Then

$$\lambda_k = \frac{\lim_{d\to 0} E(Y \mid c_k > I > c_k - d) - \lim_{d\to 0} E(Y \mid c_k + d > I > c_k)}{\lim_{d\to 0} \Pr(T = 1 \mid c_k > I > c_k - d) - \lim_{d\to 0} \Pr(T = 0 \mid c_k + d > I > c_k)} \qquad (4)$$

Since we have a sharp design, the elements of the denominator sum to one, leaving

$$\lambda_k = \lim_{d\to 0} E(Y \mid c_k > I > c_k - d) - \lim_{d\to 0} E(Y \mid c_k + d > I > c_k) \qquad (5)$$

where c_1 = £30,810 , c_2 = £25,521, c_3 = £20,817 and the last equality holds because nobody with income above the cut-off receives a higher allowance.[2]

RDD estimates can be obtained either using parametric or non-parametric methods (see Hahn et al., 2001; Lee and Lemieux, 2010; and McCall and Bielby, 2012). For parametric specifications we use linear, quadratic and cubic polynomial specifications of the running variable. Furthermore, we only include observations whose household income is within a pre-specified bandwidth of the cut-point. We check the robustness of the estimates to different choices of bandwidth. As this bandwidth decreases, we expect the bias of the estimates to be reduced, but since the number of observations falls concomitantly, the variance of the estimates increases. Moreover, at lower bandwidths the estimates are based on data that is close to only one cut-off point.

More specifically, we estimate the following regression models:

$$y = \alpha_k I(x_k \le 0) + \sum_{s=0}^{S} \beta_k^s x_k^s + \varepsilon \qquad \text{for} \qquad x_k \in \left[-d^j, d^j\right] \qquad (6)$$

for s = 1, 2, 3 and j = 1, 2, 3, 4 where d^1 = £3750, d^2 = £2500, d^3 = £1750, and d^4 = £1750 and x_k is the running variable measured as a deviation from the cut-off value c_k. The

2 For the first couple years of the program, the cut-off levels were slightly different. These differences are incorporated into the empirical work below.

parameter α_k measures the local average treatment effect on the probability of enrollment of an increase in the EMA amount that an individual is eligible to receive from 10 × (*k*-1) to 10 × *k, k=1,2,3.* If near the cut-off points individuals are approximately randomly distributed on either sides of the cut-off point, adding additional regressors to the model should not substantially change the parameter estimates α_k.[3] Nonetheless, we re-estimate the model with regressors to provide a robustness check (see Appendix 3.A, Table 3.A2); we find that adding these regressors does not substantively alter our estimates. We also check whether individual households are manipulating their household income, i.e., setting it just below one of the cut-off points, by investigating the distribution of household income near these cut-off points to see if there is any evidence of discontinuities (McCrary, 2008). Non-parametric methods that employ local linear regression (Fan and Gijbels, 1996) are used to further check the robustness of our empirical results. For these estimates we use the Epanechnikov kernel and compute robust z-statistics according to the methodology recommended by Calonico et al. (2014).

Data

The data for this study comes from the Family Resources Survey (FRS). Sponsored by the government's Department for Work and Pensions, the FRS consists of annual cross-sectional surveys of about 25,000 households (per year) across the United Kingdom. Households are selected from the Royal Mail's database of addresses that receive fewer than 500 items per day and are not labelled as organizations. The primary sampling units are post-code sectors, of which there are 1,848; each sector's probability of selection is weighted in proportion to its size. In addition, sampling units are stratified according to region, average income, and average cost of living. Interviews are conducted with all household residents

3 The precision of the estimates, however, should increase.

Table 3.1: Descriptive statistics

Variable	(1) All	(2) Within £2500 below £30 cut-off	(3) Within £2500 above £30 cut-off	(4) Within £2500 below £20 cut-off	(5) Within £2500 above £20 cut-off	(6) Within £2500 below £10 cut-off	(7) Within £2500 above £10 cut-off
Male	0.505	0.492	0.479	0.470	0.534	0.533	0.557
Age							
16	0.428	0.416	0.447	0.360	0.457	0.423	0.385
17	0.340	0.395	0.337	0.375	0.290	0.330	0.354
Parent has university degree	0.212	0.098	0.130	0.161	0.145	0.192	0.163
Year							
2005	0.133	0.165	0.145	0.114	0.139	0.140	0.157
2006	0.165	0.125	0.148	0.189	0.191	0.176	0.160
2007	0.158	0.162	0.145	0.155	0.164	0.154	0.145
2008	0.165	0.159	0.183	0.186	0.157	0.118	0.191
2009	0.156	0.153	0.172	0.136	0.142	0.137	0.111
2010	15.280	0.159	0.130	0.164	0.142	0.184	0.169
Region							
North West & Merseyside	0.155	0.190	0.154	0.219	0.151	0.151	0.132
Yorkshire & Humberside	0.103	0.116	0.089	0.120	0.073	0.115	0.086
East Midlands	0.094	0.107	0.124	0.068	0.110	0.085	0.139
West Midlands	0.110	0.122	0.130	0.102	0.123	0.115	0.154
Eastern	0.112	0.089	0.148	0.117	0.155	0.129	0.074
London	0.123	0.107	0.098	0.105	0.098	0.085	0.114
South East	0.163	0.144	0.139	0.139	0.164	0.165	0.151
South West	0.088	0.064	0.074	0.071	0.088	0.107	0.105
Sample size	9403	327	338	324	317	364	325
p-value of equality test			0.667		0.057		0.112

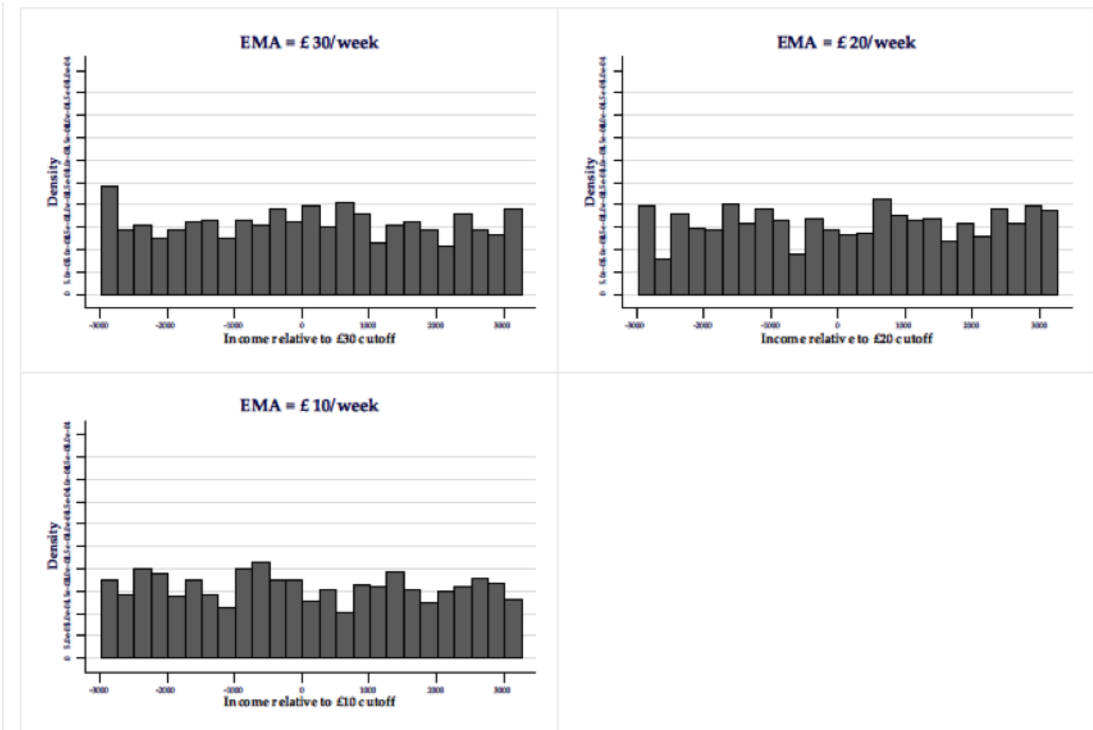
Source: Family Resource Surveys: 2004-2010 aged 16 or above, on a face-to-face basis, and then cross-validated against official records. In addition, the FRS collects detailed information on the educational enrollment status of all household members, their age, and whether they are receiving EMA. Further details on the FRS' survey methodology are available from the Department for Work and Pensions (2012).

We exclude households from Northern Ireland, Scotland and Wales. We also limit our sample to those who are 16-18 years old at the time of the survey. Moreover, we limit our data to the survey years 2004/2005 – 2010/2011 since the EMA began nationwide in England in 2004 and stopped accepting applications in January 2011, although those who applied before that date continued receiving payments for the remainder of the school year. The final sample size is 9,403. Summary statistics are provided in column (1) of Table 3.1.

The FRS collects a detailed breakdown of income data from all household members, as reported to Her Majesty's Revenue and Customs for tax purposes. This allows us to construct measures matching the exact criteria under which eligibility for EMA is assessed. Using household identifiers, we merge on information about parental income (including step-parents) to these youths and determine family income as the sum of parental (and step-parent) net income, where parental net income equals their gross income minus income from tax credits, disability benefits and other government provided benefits such as jobseekers allowances, maternity allowances and incapacity benefits. Income data is reported weekly to the closest pound; we convert this into a yearly amount and, because of data heaping and possible measurement error, we further group the income data into £250 bins.

In order to determine whether there is any empirical evidence of household manipulation around the EMA cut-off points, Figure 3.1 displays the histogram of household income around each of the three cut-off points. As can be seen from the graphs, there is no evidence of unusual jumps in the histogram just below the cut-off points. As a further robustness check, in columns (2) – (7) of Table 3.1 we report the results of tests to judge whether there is any empirical evidence that the distribution of a set of categorical variables differ between those just above and just below the cut-off points.[4] As indicated by the p-values reported at the bottom of the table, we do not reject the null hypothesis of no differences at the 5% significance level, although for the £20 cut-off we can reject the null hypothesis at the 10% level (p-value = 0.057).

Figure 3.1: Distribution of income



Source: Family resources Survey 2004–2010.

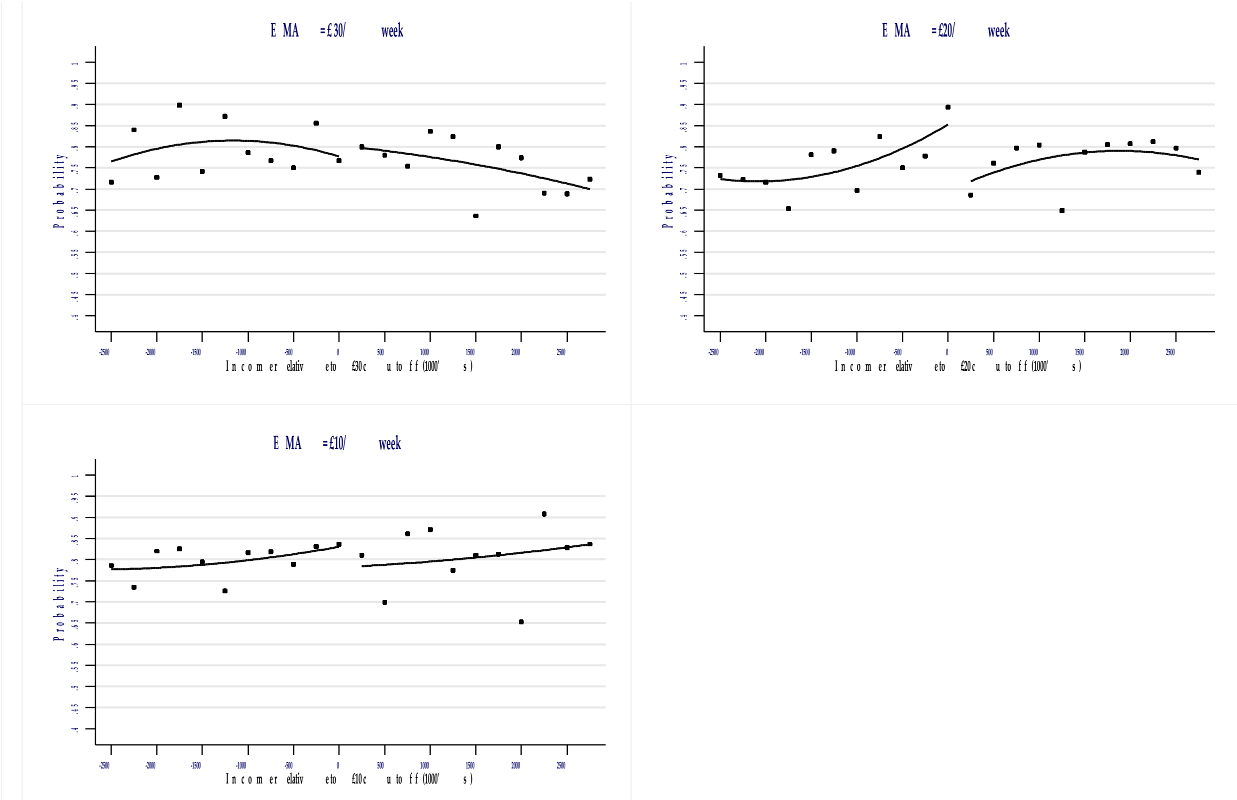
Results

Of the potential methods for calculating household income, here we present results when calculating taxable household income on the basis of both biological and step-parents' income. Appendix 3.A presents alternative estimates, the first of which instead calculates taxable income solely on the basis of biological parents' income (Table 3.A1). In addition, we

4 We define those just above to be no more than £2500 above the cut-off point and those just below to be no more than £2500 below the cut-off point.

check the sensitivity of our eligibility measure by adding model controls for gender, age, parental education, region, and the month and year in which the survey was conducted (Table 3.A2). Neither substantially alters the results of the model estimates. We also estimate all models with non-parametric local linear regressions (Table 3.A3). These results provide point estimates of a similar magnitude, albeit with more conservative levels of statistical significance.

Figure 3.2: Probability of school enrollment by income



Source: Family resources Survey 2004–2010.

In order to test the robustness of our results, we estimate models with varying interval sizes and polynomial orders. The interval sizes are £1,250, £1,750, £2,500, and £3,750 either side of each eligibility cut-off. We re-estimate each of these models allowing for the running variable to have a linear, quadratic and then cubic effect on the dependent variable, enrollment.[5]

Table 3.2: RDD estimates of the effect of EMA eligibility on enrollment

Interval	(1) Linear		(2) Quadratic		(3) Cubic	
£30 cut-off point						
±3750	**0.004**		0.004		-0.008	
	(0.054)		(0.050)		(0.054)	
±2500	**0.007**		0.009		-0.049	
	(0.054)		(0.059)		(0.059)	
±1750	**-0.024**		-0.024		0.020	
	(0.050)		(0.067)		(0.065)	
±1250	-0.057		-0.056		**0.142**	*
	(0.063)		(0.078)		**(0.086)**	
£20 cut-off point						
±3750	-0.029		-0.029		**0.105**	
	(0.064)		(0.064)		**(0.072)**	
±2500	0.044		0.046		**0.153**	*
	(0.072)		(0.071)		**(0.080)**	
±1750	**0.121**	*	0.122	*	0.186	**
	(0.073)		(0.072)		(0.095)	
±1250	**0.161**	*	0.161	*	0.254	**
	(0.084)		(0.082)		(0.104)	
£10 cut-off point						
±3750	**-0.006**		-0.007		0.036	
	(0.044)		(0.045)		(0.053)	
±2500	**0.024**		0.026		0.094	*
	(0.049)		(0.050)		(0.057)	
±1750	0.054		0.049		**0.149**	**
	(0.057)		(0.056)		**(0.063)**	
±1250	0.115	*	**0.110**	**	0.114	**
	(0.059)		**(0.053)**		(0.058)	

Notes: Dependent variable whether or not an individual is enrolled in full-time education. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level, respectively. Source: Family Resource Survey 2004-2010. Estimates presented in bold are those which we believe to be the most reliable at each interval level, as per the rationale laid out in footnote 5.

5 Given that we present multiple estimates for each cut-off point, we thought it worth providing some guidance as to which estimates we consider to be the most reliable. The preferred model is the linear model if the estimated coefficient for the quadratic is not significant and the joint test of the coefficients for the quadratic and cubic terms are not significant at 10% level. The preferred model is quadratic if the coefficient for the quadratic term is significant but the coefficient for the cubic term is not significant at 10% level. The cubic model is preferred if both the coefficient for the cubic term is significant and the joint test of quadratic and cubic coefficients is significant at 10% level. In Table 3.2, we use bold text to indicate which estimate we consider to be the most reliable estimate for each interval level.

Figure 3.2 provides a visual representation of the pattern of enrollment behavior around each of the cut-offs, allowing for the running variable to have cubic effects. Table 3.2 presents regression-discontinuity estimates for all models estimated on the full sample, without controls. Both provide limited evidence of the £30 EMA payments having any impact on enrollment. Whereas the model estimate with a £1,250 interval and a cubic function is significant at the 10 percent level, this is not supported by the other estimates for this cut-off since as many are negative as are positive, and for each the standard error exceeds the point estimate.

In contrast, the visible discontinuities that are apparent in the figures around the £20 and £10 cut-offs are borne out by the model estimates: it appears that both the £20 and £10 EMA payments have positive impact on net enrollment. For the £20 payment, seven of the 12 estimates are significant, and each of these was at one of the two lower interval levels. Also, the size of the estimates increases as the interval size decreases. Even the lower bound of the 95% confidence interval is approximately five percentage points when using a model with a £1,250 interval and a cubic function. However, it is worth noting that the other significant estimates are more conservative, with lower bounds closer to one percentage point. There is also evidence that the £10 payment influenced enrollment behavior. As with estimates for the £20 cut-off, estimates are significant for all models with the interval of £1,250. Significant point estimates range from nine to 15 percentage points, although the lower bound for these confidence intervals are around one percentage point.

We also test for heterogeneous effects of EMA payments according to recipients' region, gender, age, and parental education. For region, we combine youth in the north and Midlands into a single sample due to restricted sample sizes,. Model estimates for heterogeneous effects by region are presented in Table 3.3. Columns (1) – (3) present model estimates for youth in the south of England, and columns (4) – (6) present model estimates for youth in the north and the Midlands. The major notable difference between the regions is

Table 3.3: RDD estimates for heterogeneous effects by region

	South			North and Midlands		
Interval	(1) Linear	(2) Quadratic	(3) Cubic	(4) Linear	(5) Quadratic	(6) Cubic
			£30 cut-off point			
±3750	-0.046	-0.048	-0.092	0.045	0.038	0.059
	(0.062)	(0.071)	(0.079)	(0.069)	(0.069)	(0.064)
±2500	-0.068	-0.067	-0.110	0.065	0.066	-0.004
	(0.073)	(0.087)	(0.086)	(0.067)	(0.080)	(0.076)
±1750	-0.073	-0.068	0.011	0.014	0.015	0.038
	(0.074)	(0.100)	(0.078)	(0.060)	(0.090)	(0.075)
±1250	-0.064	-0.054	0.118	-0.044	-0.047	0.170*
	(0.070)	(0.115)	(0.102)	(0.071)	(0.107)	(0.098)
			£20 cut-off point			
±3750	0.056	0.060	0.193*	-0.105	-0.105	0.041
	(0.089)	(0.082)	(0.101)	(0.088)	(0.089)	(0.114)
±2500	0.151	0.153	0.184	-0.045	-0.028	0.142
	(0.100)	(0.097)	(0.121)	(0.116)	(0.106)	(0.128)
±1750	0.193*	0.195*	0.223*	0.073	0.083	0.168
	(0.112)	(0.109)	(0.134)	(0.122)	(0.108)	(0.149)
±1250	0.183	0.187	0.394***	0.160	0.165	0.162
	(0.124)	(0.116)	(0.150)	(0.137)	(0.116)	(0.156)
			£10 cut-off point			
±3750	-0.030	-0.040	-0.051	0.009	0.010	0.091
	(0.047)	(0.038)	(0.054)	(0.069)	(0.068)	(0.079)
±2500	-0.031	-0.046	0.006	0.069	0.069	0.111
	(0.054)	(0.045)	(0.056)	(0.074)	(0.075)	(0.098)
±1750	-0.041	-0.030	-0.009	0.092	0.087	0.227**
	(0.049)	(0.054)	(0.079)	(0.087)	(0.079)	(0.104)
±1250	-0.008	-0.029	-0.081*	0.165*	0.172**	0.184*
	(0.062)	(0.048)	(0.058)	(0.091)	(0.080)	(0.111)

Notes: Dependent variable whether or not an individual is enrolled in full-time education. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level, respectively. Source: Family Resource Survey 2004-2010.

that there is consistent evidence that the £10 payment has a positive impact on enrollment among youth in the north and the Midlands, since four estimates were positive and significant at the 10% level, including all three estimates with the narrowest interval. In contrast, there is not comparable evidence of the £10 payment having a significant impact among youth in the south. This difference between the regions was statistically significant at the .01 level.

Columns (1) – (6) of Table 3.4 present separate model estimates for females and males, respectively. On the basis of these estimates, it would appear that the impact of EMA payments varies according to gender. For females, there is scant evidence that EMA payments influence their enrollment behavior at any payment level. In contrast, model estimates are consistently significant for males at the £10 and £20 payment levels when using lower interval values, and of a similar magnitude to those for the whole group. However, statistical tests for differences between the groups were not significant.

For age, columns (7) – (12) of Table 3.4 present findings for 17- and 18-year-olds, respectively.[6] There are more consistent positive results for 17 year-olds than for 18 year-olds, at both the £20 and £10 levels. As with the models for the full sample and by gender, for both age groups there is little if any evidence that the increase from the £20 to the £30 payment has a significant impact on enrollment behavior. Statistical tests for differences between the groups were only significant in one instance, namely for the £10 payment when allowing for intervals of £1,250 on either side of the cut-off, in which case the payment had a higher impact on 18 year-olds.

6 Estimates for 16 year-olds are presented in Table 3.A4 of Appendix 3.A. These provided insignificant model estimates, suggesting a difference from 17- and 18-year-olds. However, the absence of birth month makes it impossible for us to distinguish between 16 year-olds who are and are not still in compulsory education (tenth grade).

Table 3.4: RDD estimates for heterogeneous effects by gender and age

	Females			Males			Age 17			Age 18		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Interval	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic	Linear	Quadratic	Cubic
£30 cut-off point												
±3750	0.00	0.02**	-0.05	0.01	0.01	-0.02	-0.05	-0.05	-0.08	0.02	0.02	0.14
	(0.06)	(0.08)	(0.09)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.11)	(0.14)	(0.14)
±2500	0.02	0.00	-0.05	0.00	0.00	-0.05	-0.07	-0.07	-0.123	0.12	0.14	0.18
	(0.07)	(0.09)	(0.09)	(0.07)	(0.09)	(0.09)	(0.08)	(0.10)	(0.09)	(0.14)	(0.16)	(0.14)
±1750	0.01	0.01	-0.02	-0.06	-0.06	0.07	-0.12	-0.12	-0.01	0.15	0.16	0.23
	(0.06)	(0.08)	(0.09)	(0.09)	(0.11)	(0.09)	(0.078)	(0.11)	(0.10)	(0.13)	(0.19)	(0.17)
±1250	-0.04	-0.04	0.06	-0.07	-0.07	0.22*	-0.10	-0.10	0.20	0.14	0.15	0.12
	(0.07)	(0.10)	(0.12)	(0.10)	(0.12)	(0.11)	(0.11)	(0.13)	(0.13)	(0.12)	(0.22)	(0.19)
£20 cut-off point												
±3750	-0.07	-0.06	0.11	0.00	0.00	0.09	0.01	0.01	0.21*	0.02	0.02	0.27***
	(0.08)	(0.08)	(0.11)	(0.07)	(0.07)	(0.09)	(0.12)	(0.11)	(0.11)	(0.10)	(0.10)	(0.10)
±2500	0.07	0.07	0.19**	0.02	0.02	0.19**	0.15	0.15	0.28**	0.18*	0.18*	0.27**
	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	(0.11)	(0.11)	(0.12)	(0.10)	(0.10)	(0.12)
±1750	0.12	0.12	0.11	0.13	0.13	0.25***	0.22*	0.22**	0.35**	0.24**	0.24**	0.10
	(0.10)	(0.09)	(0.15)	(0.08)	(0.09)	(0.08)	(0.12)	(0.11)	(0.15)	(0.11)	(0.12)	(0.18)
±1250	0.11	0.11	0.23	0.22***	0.22**	0.26***	0.34***	0.33***	0.35**	0.17	0.15	0.10
	(0.12)	(0.11)	(0.17)	(0.08)	(0.08)	(0.09)	(0.13)	(0.12)	(0.16)	(0.13)	(0.12)	(0.18)
£10 cut-off point												
±3750	-0.05	-0.05	0.01	0.03	0.03	0.05	-0.08	-0.08	0.05	-0.08	-0.08	0.05
	(0.07)	(0.07)	(0.08)	(0.05)	(0.05)	(0.07)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
±2500	0.01	0.02	0.14	0.04	0.04	0.06	0.02	0.023	0.12*	0.07	0.08	0.26***
	(0.08)	(0.08)	(0.09)	(0.07)	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)	(0.16)	(0.17)	(0.06)
±1750	0.09	0.09	0.14	0.03	0.03	0.15***	0.08	0.08	0.13	0.11	0.13	0.46**
	(0.08)	(0.08)	(0.12)	(0.07)	(0.07)	(0.05)	(0.07)	(0.07)	(0.08)	(0.19)	(0.19)	(0.22)
±1250	0.12	0.12	0.09	0.11**	0.11	0.14**	0.08	0.08	0.23***	0.32*	0.30*	0.32
	(0.10)	(0.09)	(0.13)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.18)	(0.17)	(0.23)

Notes: Dependent variable whether or not an individual is enrolled in full-time education. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level, respectively. Source: Family Resource Survey 2004-2010.

In addition, we tested for differences according to parental education, the results of which are presented in Tables 3.A5 and 3.A6 of Appendix 3.A. Although a greater number of model estimates are significant among youths without a degree-educated parent, just one of these estimates proved significantly different between the groups: the £10 payment when allowing for intervals of £1,250 on either side of the cut-off, in which case the payment had a higher impact on youths without a degree-educated parent. However, it is worth noting that very few youths in the sample had both a degree-educated parent and a sufficiently low household income to be eligible for the EMA, and so we choose not to put too much stock in this finding.

Discussion and conclusions

To the best of our knowledge, this is the first large-scale inferential analysis of the EMA during its national rollout or, for that matter, of any conditional cash transfer carried out in the entirety of a high-income country. Our findings are all the more compelling because we use nationally representative data collected during the program's full implementation. This advances prior literature on the EMA's pilot phase because it enables a more geographically and temporally expansive view of the program's impact. In addition, our use of a regression-discontinuity design is likely to provide the most robust casual estimates in the absence of a randomized controlled trial.

Our findings indicate that the EMA had a positive impact on enrollment in education among eligible youths, thus corroborating claims by the most robust prior research on the policy, by Dearden et al. (2009) and Middleton et al. (2003). While the size of some confidence intervals means that we do not attach too much weight to the exact estimates provided by the models, the consistency of significant non-zero results provides strong grounds to reject the null hypothesis of no impact of the EMA on enrollment. These findings verify a core concept of human capital theory: that youth evaluate short-term financial costs when making enrollment decisions and are responsive to reductions in these costs.

Of our tests for heterogeneous treatment effects, perhaps the most interesting is the apparent regional difference for the £10 payment: it seems to have a larger impact on enrollment among youth in the North and the Midlands than among youth in the South. This difference is logical; the cost of living and earnings are typically lower in the North and the Midlands, making the relative value of the payment was greater for youth in those regions than for youth in the South. This suggests that youth are sensitive to the value of conditional cash transfers, and calls into question whether the EMA might have been more effective had its value not been held constant for its duration.

Unfortunately, we are unable to test whether the EMA affected achievement as well as enrollment. As far as we know, the Family Resources Survey is the only data source with a sufficient number of respondents and detailed information on earnings to make an RD approach feasible. However, its cross-sectional nature means that it is not possible to track youth around the eligibility thresholds to see whether the EMA also influenced their achievement and receipt of qualifications. Since our findings confirm that increases in enrollment during the pilot extended to the national rollout, it is plausible that this increase was reflected in the increases in achievement identified by Chowdry et al. (2007). Unfortunately though, the absence of appropriate data means that we do not think it is possible to test this empirically.

Nonetheless, future research could extend this paper's analysis in several ways. More qualitative approaches could explore the mechanisms by which the EMA influences enrollment behavior. For example, it would be useful to policymakers to know whether knowledge of EMA in early years of schooling expanded the program's impact, or how the application and payment processes might influence take-up. In addition, both qualitative and quantitative approaches might further investigate whether the EMA's impact is greater for especially disadvantaged subgroups of the population, such as youths with disabilities. With a longitudinal approach, future research could extend our understanding of the EMA's effect on future outcomes, such as enrollment in higher education or employment. Such analyses of outcomes in later years would improve our ability to make meaningful assessments of the cost-effectiveness of the EMA.

It is important though to reiterate that the regression-discontinuity design provides local area treatment effects. Since we can only focus on youths close to the eligibility thresholds, it is not possible to make claims about the EMA's impact on the most financially disadvantaged youths. The incumbent coalition government's replacement to the EMA – the 16–19 bursary fund – focuses on youths below the income threshold for the £30 EMA payment. Unfortunately, our analysis strategy does not make it possible to estimate the EMA's impact on equivalent youths in order to offer a point of comparison with any future research on the replacement bursary policy.

Nonetheless, our findings also add to current debates on education policy and access in England. As Britain nears a national election, tuition fee levels have attracted much of the attention in education policy. Yet, although reducing the direct costs of university would appear to be a pro-access measure, it may represent a greater de facto subsidy to the middle classes than more needs-targeted programs such as the EMA. Past research (for example, Anders, 2012; Chowdry et al., 2013; Galindo-Rueda, Marcenaro-Gutierrez, & Vignoles, 2004) provides compelling evidence that disparities do not occur at the point of making applications, but years before. Access policies must thus focus on the years preceding university.

It remains unclear whether the raising of the school-leaving age and the 16–19 bursary fund may help students progress to higher education. The first of these policies only mandates a level of education or training far lower than the type of full-time education necessary to progress to higher education. The second transfers a far lower total sum to students than the EMA did, and a lack of clarity over its conditions may lessen the extent to which it influences students' post-16 decisions. However, this is speculation; to date, there is not yet research on whether either policy has widened access to post-compulsory education. In the meantime, our findings suggest that the English government has abolished a policy that is accurate in targeting, and effective in supporting, poorer students to progress in full-time, post-compulsory education.

References

Anders, J. (2012a). The Link between Household Income, University Applications and University Attendance. *Fiscal Studies*, *33*(2), 185–210.

Archer, L., Hutchings, M., & Ross, A. (2003). *Higher education and social class: Issues of exclusion and inclusion*. London: Routledge Falmer.

Ashworth, K., Hardman, J., Liu, W. C., Maguire, S., Middleton, S., Dearden, L., & Meghir, C. (2001). *Education Maintenance Allowance: the first year: a quantitative evaluation*. Research report RR257. London: Department for Education and Employment.

Attanasio, O., Fitzsimons, E., Gomez, A., Gutierrez, M. I., Meghir, C., & Mesnard, A. (2010). Children's schooling and work in the presence of a conditional cash transfer program in rural Colombia. *Economic development and cultural change*, *58*(2), 181-210.

Attanasio, O. P., Meghir, C., & Santiago, A. (2012). Education choices in Mexico: using a structural model and a randomized experiment to evaluate Progresa. *The Review of Economic Studies*, *79*(1), 37-66.

Ball, S. J., Reay, D. and David, M. (2002). "'Ethnic choosing': minority ethnic students, social class and higher education choice." *Race, ethnicity and education*, 5(4): 333-357.

Barr, N. (2009). Financing higher education: lessons from economic theory and reform in England. *Higher education in Europe*, *34*(2), 201-209.

Barr, N. (2010). *Paying for higher education: What policies, in what order?* (Submission to the Independent Review of Higher Education Funding and Student Finance). Retrieved from http://econ.lse.ac.uk/staff/nb/Barr_HEReview100215.pdfBecker, 1980;

Barr, N. (2011). *Breaking the logjam* (Submission to the Business, Innovation and Skills Committee, The Future of Higher Education, Session 2011-12, HC885). Retrieved from http://econ.lse.ac.uk/staff/nb/Barr_BISSelectComm110522.pdf

Becker, G. S. (1967). *Human capital and the personal distribution of income*. Ann Arbor, MI: Institute of Public Administration.

Becker, G. S. (1976). *The economic approach to human behavior*. Chicago, IL: The University of Chicago Press.

Becker, G. S. (1980). *Human capital: A theoretical and empirical analysis, with special reference to education* (2nd ed.). Chicago, IL: The University of Chicago Press.

Becker, G. S., & Tomes, N. (1986). Human capital and the rise and fall of families, *Journal of Labor Economics*, *4*(3), S1–S39.

Blanden, J., & Machin, S. (2004). Educational inequality and the expansion of UK higher education. *Scottish Journal of Political Economy*, *51*(2), 230–249.

Blundell, R., Dearden, L., Goodman, A., & Reed, H. (2000). The returns to higher education in Britain: evidence from a British cohort. *The Economic Journal*, *110*(461), 82-99.

Blundell, R., Dearden, L., & Sianesi, B. (2005). Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *168*(3), 473-512.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression‐Discontinuity Designs. *Econometrica*,*82*(6),

2295-2326.
Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics*, 3, 1801-1863.
Card, D. (2001). Estimating the return to schooling: progress on some persistent econometric problems. *Econometrica*, 69(5), 1127-1160.
Carneiro, P., & Heckman, J. J. (2002). The evidence on credit constraints in post‐secondary schooling. *The Economic Journal, 112*(482), 705-734.
Carneiro, P. and Heckman, J. J. (2003) Human capital policy. In J.J. Heckman, A. Krueger, & B. Friedman (Eds.), *Inequality in America: What role for human capital policies?* (77–239). Cambridge, MA: MIT Press.
Checchi, D. (2006). *The economics of education: Human capital, family background and inequality*. Cambridge: Cambridge University Press.
Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2013). Widening participation in higher education: analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(2), 431-457.
Chowdry, H., Dearden, L., & Emmerson, C. (2007). *Education Maintenance Allowance: Evaluation with Administrative Data The impact of the EMA pilots on participation and attainment in post-compulsory education*. London: Institute for Fiscal Studies.
Dearden, L. (1999). The effects of families and ability on men's education and earnings in Britain. *Labour Economics*, *6*(4), 551-567.
Dearden, L., Emmerson, C., Frayne, C., & Meghir, C. (2009). Conditional cash transfers and school dropout rates. *Journal of Human Resources*, *44*(4), 827-857.
Dearden, L., & Heath, A. (1996). Income support and staying in school: what can we learn from Australia's AUSTUDY experiment?. *Fiscal Studies*, *17*(4), 1-30.
Dearden, L., Machin, S., & Vignoles, A. (2011). The contribution of the economics of education to education. In J. Furlong & M. Lawn (Eds.), *Disciplines of education: Their role in the future of education research* (pp. 85–102). London: Routledge.
Dearden, L., McIntosh, S., Myck, M., & Vignoles, A. (2002). The returns to academic and vocational qualifications in Britain. *Bulletin of economic research*, *54*(3), 249-274.
Department for Education (2011). *Plans to end the Education Maintenance Allowance (EMA) programme*. Retrieved from https://www.gov.uk/government/news/plans-to-end-the-education-maintenance-allowance-ema-programme
Department for Work and Pensions (2012). Family Resources Survey. Retrieved from https://www.gov.uk/government/uploads/.../frs_2010_11_report.pdf
DiNardo, J., & Lee, D. S. (2011). Program evaluation and research designs. *Handbook of labor economics*, *4*, 463-536.
Dolton, P., & Vignoles, A. (2000). The incidence and effects of overeducation in the UK graduate labour market. *Economics of Education Review*, *19*(2), 179-198.
Fan, J., & Gijbels, I. (1996). *Local polynomial modelling and its applications: monographs on statistics and applied probability*. New York, NY: Chapman and Hall.
Friedman, M. (1955). The role of government in education. In R. Solo (Ed.), *Economics and the public interest* (pp. 123-44). New Brunswick, NJ: Rutgers

University Press.
Galindo-Rueda, F., Marcenaro-Gutierrez, O., & Vignoles, A. (2004). The widening socio-economic gap in UK higher education. *National Institute Economic Review*, *190*(1), 75-88.
Glennerster, H. (1991). Quasi-markets for education? *The economic journal*, *101*(408), 1268- 1276.
Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression–discontinuity design. *Econometrica*,*69*(1), 201-209.
Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). Earnings functions, rates of return and treatment effects: the Mincer equation and beyond. *Handbook of the Economics of Education*, *1*, 307-458.
Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: a guide to practice. *Journal of Econometrics*, *142*(2), 615-635.
Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355.
Legard, R., Woodfield, K., & White, C. (2001). *Staying away or staying on?: a qualitative evaluation of the Education Maintenance Allowance.* Research report RR256. London: Department for Education and Employment.
Machin, S. (2008). The new economics of education: methods, evidence and policy. *Journal of Population Economics*, *21*(1), 1-19.
Marcenaro-Gutierrez, O., Galindo-Rueda, F., & Vignoles, A. (2007). Who actually goes to university? *The Economics of Education and Training, 32*(2), 333–357.
McCall, B. P., & Bielby, R. M. (2012). Regression discontinuity design: recent developments and a guide to practice for researchers in higher education. In Smart, J.C, & Paulsen, M. B. (Eds.), *Higher education: Handbook of theory and research* (pp. 249-290). Springer Netherlands.
McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698-714.
Middleton, S., Maguire, S., Ashworth, K., Legge, K., Allen, T., & Perren, K. (2003). *The evaluation of Education Maintenance Allowance Pilots: three years evidence: a quantitative evaluation*. Retrieved from https://dspace.lboro.ac.uk/dspace/bitstream/2134/2477/1/RR499.pdf
Mincer, J. (1974). *Schooling, Experience and Earnings*. New York, NY: Columbia University Press.
Morris, P., Aber, J. L., Wolf, S., & Berg, J. (2012). Using Incentives to Change How Teenagers Spend Their Time: The Effects of New York City's Conditional Cash Transfer Program. *MDRC*.
National Foundation for Educational Research (2011). Education committee new inquiry: 16-19 participation in education and training. Submission of evidence by the National Foundation for Educational Research. Retrieved from: https://www.nfer.ac.uk/nfer/publications/99911/99911.pdf
Neyman, J. S. (1923). Statistical problems in agricultural experimentation. *Journal of the Royal Statistical Society*, Series B, *2*(2), 107–180.
Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.
Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, *66*(5), 688–701.

Schultz, T. (1961). Investment in human capital. *The American Economic Review 51*(1), 1–17.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
Smith, A. (2007). *An inquiry into the nature and causes of the wealth of nations*. Petersfield: Harriman House. (Original work published 1776.)
Spielhofer, T., Golden, S., Evans, K., Marshall, H., Mundy, E., Pomati, M., & Styles, B. (2010). *Barriers to participation in education and training*. London: Department for Education. Retrieved from http://dera.ioe.ac.uk/82/1/DFE-RR009.pdf
Universities and Colleges Admissions Service (2013). *Applicant and accepted applicant regional analysis, home and overseas*. Retrieved from http://www.ucas.com/about_us/stat_services/stats_online/data_tables/residence/ukregion
Vignoles, A., & Powdthavee, N. (2009). The socioeconomic gap in university dropouts. *The B.E. Journal of Economic Analysis & Policy*, *9*(1).
Walker, I., & Zhu, Y. (2011). Differences by degree: evidence of the net financial rates of return to undergraduate study for England and Wales. *Economics of Education Review*, 30(6), 1177-1186.

Appendix 3.A: Supplementary model estimates

Table 3.A1: RDD Estimates of the effect of EMA eligibility on enrollment, when excluding step-parental income from household income.

Interval	Linear (1)		Quadratic (2)		Cubic (3)	
			£30 cutoff point			
±3750	0.009		0.009		0.043	
	(0.054)		(0.049)		(0.056)	
±2500	0.044		0.045		0.009	
	(0.056)		(0.059)		(0.061)	
±1750	0.017		0.016		0.115	*
	(0.057)		(0.067)		(0.069)	
±1250	0.016		0.017		0.228	***
	(0.065)		(0.079)		(0.088)	
			£20 cutoff point			
±3750	-0.042		-0.042		0.052	
	(0.049)		(0.049)		(0.058)	
±2500	0.008		0.012		0.072	
	(0.055)		(0.051)		(0.060)	
±1750	0.058		0.062		0.085	
	(0.062)		(0.051)		(0.062)	
±1250	0.068		0.068		0.103	
	(0.062)		(0.060)		(0.072)	
			£10 cutoff point			
±3750	-0.027		-0.031		0.011	
	(0.045)		(0.044)		(0.052)	
±2500	-0.003		-0.006		0.080	*
	(0.050)		(0.049)		(0.047)	
±1750	0.047		0.037		0.124	**
	(0.053)		(0.053)		(0.049)	
±1250	0.083	*	0.080	*	0.146	***
	(0.047)		(0.047)		(0.056)	

Source: Family Resource Surveys: 2004-2010

Notes: Dependent variable whether or not an individual is enrolled in education. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level, respectively.

Table 3.A2: RDD Estimates of the effect of EMA eligibility on enrollment, with additional regressors included.

Interval	Linear (1)		Quadratic (2)		Cubic (3)	
£30 cut-off point						
±3750	0.002		0.002		-0.004	
	(0.051)		(0.047)		(0.054)	
±2500	0.004		0.006		-0.038	
	(0.053)		(0.056)		(0.050)	
±1750	-0.033		-0.032		0.055	
	(0.054)		(0.064)		(0.056)	
±1250	-0.026		-0.024		0.155	*
	(0.058)		(0.076)		(0.078)	
£20 cut-off point						
±3750	-0.001		-0.001		0.136	**
	(0.064)		(0.064)		(0.069)	
±2500	0.091		0.099		0.202	***
	(0.074)		(0.069)		(0.076)	
±1750	0.172	**	0.173	**	0.262	***
	(0.080)		(0.080)		(0.084)	
±1250	0.249	***	0.249	***	0.352	***
	(0.066)		(0.064)		(0.080)	
£10 cut-off point						
±3750	-0.029		-0.030		0.011	
	(0.039)		(0.039)		(0.045)	
±2500	0.008		0.011		0.062	
	(0.041)		(0.041)		(0.057)	
±1750	-0.006		-0.007		0.069	
	(0.064)		(0.062)		(0.071)	
±1250	0.054		0.051		0.083	
	(0.067)		(0.065)		(0.080)	

Source: Family Resource Surveys: 2004-2010

Notes: Dependent variable whether or not an individual is enrolled in full-time education. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level, respectively. Estimates include controls for gender, age, parent's education, region of the country, and month and year of survey

Table 3.A3: RDD Estimates of the effect of EMA eligibility on enrollment for the full sample, using local linear regression estimates.

	£30 cutoff point	£20 cutoff point		£10 cutoff point
Sample	(1)	(2)		(3)
All	-0.036	0.202	*	0.114
	(-0.462)	(1.690)		(1.107)
Males	0.084	0.205		0.118
	(0.695)	(1.002)		(0.882)
Females	-0.096	0.179		0.115
	(-0.651)	(1.257)		(0.147)
Aged 16	0.005	0.176		0.164
	(0.223)	(1.415)		(0.699)
Aged 17	0.016	0.254		-0.014
	(0.282)	(1.110)		(0.170)
Aged 18	-0.303	0.321		0.233
	(-0.898)	(0.789)		(0.692)
Parent with Uni Degree	-0.143	-		0.110
	(-1.226)			(0.122)
No Parent with Uni Degree	0.243	0.239	*	0.147
	(1.319)	(1.858)		(1.169)

Source: Family Resource Surveys: 2004-2010

Notes: Dependent variable whether or not an individual is enrolled in school. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level, respectively. Robust z-statistics are reported in parentheses. Estimates based on Epanechnikov kernel.

Table 3.A4: RDD Estimates of the effect of EMA eligibility on enrollment among 16 year-olds.

Interval	Linear (1)	Quadratic (2)	Cubic (3)	
	£30 cutoff point			
±3750	0.060	0.059	0.007	
	(0.056)	(0.055)	(0.069)	
±2500	0.032	0.031	-0.088	
	(0.066)	(0.061)	(0.091)	
±1750	-0.019	-0.020	-0.099	
	(0.076)	(0.073)	(0.113)	
±1250	-0.099	-0.097	0.017	
	(0.097)	(0.092)	(0.156)	
	£20 cutoff point			
±3750	-0.016	-0.017	-0.025	
	(0.060)	(0.061)	(0.085)	
±2500	-0.068	-0.066	0.002	
	(0.078)	(0.078)	(0.107)	
±1750	0.004	0.010	0.103	
	(0.086)	(0.087)	(0.103)	
±1250	0.044	0.023	0.177	**
	(0.107)	(0.099)	(0.087)	
	£10 cutoff point			
±3750	-0.007	-0.017	-0.015	
	(0.047)	(0.043)	(0.057)	
±2500	-0.027	-0.028	-0.037	
	(0.046)	(0.047)	(0.060)	
±1750	-0.038	-0.041	0.015	
	(0.054)	(0.053)	(0.069)	
±1250	-0.008	-0.012	-0.029	
	(0.070)	(0.059)	(0.076)	

Source: Family Resource Surveys: 2004-2010

Notes: Dependent variable whether or not an individual is enrolled in school. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level, respectively.

Table 3.A5: RDD Estimates of the effect of EMA eligibility on enrollment among those whose parents do not have a degree.

Interval	(1) Linear		(2) Quadratic		(3) Cubic	
			£30 cut-off point			
±3750	-0.002		-0.004		-0.028	
	(0.055)		(0.054)		(0.058)	
±2500	-0.013		-0.013		-0.068	
	(0.059)		(0.064)		(0.065)	
±1750	-0.044		-0.044		0.011	
	(0.055)		(0.073)		(0.078)	
±1250	-0.073		-0.071		0.127	
	(0.070)		(0.085)		(0.105)	
			£20 cut-off point			
±3750	-0.007		-0.007		0.130	
	(0.070)		(0.070)		(0.082)	
±2500	0.074		0.079		0.194	**
	(0.083)		(0.078)		(0.086)	
±1750	0.156	*	0.157	**	0.214	**
	(0.082)		(0.080)		(0.099)	
±1250	0.197	**	0.195	**	0.304	***
	(0.090)		(0.086)		(0.234)	
			£10 cut-off point			
±3750	0.009		0.008		0.031	
	(0.045)		(0.045)		(0.059)	
±2500	0.031		0.031		0.071	
	(0.053)		(0.053)		(0.064)	
±1750	0.041		0.039		0.147	**
	(0.064)		(0.062)		(0.061)	
±1250	0.098		0.096	*	0.174	***
	(0.062)		(0.055)		(0.061)	

Source: Family Resource Surveys: 2004-2010
Notes: Dependent variable whether or not an individual is enrolled in full-time education. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level, respectively.

Table 3.A6: RDD Estimates of the effect of EMA eligibility on enrollment among those whose parents have a degree.

Interval	Linear (1)		Quadratic (2)	Cubic (3)	
£30 cut-off point					
±3750	0.039		0.060	0.156	*
	(0.109)		(0.107)	(0.094)	
±2500	0.152		0.155	0.139	
	(0.081)		(0.109)	(0.087)	
±1750	0.153	*	0.145	0.142	
	(0.083)		(0.104)	(0.096)	
±1250	0.120		0.120	0.192	*
	(0.090)		(0.147)	(0.111)	
£20 cut-off point					
±3750	-0.198	*	-0.193	-0.025	
	(0.113)		(0.117)	(0.151)	
±2500	-0.116		-0.134	-0.161	
	(0.130)		(0.131)	(0.141)	
±1750	-0.096		-0.101	-0.134	
	(0.121)		(0.130)	(0.196)	
±1250	-0.116		-0.085	-0.088	
	(0.143)		(0.171)	(0.234)	
£10 cut-off point					
±3750	-0.121		-0.118	-0.031	
	(0.066)		(0.064)	(0.070)	
±2500	-0.060		-0.048	0.044	
	(0.071)		(0.063)	(0.114)	
±1750	0.028		-0.008	-0.023	
	(0.185)		(0.071)	(0.141)	
±1250	0.048		-0.001	-0.395	**
	(0.169)		(0.113)	(0.154)	

Source: Family Resource Surveys: 2004-2010
Notes: Dependent variable whether or not an individual is enrolled in full-time education. One, two, and three asterisks indicate statistical significance at the 10, 5, and 1 percent level, respectively.

Conclusion

Educational access continues to be a major policy concern in England. Inequalities remain entrenched: for all the efforts of practitioners, politicians and researchers, a student's background remains a strong determinant of her progression to higher education. Effective policymaking is crucial to redressing this, but effective policymaking is reliant on well-informed theory, grounded in empirical research.

My thesis comprises three distinct, but related, studies of influences on student behavior that may alter access inequalities. I hope that their collective impact will be to broaden the ways in which our research field investigates access. Too often, theoretical divisions also mirror methodological divisions. This is a problem because the complementary affordances of different methodologies are rarely applied to a given theorization or hypothesis. In an attempt to ameliorate this shortcoming, I apply inferential methods to assess theories that researchers in the UK have only studied with qualitative methods to date. I hope that this will encourage other researchers to be more flexible in how they pair theory and methods. In addition, I make original use of an administrative dataset – primarily focused on employment, earnings, and wealth – to evaluate the national impact of an education policy, when we had previously been reliant on data from that policy's pilot phase. I hope this will encourage other education researchers to consider how to make more innovative use of the rich data in England that do not necessarily focus on education. Such changes in thinking can expand the evidentiary base from which policy and research are developed.

In chapter 1, I assess the role of teacher encouragement to enroll in the final, non-compulsory years of high school on students' observed enrollment behavior, both in those final years and in undergraduate degree programs. Education researchers in England have invoked Bourdieu's theorization of habitus to claim that social interactions influence students' enrollment behavior by altering the extent to which they think further education is feasible and desirable, and that the effect of these interactions depends on students' social class. My findings corroborate both of these claims: it seems that teacher encouragement does change students' ensuing enrollment behavior, and that this impact is greatest for students from lower socio-economic groups. Furthermore, the fact that encouragement can have a positive effect supports researchers who have diverged from Bourdieu's work by claiming that teachers can challenge educational inequalities.

These findings offer robust quantitative evidence for a factor that, to the best of my knowledge, prior research has only tested with small-scale, cross-sectional, qualitative methods. Thus, the scale and longitudinal nature of the present inferential study may help to widen the audience who would countenance the role of teacher encouragement in students' progress to higher education. It may also expand debates on the role of teachers in supporting more disadvantaged student groups: where there is currently a narrow focus on teachers raising exam attainment among pupils, the relational aspects of their work might now receive greater attention too.

In chapter 2, I assess the impact of discrete grade labels from national eighth grade exams on students' enrollment in undergraduate degree programs five to seven years later. I find that grade labels lead to a bifurcation in achievement and enrollment behavior among students who are otherwise indistinguishable from one another. This finding supports theorizations from a range of conceptual traditions in psychology, economics, and sociology. In addition, effects appear to be greater among students from lower socio-economic groups, supporting the theorization that working-class students are particularly susceptible to the summative judgments of their ability that they receive in formal education.

This finding is especially worrying given that the purported goal of these examinations was to judge school quality rather than student progress. Although it could be argued that the labelling effect is as much a boon for those receiving the higher grade as it is a detriment to those receiving the lower grade, past sociological research emphasizes the latter. This paper's quasi-experimental design and longitudinal approach bolsters the research literature, and might encourage researchers and policymakers to think more carefully about how one conducts and provides feedback on the many instances of formal assessment that currently occur in the English education system. The case for reform is all the stronger given that these ostensibly low-stakes exams have negative consequences for historically underrepresented students.

Chapter 3 is more conventional in the sense that its theoretical and empirical approaches are commonly used in conjunction with one another in the research literature. Human capital theory suggests that conditional cash transfers will increase enrollment in education, as in students' decision-making they offset the short-term costs that risk outweighing the longer-term benefits of education. Robust inferential research already exists to suggest that the UK's primary conditional cash transfer policy, the Education Maintenance Allowance (EMA), was able to increase enrollment during its pilot stage. In this chapter, Brian McCall and I extend this research considerably by assessing the EMA's impact during its full implementation, i.e., at the national scale and for multiple student cohorts. We are able to do this by using administrative datasets whose scale and detailed income information permit robust regression-discontinuity estimates with nationally representative data.

While Northern Ireland, Scotland and Wales have maintained more generous variants of the Education Maintenance Allowance, England's current government has discontinued the program, claiming that is was ineffective and expensive. Our findings contradict the first of these claims. We find that the EMA increased full-time enrollment in post-compulsory education, thus corroborating both human capital theory and prior evaluations of the EMA during its pilot stage. However, we do not fully engage with the latter claim largely because any cost-benefit analysis should account for the true impact of this apparent boost in enrollment. Future research could aid this shortcoming by investigating the impact of the EMA on later educational qualifications and employment outcomes. If such work were to find a positive impact commensurate with our findings on enrollment, it would lend further support to the decision of the other governments of the United Kingdom to continue supporting their EMA programs.

As a set, these chapters demonstrate the multi-faceted nature of inequalities in access to higher education. By working from varied theoretical traditions, I provide evidence for a range of conceptualizations of behavior and influencing factors. My emphasis on socio-economic class disparities should not mask the fact that there are numerous sources of disadvantage that are likely to act in mutually dependent and re-enforcing ways. Yet, even when focusing on a single element of inequality, my findings demonstrate that many factors play a role in shaping disparities. Research ought to reflect this complexity. The emphasis here on quantitative inference should not be mistaken for empirical dogmatism: a breadth of methodological approaches is of the utmost importance to effective policy research. The entrenched nature of disadvantage means that researchers and policymakers face an uphill struggle in helping rectify inequalities. If we are to truly make inroads, policy researchers must take a methodologically expansive, diverse, and cooperative approach to their work.