Short and long term barriers to university education in England. What do expectation data tell us?

Laura Fumagalli
ISER, University of Essex

February 1, 2019

Abstract

Using data from the Longitudinal Study of Young People in England (LSYPE) and the National Pupil Database (NPD), this paper provides evidence on the role of short-term and long-term barriers in determining the SES gap in university enrollment. We focus on the SES gap in expectations at age 14-15 over future university application and admission. To shed light on the role of short-term barriers we look at whether short-term costs of education affect the perceived likelihood of applying to university. To do this, we use quasi-random variation in family's awareness of and access to available financial aid, namely the Educational Maintenance Allowance (EMA). To shed light on the role of long-term barriers we look at the determinants of perceived likelihood of being admitted to university. This likelihood is elicited conditional on university application, and thus it should not be affected by short-term costs of applying to university. We find evidence that both long-term and short-term barriers are in place.

1 Introduction

In many industrialized countries, students from well-off families are more likely to attend university than less affluent peers (Blossfeld and Shavit, 1993; Cameron and Heckman, 2001; Carneiro and Heckman, 2002; Blanden and Machin, 2004; Machin and Vignoles, 2004; Anders and Micklewright, 2013; Chowdry et al., 2013). Indeed, research suggests in the USA and in the UK the expansion of higher education (HE) has exacerbated -and not reduced- the socio-economic gap in university participation (Carneiro and Heckman, 2002; Blanden and Machin, 2004; Machin and Vignoles, 2004).

Low university enrollment among low socio-economic status (SES) students may be due to long-term and short-term barriers (see: Cameron and Heckman, 1999, 2001; Carneiro and Heckman, 2002). Long-term barriers stem from low 'permanent' family income. Low permanent family income -and thus resources available for education throughout children's formative years- translate into poor home learning environment. This leads students from less affluent families to develop lower cognitive and non-cognitive skills, and lower levels of HE readiness. Short-term barriers are the direct (e.g., tuition fees) and opportunity (e.g., income lost from delayed entry in the labour market) costs of staying in education. These short-term barriers can prevent HE ready low SES students from applying to university, especially in presence of credit constraints (Dearden et al., 2009; Deming and Dynarski, 2009; Chevalier et al., 2010; Kaufmann, 2010).

Using data from a nationally representative cohort study: the Longitudinal Study of Young People in England (LSYPE), as well as administrative data from the National Pupil Database (NPD), this paper provides evidence on the role of short-term and long-term barriers in determining the SES gap in university enrollment. We focus on the SES gap in expectations at age 14-15 over future university application and admission. To shed light on the role of short-term barriers we look at whether short-term costs of education affect the perceived likelihood of applying to university. To do this, we use quasi-random variation in family's awareness of and access to available financial aid, namely the Educational Maintenance Allowance (EMA). To shed light on the role of long-term barriers we look at the determinants of the perceived likelihood of being admitted to university. This likelihood is elicited conditional on university application, and thus it should not be affected by short-term costs of applying to university.

In the first part of the paper, we use the introduction of the Education Maintenance Allowance (EMA) to investigate the role of short-term barriers, in the form of short-term costs of education. The EMA was a mean tested weekly allowance introduced in the academic year 2004-2005 aimed at keeping low income people between the age of 16 and 19 in post-compulsory academic education until the age for university education. The EMA did not directly cover the costs of university. However, it was meant to compensate the costs of being in post-compulsory, non-university education. As completion of post-compulsory, non-university education is required to apply to university, the EMA had the potential of affecting expectations over university application, particularly among low income students.

We do not have variation in the access of EMA for LSYPE respondents, who were all around 14 on the first year of the EMA national roll-out. However, before its national introduction, the EMA was piloted in a number of Local Education Authorities (LEAs). This creates variation in i) knowledge of the policy among LSYPE respondents; ii) access to financial aid for family members older than LSYPE respondents. We hypothesize households in pilot areas, and particularly those with children older than the respondents, are more likely to both be aware of the EMA and to have used it to keep older children in education, possibly until university.

In the second part of the paper we look at the perceived likelihood of being admitted to university given a hypothetical application to investigate the role of long-term barriers (Cameron and Heckman, 1999, 2001; Carneiro and Heckman, 2002). The perceived likelihood of being admitted to university given a hypothetical application should not be affected by short-term barriers to education. Therefore, any difference in the perceived likelihood of success in the application between students from different SES should be due to constraints linked to family permanent income (e.g., due to low investment in cognitive and non-cognitive skills early in students' lives or to lack of information about requirements for university admission), rather than short-term costs and credit constraints.

The literature (e.g., Hoxby and Avery, 2012; Hoxby and Turner, 2014; Goodman, 2016) suggests low-SES high-achievers underestimate their suitability for selective colleges and generally apply to less prestigious institutions than richer peers with similar attainment. This is in line with results by Jacob and Wilder (2010) who find that low-SES students do not update their expectations on attainment at university based upon changes in GPA.

What explains this miss-alignment between academic performance and (expectations over) university enrollment for low SES students? There are various reasons. The miss-alignment may be due to short-term monetary barriers. For example, more prestigious universities may levy higher fees and may be located in more expensive cities than less prestigious ones (Hoxby and Avery, 2012), or low SES students may misunderstand the costs of college, and lack awareness about the available financial help (Jacob and Wilder, 2010). At least for the USA, short-term monetary barriers do not seem to be the only explanation: prestigious universities often offer more generous financial aid than less prestigious institutions (Hoxby and Avery, 2012) and increased generosity in merit programs had a limited effect on the number of students who expect to enroll in education (Jacob and Wilder, 2010). Alternatively, the miss-alignment may be due to long-term barriers. For example, low-SES people may not be aware of the requirements for university admission (see also: Hastings et al., 2007). By looking at a measure of expectation arguably unaffected by short-term monetary barriers, we provide evidence that the latter explanation may be relevant.

The LSYPE does not elicit information on the perceived likelihood of being admitted to university for respondents who declare they are 'not at all likely' to apply. This sample selection is potentially non-random. We address the problem by estimating a selection model where identification is achieved through functional form assumptions and exclusion restrictions: namely indicators of EMA pilot regions and the local unemployment rate.

The results on the effect of the EMA contribute to the literature on the effect of short-term barriers on human capital investment. In particular, they contribute to the literature on the effect of awareness (Dinkelman and Martínez, 2011; Bettinger et al., 2012) and availability (Dynarski, 2008; Dearden et al., 2009; Sjoquist and Winters, 2015; Castleman and Long, 2016; Barr, 2019) of financial support. Unlike most existing work, our paper relies on a nationally representative sample, focuses on financial aid offered to stay on in post-compulsory, non-university education, and looks at educational expectations, rather than enrollment or attainment.[1]

The results on the likelihood of being admitted to university conditional on application contribute the literature emphasizing the role of long-term barriers on human capital investment. In particular, this paper is the first attempt of analyzing expectations over university admission to use them to isolate the role of short-term costs from the role of log-term barriers.[2] Therefore, we contribute to the growing literature on education expectations (see, for example: Chowdry et al., 2009, 2010; Jacob and Wilder, 2010; Zafar, 2011; Stinebrickner and Stinebrickner, 2012).[3]

We find that living in a EMA pilot area increases the perceived likelihood of applying to university measured at age 14 (the first year of implementation of the EMA), especially for low SES students. We interpret this as evidence that being aware of mechanisms potentially alleviating short-term direct and opportunity costs of being in post-compulsory education boosts low SES students' expectations of university application. We also find students in EMA pilot areas with older siblings expect lower likelihood of applying to university than students in EMA pilot areas without older siblings. We speculate the EMA may have helped older siblings to access university. Coherently with research on order effects in education (e.g., Black et al., 2005, 2018; Price, 2008), this might have drained the resources available for higher education participation for younger siblings.

The positive impact of residence in EMA pilot areas on the perceived likelihood of applying to university disappears when we look at expectations at age 15, when the EMA was fully in

[1] Dearden et al. (2009) also focuses on pre-university financial aid and (Dinkelman and Martínez, 2011) also looks at educational expectations.

[2] For a more recent attempt of eliciting information on success at university conditional on application, see: Auspurg et al. (2013). The existing literature using LSYPE data (e.g., Chowdry et al., 2009, 2010) has addressed the problem of -potentially non random- missing information on the perceived likelihood of being admitted to university by combining in a single variable information on the perceived likelihood of applying to university and on the perceived likelihood of success of such application. This is not ideal, as policies aimed at widening school participation can only be successful if the determinants of the perceived chances of success are understood separately from other factors affecting the choice of university application.

[3] An early contribution on the importance of educational expectations for schooling choice is Manski (1993). Educational expectations have been found to be correlated with family socio-economic status (Reynolds and Pemberton, 2001; Duckworth et al., 2009; Chowdry et al., 2010; Jacob and Wilder, 2010; Fowley, 2011; Jerrim, 2011), perceived costs of education (Dinkelman and Martínez, 2011), and education wage premia (Carneiro and Heckman, 2002; Attanasio and Kaufmann, 2009; Jensen, 2010; Nguyen, 2010).

place in the whole country. The negative effects of having a older sibling for students in EMA pilot areas does not. This is coherent with the information advantage of living in EMA areas fades with time, while the shift of limited parental resources towards the older siblings does not.

Finally, we find the link between the perceived likelihood of being admitted to university and students' academic performance is weaker for low SES students, that is, low SES students with low (high) academic performance expect higher (lower) probability of being admitted to university than high SES students. We interpret this as evidence that the SES gap in university application is not due to short-term monetary barriers only, and that low SES students may also lack information on the requirements for university admission.

Distinguishing the role of short-term and long-term barriers in important to inform policies to reduce the SES gap in university participation. Our results on the EMA suggest students do anticipate short-term barriers to university participation. Anticipated barriers to university participation could lead to poor investment in human capital in teenage years, making it difficult to enter Higher Education later on (Chowdry et al., 2013). This is particularly salient in the case of English students aged 14 and 15 who are asked to choose the subjects they want to study in the future: a choice with long-term consequences on students' likelihood to be admitted to university (Dilnot, 2017). Therefore, removing anticipated short-term barriers can increase students' enrollment and reduce the SES gap in university participation.[4]

Our results on the perceived likelihood of being admitted to university conditional on a hypothetical application suggest the SES gap in university participation is also determined by long-term barrier. Therefore, it will not disappear if short-term barriers are removed. The link between the perceived likelihood of success in a university application and academic performance is found to be weaker for low SES students. This result can help design better policies to promote equal access to higher education. If disadvantaged students do not form

[4]While not the main objective of the paper, our results also provide an important contribution to the debate on the cost-effectiveness of the EMA. The EMA was abolished in 2010 as it was deemed too expensive. This decision was criticized, claiming it was only based on estimates of the impact of the policy on participation in post-compulsory education, and overlooked potential indirect benefits, such as increased effort at school (see: Chowdry and Emmerson, 2010). In this respect, Chowdry et al. (2008) find the EMA significantly increased students' performance at A-levels, particularly among pupils from deprived backgrounds. Our analysis of expectations complements -and potentially provides an explanation for- this boost in attainment.

their expectations of success in a university application based on their academic performance, closing the SES gap in attainment may not be enough to close the gap in university application. Other interventions (such as providing information about the application requirement or the returns to a university degree) may be needed, as suggested by Hastings et al. (2007); Hoxby and Avery (2012); Hoxby and Turner (2014); Goodman (2016) to explain why, in the USA, low-income-high achievers underestimate their suitability for selective schools and universities and generally apply to less prestigious institutions than richer peers with similar attainment.

2 Structure and costs of the English educational system at the time of our data

The English educational system is divided into 'key stages' (KS), with each stage reflecting the level of development the pupils are expected to achieve. Table 1 shows how the system is organized, the median age of the pupils in each stage, and the correspondent level in the UNESCO's International Standard Classification of Education (ISCED). Key stages 2, 3, 4 and 5 (from now on: 'KS2', 'KS3', 'KS4', 'KS5') end with exams.[5] KS2 and KS3 final exams are designed to show schools and parents whether the children are 'working at, above or below the target level for their age'.[6] In contrast, KS4 and KS5 final exams lead to certifications, respectively, the 'General Certificate of Secondary Education (GCSE)' and the 'Advanced Level General Certificate of Education (A-level)', with practical consequences both in the labor market and for university admission.

All students study the same subjects until grade 9. At the end of grade 9, students choose their GCSE subjects. Subjects studied and performance in GCSEs determine the range of A-level students can take. Performance in GCSE and A-levels affects the chances of being admitted to university. Universities admission requirements generally include the number and the subject of GSCEs and A-levels to be held, as well as the performance in such exams.[7]

[5]Until 2008, all exams were marked by external markers. KS3 taken after 2008 were marked by teachers.
[6]see http://www.direct.gov.uk.
[7]A personal statement must also be provided by all students. Only few institutions (e.g. medical schools, Oxford and Cambridge) interview applicants. See also Chowdry et al. (2013) and http://www.ucas.com/.

Table 1: English educational system

Key Stage	Ages	School years	ISCED classification	Type
0	3-5	0	ISCED 0	Pre-primary
1	5-7	1-2	ISCED 1	Primary
2	7-11	3-6	ISCED 1	
3	11-14	7-9	ISCED 2	Secondary
4	14-16	10-11	ISCED 3	
5	16-18	12-13	ISCED 3-4	Sixth form; Further Education (FE)
na	18+	na	ISCED 4-5-6	Further Education (FE) and Higher Education (HE)

Note: the shaded area indicates compulsory education at the time we consider.
Source: http://www.unesco.org

Therefore, choices at year 9 and 10 can potentially have far reaching consequences on the likelihood of being admitted to university. For this reason this paper analyzes young people in that particular age group.

In England, education is mainly free for pupils aged 3 to 18, as more than 90% of the students attend state-run schools financed through national taxation. While direct costs may be limited, the opportunity costs of being in in post compulsory education may be sizable, especially for low income families. This concern informed the introduction, in the academic year 2004/2005, of the 'Educational Maintenance Allowance' (EMA): a mean tested 30, 20 or 10 weekly allowance (depending on family income) aimed at keeping low-income people between the age of 16 and 19 in post-compulsory academic education until they reached the age for university education.

Low income students were considered those whose family income did not exceeds £30,000 for school year 2004/2005, and £30,810 for the following years. From school year 2006/2007, the EMA was extended also to students enrolled in vocational programs like 'Entry to Employment' (E2E) and 'Programme Led Apprenticeship' (PLA). The EMA was discontinued in the academic year 2011/2012 and replaced with a '16 to 19 Bursary Fund'. In 1999, the EMA was piloted in 15 Local Education Areas. The pilot targeted pupils who completed Year 11 in the summer of 1999. The piloting was then extended in 2000 (see, for example, Ashworth et al., 2001 and Dearden et al., 2009).[8]

[8]The LEAs where EMA was piloted in 1999 are: Bolton, Doncaster, Gateshead, Middlesbrough, Nottingham, Oldham, Southampton, Stoke-on-Trent, Walsall, Cornwall, Leeds, Lambeth, Southwark, Levisham, and Greenwich (Dearden et al., 2009. The LEAs where EMA was piloted in 2000 are: Northumberland, North

Free education ends with higher education. The first waves of LSYPE were collected in a time of significant changes in the costs of HE (see: Bolton, 2008a,b). In the early 2000s students entering full time HE were expected to contribute about £1000 a year towards the costs of their education, and the amount levied by universities was linked to students' family income. Students were eligible for zero real interest rate Income Contingency Loans (ICL) and for Maintenance Loans which were partly income assessed. Loans only had to be repaid after graduation and only by individuals whose annual income exceeded a threshold (set to £15,000 in 2000).

Between 2003 and 2004 changes in the funding to HE were discussed in the parliament. The Higher Education Act 2004 received royal assent on 1 July 2004. The act stated that, starting from the academic year 2006/2007, English HE institutions could levy fees ranging from 0 to 3000 pounds. Students could take up zero real interest rate Fees Loans and Maintenance Loans. Again, repayments were only due after graduation and only by individuals with annual income greater than £15,000. Income assessed maintenance grants were also introduced. Finally, institutions charging fees exceeding £2,700 were required to offer bursaries to recipients of the maximum Maintenance Loan.[9] In sum, the 2004 changes, affecting our sample members, pointed towards higher university fees, not anymore linked to students' family income (the link to students' income remained for maintenance grants and bursaries). This may have increased the perceived costs of university education, especially for low-SES students.

3 Data

The core of our analysis is based on the first two waves of the Longitudinal Study of Young People in England (LSYPE), a cohort study containing detailed information for about 15000

Tyneside, South Tyneside, Sunderland, Manchester, Salford, Tameside, Wigan, Lancashire, Knowsley, Liverpool, St Helens, Wirral, Halton, City of Kingston-Upon-Hull, Barnsley, Bradford, Wakefield, Leicestershire, Lincolnshire, Worcestershire, Coventry, Birmingham, Sandwell, Wolverhampton, Suffolk, Camden, Islington, Wandsworth, Hackney, Tower Hamlets, Barking and Dagenham, Newham, Haringey, Waltham Forest, Hammersmith and Fulham, Brent, Ealing.

[9] Various changes were made after 2004, both to fees and loans. However, these changes were announced after the time of our data collection and thus are irrelevant for our analysis

pupils living in England and born between 1^{st} September, 1989 and 31^{st} August, 1990. LSYPE is a two-stage probability sample with schools as primary sampling units (PSU) and students within schools as secondary sampling units (SSU).[10] Sample members were first interviewed in 2004 when they were in year 9 at school and then interviewed yearly.

LSYPE can be linked to the National Pupil Database (NPD): a longitudinal administrative dataset covering all pupils in state schools (from nurseries to secondary schools) and containing information on pupils' attainment and characteristics. In particular, the NPD contains information on free school meals eligibility (FSM) status. FSM eligible children are those whose parents receive some kind of support payments, such as: income support, income related 'job-seeker' and 'employment and support' allowance, child tax credit, working tax credit, or universal credit.

FSM eligibility is the variable commonly used to identify low income students (see: Kounali et al., 2008; Chowdry et al., 2013). As such, it has been widely used in past evaluations of the EMA (e.g., Chowdry et al., 2008) to indicate low-SES students. It is estimated that around 90 percent of FSM eligible students are also EMA recipients (Bolton, 2011). For this reason, and due to the fact that LSYPE does not have a measure of income which is consistent over time (see also: Meschi et al., 2011), FSM eligibility is used in this paper as a main proxy of low SES.[11]

The NPD also contains information on KS2 and KS3 exam results. KS2 exams are taken at the end of year 6, thus KS2 results are the last piece of information on respondents' ability prior to year 9 (wave 1); KS3 tests are taken at the end of year 9, thus between wave 1 and wave 2. In both KS2 and KS3 tests students are assessed in English, Mathematics and

[10] For the maintained sector, PSUs are stratified by deprivation level, measured as number of people receiving free-school meals. Deprived schools are over-sampled and so are ethnic minority students (Indian, Pakistani, Bangladeshi, Black Caribbean, Black African and Mixed). For the independent sector, schools are stratified by GCSE results in 2003, boarding status and students' gender. Pupil Referral Units are a stratum of their own. The selection probability of both independent schools and PRUs is proportional to the number of pupils aged 13 enrolled in the institution. Children educated at home are excluded from the sample and so are children attending very small schools (having fewer than 10 pupils for the maintained sector and fewer than 6 for the independent sector), boarders and children in England just for education purposes.

[11] There are two limits of the use of FSM as a proxy for low income status (Kounali et al., 2008; Hobbs and Vignoles, 2007). FSM eligibility status may not be able to fully capture persistent income deprivation and only identifies those claiming FSM (in 2013 they were about 90 percent of the eligible students, as estimated by Lord et al. (2013)). For these reasons, in a robustness check we construct an index of income deprivation and we use it in place of the FSM eligibility indicator.

Science. The Department of Education uses students' performance in the tests to compute a 'fine grading score' for each subject and a 'fine grading average score' computed as the average of the three.KS2 and KS3 'fine grading average score' are the measures of academic performance used in this paper. Figure 1 shows the timing of data collection in the first three waves, the timing of the events experienced by sample members, and an indication of respondents' age computed for a hypothetical young person born on Jan 1st 1990.

Figure 1: Young people's life events and timing of data collection

<table>
<tr><td>Calendar year</td><td colspan="12">2004</td><td colspan="12">2005</td><td colspan="12">2006</td></tr>
<tr><td>Month</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td></tr>
<tr><td>Age</td><td colspan="12">14</td><td colspan="12">15</td><td colspan="12">16</td></tr>
<tr><td>School year</td><td colspan="7">9</td><td colspan="12">10</td><td colspan="12">11</td><td colspan="5">12</td></tr>
<tr><td>Key stage</td><td colspan="7">3</td><td colspan="24">4</td><td colspan="5">5</td></tr>
<tr><td>Wave</td><td colspan="2"></td><td colspan="8">1</td><td colspan="5"></td><td colspan="6">2</td><td colspan="5"></td><td colspan="7">3</td><td colspan="3"></td></tr>
</table>

Our analysis uses two LSYPE questions on the perceived likelihood of applying to university and on the perceived likelihood of being admitted given a hypothetical application. The former asks the young person: 'How likely do you think it is that you will ever apply to go to university to do a degree?'. Four possible answers are provided: 'not at all likely', 'not very likely', 'fairly likely', 'very likely'.[12] Let A_{it} be the variable summarizing these four possible outcomes for individual i at time t (year 9 or 10), and $E_{it}(p^A)$ a continuous latent variable indicating the perceived likelihood of ever applying to university. We can write:

$$A_{it} = \begin{cases} \text{`very likely' if } \alpha_3 < E_{it}(p^A) \leq \alpha_4 \ . \\ \text{`fairly likely' if } \alpha_2 < E_{it}(p^A) \leq \alpha_3. \\ \text{`not very likely' if } \alpha_1 < E_{it}(p^A) \leq \alpha_2. \\ \text{`not at all likely' if } \alpha_0 < E_{it}(p^A) \leq \alpha_1. \end{cases} \tag{1}$$

where α_0,..., α_4 are unknown thresholds values of $E_{it}(p^A)$. Only students answering they are 'very likely', 'fairly likely' or 'not very likely' to apply to university ($E_{it}(p^A) > \alpha_1$) are asked: 'How likely do you think it is that if you do apply to go to university you will get in?'. The consequence of this question routing is that respondents are selected into the question about

[12]'Don't know' is not considered as a separate outcome. As a consequence, the 'Don't know' category is treated as missing, and the corresponding observations are excluded from the analysis.

the perceived likelihood of being admitted to university given a hypothetical application on the basis on their perceived likelihood of applying. In formulas:

$$S_{it} = \begin{cases} 0 \text{ if } \alpha_0 < E_{it}(p^A) \leq \alpha_1. \\ 1 \text{ otherwise.} \end{cases} \tag{2}$$

where S_{it} indicates whether student i at time t is asked the question about her perceived likelihood of being admitted to university if she applies. Therefore, the information we have on young people's perceived likelihood of being admitted to university given a hypothetical application can be described as follows:[13]

$$I_{it} = \begin{cases} \text{missing if } S_{it} = 0. \\ \text{`very likely' if } \mu_3 < E_{it}(p^I) \leq \mu_4 \text{ and } S_{it} = 1. \\ \text{`fairly likely' if } \mu_2 < E_{it}(p^I) \leq \mu_3 \text{ and } S_{it} = 1. \\ \text{`not very likely' if } \mu_1 < E_{it}(p^I) \leq \mu_2 \text{ and } S_{it} = 1. \\ \text{`not at all likely' if } \mu_0 < E_{it}(p^I) \leq \mu_1 \text{ and } S_{it} = 1. \end{cases} \tag{3}$$

where $E_{it}(p^I)$ is the continuous latent variable measuring perceived likelihood (at time t) of being admitted to university given a hypothetical application, and μ_0,...,μ_4 are the relevant thresholds.

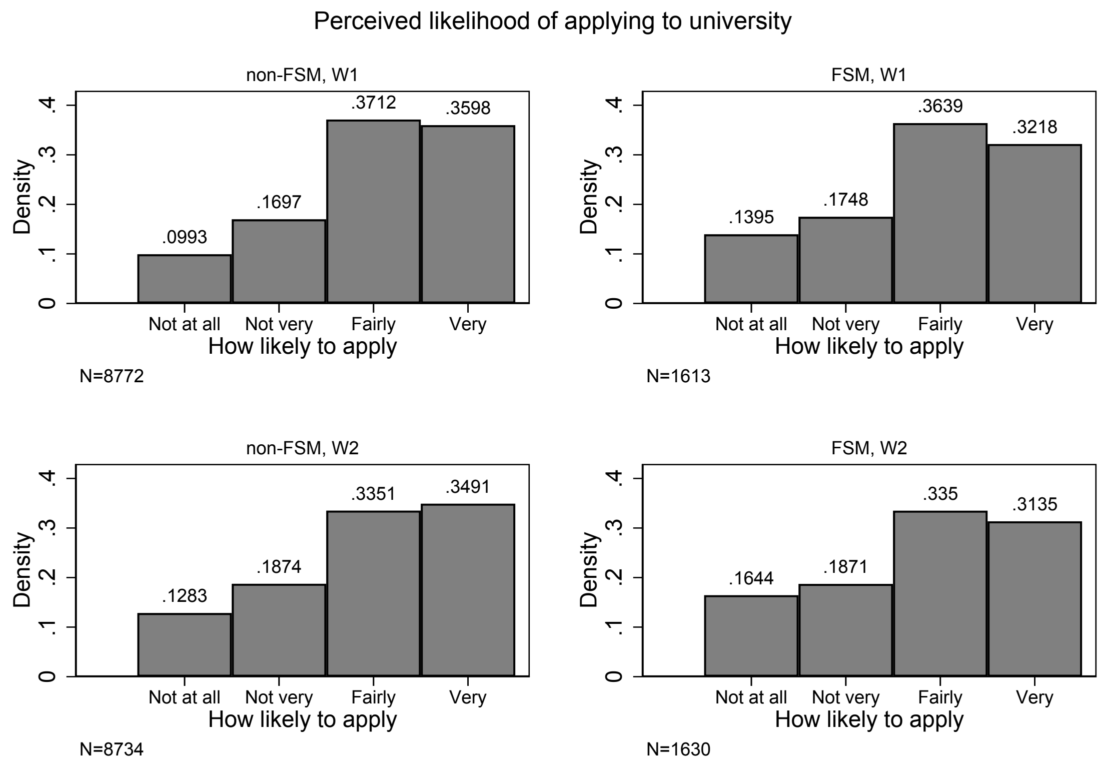
3.1 Descriptive statistics

Let us first look at the perceived likelihood of applying to university. Is there a SES gap? Figure 2 suggests there is, and that FSM eligible students report lower perceived likelihood of applying to university. For example, at wave one, just over 30% of FSM students (top right panel of figure 2) declare they are not at al likely or not very likely to apply to university, and a similar percentage declare they are very likely to apply to university. Among non-FSM students (top left panel), 27% are not at all likely or not very likely to apply to university,

[13] Also in this case, the 'Don't know' category is treated as missing, and the corresponding observations are excluded.

while 36% are very likely to apply to university. At wave two, expectations worsen for both FSM eligible and non-FSM eligible students. However, for the case of non-FSM students, the share of students who are very likely to apply to university (35%) is still greater than the share of students who are not at all or not very likely to apply to university (32%); in contrast, only 31% of the FSM eligible students answer they are very likely to apply to university, while 35% answer they are not at all or not very likely to apply.

Figure 2: Perceived likelihood of applying to university, by FSM eligibility



Source: University College London. UCL Institute of Education. Centre for Longitudinal Studies:
Next Steps: Sweeps 1 and 8, 2001 and 2016: Geographical Identifiers, 2001 Census Boundaries: Secure Access [computer file]. Colchester, Essex: UK Data Archive [distributor], June 2017. SN: 8189.
Next Steps: Sweeps 1-8, 2004-2016: Secure Access. [data collection]. 4th Edition. UK Data Service, 2018. SN: 7104; http://doi.org/10.5255/UKDA-SN-7104-3
Wave 1 and 2 data

Various factors could explain the SES gap in the perceived likelihood of applying to university. An obvious one is short-term barriers: staying in education after compulsory education and then attend university is simply too costly for low-SES (FSM eligible) students. Short-term monetary barriers are unlikely to be the only responsible of the SES gap. Even in the absence of short-term barriers, students from different SES may differ in the perceived likelihood of being admitted to university, and thus in their perceived likelihood of applying. Table

2 shows the relationship between the perceived likelihood of applying to university and the perceived likelihood of being admitted by year (9 and 10), and FSM eligibility status. What is the relationship between the two? Are they likely to be affected by the same factors?

Three aspects are worth noting. First, the two measures of expectations are positively related: when both questions are answered, the highest transition probabilities are generally on the diagonal. For example, in the case of FSM eligible students in year 9 (panel B.), 51.16% of those who answer they are 'very likely to apply' to university also answer that they are 'very likely' to be admitted, and 80.53% of those who answer they are 'fairly likely' to apply to university also answer they are 'fairly likely' to be admitted. This is not surprising, as the perceived likelihood of being admitted to university given a hypothetical application is a key determinant of the decision (and thus the perceived likelihood) of applying. Second, answers on the perceived likelihood of being admitted to university are less extreme, which could reflect the uncertainty due to the role of universities in determining acceptance. Third, for each level of perceived likelihood of applying to university, FSM eligible students report higher perceived likelihood of being admitted. This is consistent with low-SES students requiring a higher perceived likelihood of being admitted to be induced to apply, and could signal differences in monetary or psychological costs of education.

Who are the people excluded from the question on the perceived likelihood of being admitted to university given a hypothetical application? As mentioned, the LSYPE does not collect information on the perceived likelihood of being admitted to university for respondents who are 'not at all likely' to apply to university. This implicitly assumes these are discouraged students with no chance of applying. Table 3 shows the transition matrices from year 9 to 10 for the perceived likelihood of applying to university, and suggests this is not the case. Around 43% (40%) of non-FSM eligible (FSM eligible) students who answer they are 'not at all likely to apply to university' at year 9 give more positive answers the next year, and 4.17% (3.33%) switch from 'not at all likely to apply' at year 9, to 'very likely to apply' at year 10.

Do students answer they are 'not at all likely to apply' to university because they have a low perceived likelihood of being admitted or (also) because they cannot afford university? Answering this question is not straightforward, as data on the perceived likelihood of be-

Table 2: Perceived likelihood of applying to university and perceived likelihood of being admitted given a hypothetical application

			Likelihood of being admitted				
			not applicable	very likely	fairly likely	not very likely	not at all likely
A. Year 9, non-FSM eligible			11.07	20	55.51	12.1	1.32
Lik. of applying	very likely	37.05	0	46.85	51.64	1.36	0.15
	fairly likely	36.14	0	6.43	82.9	10.13	0.54
	not very likely	15.73	0	2.02	40.77	50.45	6.77
	not at all likely	11.07	100	0	0	0	0

			Likelihood of being admitted				
			not applicable	very likely	fairly likely	not very likely	not at all likely
B. Year 9, FSM eligible			14.24	20.73	49.69	13.84	1.5
Lik. of applying	very likely	34.39	0	51.16	46.4	2.31	0.13
	fairly likely	35.19	0	7.16	80.53	11.43	0.88
	not very likely	16.18	0	3.83	33.33	55.74	7.1
	not at all likely	14.24	100	0	0	0	0

			Likelihood of being admitted				
			not applicable	very likely	fairly likely	not very likely	not at all likely
C. Year 10, non-FSM eligible			14	20.63	50.86	12.98	1.53
Lik. of applying	very likely	35.92	0	49.53	49.53	0.82	0.12
	fairly likely	32.37	0	7.8	81.57	10.21	0.42
	not very likely	17.71	0	1.79	37.63	52.95	7.63
	not at all likely	14	100	0	0	0	0

			Likelihood of being admitted				
			not applicable	very likely	fairly likely	not very likely	not at all likely
D. Year 10, FSM eligible			17.12	20.93	45.88	13.88	2.19
Lik. of applying	very likely	32.99	0	52.37	45.25	2.22	0.16
	fairly likely	32.88	0	8.73	78.41	11.27	1.59
	not very likely	17.01	0	4.6	30.37	55.52	9.51
	not at all likely	17.12	100	0	0	0	0

ing admitted to university given a hypothetical application are missing for these students. However, about one quarter of students who at year 9 declared they were 'not at all likely to apply', at year 10 declare to be at least 'fairly likely' to be admitted to university if they apply (results not presented). This suggest students may be not at all likely to apply to university and yet expect a high probability of success in case of a hypotetical application, possibly due to short-term monetary barriers affecting the perceived likelihood of applying to university, but not the perceived likelihood of being admitted given a hypothetical application. In sum, our analysis suggests that failing to control for the sample selection induced by the LSYPE question wording would miss policy relevant populations of students, i.e. 'marginal' students, that is students who may or may not apply to university later on, and potentially successfull university applicants facing short-term monetary barriers.

Table 3: Perceived likelihood of applying to university: transition probabilities

			Year 10			
			very likely	fairly likely	not very likely	not at all likely
A. non-FSM eligible			35.71	33.16	18.24	12.89
Year 9	very likely	36.85	67.96	26.01	4.31	1.73
	fairly likely	36.8	25.48	50.44	17.5	6.58
	not very likely	16.35	5.36	25.1	44	25.54
	not at all likely	10.01	4.17	9.18	30.14	56.52

			Year 10			
			very likely	fairly likely	not very likely	not at all likely
B. FSM eligible			33.42	33.42	18	15.15
Year 9	very likely	33.9	63.71	28.66	5.14	2.49
	fairly likely	36.8	27.4	49.35	15.64	7.6
	not very likely	16.63	7.94	26.35	41.9	23.81
	not at all likely	12.67	3.33	9.17	27.92	59.58

4 The formation of educational expectations

Expectations on future access to higher education can be broken down into the expected likelihood of applying to university and the expected likelihood of being admitted given a hypothetical application. Short-term monetary barriers should affect the former but not the latter. Let us consider an individual who lives until age T. At age 18, when she finishes high school, she can make two choices: to apply to university ($A = 1$) or not ($A = 0$). The individual applies to university if the expected lifetime utility from applying to university, evaluated at age 18, exceeds the expected lifetime utility from not applying, also evaluated at age 18, or $E_{18}(U(A = 1)) - E_{18}(U(A = 0)) > 0$, where:

$$\begin{aligned} E_{18}(U(A=1)) = \quad & E_{18}(p^I)\left(\int_{18}^{T} E_{18}(r_\tau^e - c_\tau^e) \exp^{-\rho(\tau-18)} \, d\tau \right) - c^A + \qquad (4) \\ & +(1 - E_{18}(p^I))\left(\int_{18}^{T} E_{18}(r_\tau^{\neg e} - c_\tau^{\neg e}) \exp^{-\rho(\tau-18)} \, d\tau \right) \\ E_{18}(U(A=0)) = \quad & \int_{18}^{T} E_{18}(r_\tau^{\neg e} - c_\tau^{\neg e}) \exp^{-\rho(\tau-18)} \, d\tau \qquad (5) \end{aligned}$$

r_τ^e (c_τ^e), and $r_\tau^{\neg e}$ ($c_\tau^{\neg e}$) are, respectively, the returns to and the costs of being (not being) in education at time $\tau : 18 \leq \tau \leq T$, $\exp^{-\rho(\tau-18)}$ is a discounting factor capturing the idea that people generally assign more weight to current rather than future utility, and c^A are the costs of putting together the application (both monetary and opportunity costs). Finally, $E_{18}(p^I)$ is the expected likelihood at age 18 of being admitted to university conditional on applying.

$E_{18}(p^I)$ is elicited conditional on a hypothetical application. Therefore, it should be independent from short-term monetary costs of university. However, $E_{18}(p^I)$ is likely to be affected by long-term factors correlated with family SES. An example of these long-term factors is the quality of the home learning environment, which can affect cognitive and non-congnitive skills and thus students' academic performance. Long-term factors correlated with family SES can also affect the way academic performance is translated into educational expectations. For example, students from different SES may differ in the access to information on the determinants of university admission: high SES students are likely to be more familiar with the requirements for university admission, as they might be exposed to a higher number of peers and family members who have applied to and attended university.

Studying expectations over university education at the time of applying is of limited relevance, as it can be too late for low achievers to catch up with higher achieving peers. A more policy relevant question is understanding the formation of $E_t(p^A)$, i.e., expectations over the likelihood of applying to university in the future, evaluated when the individual is younger than 18 ($t < 18$). In our case, and for the rest of the paper, $t =$ year 9 or 10, right before the young people choose the subjects they will be studying in years 10 and 11, which could lead them to get the credentials to be admitted to university.

Expectations in year 9 and 10 are still formed by comparing the expected lifetime utility of applying to university and the expected lifetime utility of not applying. However, as the information set at t is not the same as the information set at age 18, at t, A is a random variable whose probability distribution must be estimated. Therefore, expectations at time t over the perceived likelihood of applying to university at age 18 depend on expectations at time t over $E_t(E_{18}(U(A = 1))) - E_t(E_{18}(U(A = 0)))$ i.e., the expected difference, at time t, between the lifetime utility of applying to university and the lifetime utility of not applying, both evaluated at age 18. In formulas:

$$\begin{aligned} E_t(E_{18}(U(A=1))) - E_t(E_{18}(U(A=0))) &= E_t(U(A=1)) - E_t(U(A=0)) = \qquad (6) \\ = \int_{16}^{18} E_t(r_\tau^{\neg e} - c_\tau^{e} - r_\tau^{\neg e} + c_\tau^{\neg e}) \exp^{-\rho(\tau - t)} d\tau &- E_t(c^A) \exp^{-\rho(18-t)} + \\ + E_t(p^I) \int_{19}^{T} E_t(r_\tau^{e} - c_\tau^{e} - r_\tau^{\neg e} + c_\tau^{\neg e}) &\exp^{-\rho(\tau - t)} d\tau \end{aligned}$$

$E_t(p^I) \int_{19}^{T} E_t(r_\tau^{e} - c_\tau^{e} - r_\tau^{\neg e} + c_\tau^{\neg e}) \exp^{-\rho(\tau - t)} d\tau$ is the expected value of the difference between expected costs and returns from age 19 onwards in the two scenarios of acquiring and not acquiring university education. This expected value depends on the perceived likelihood of being admitted to university conditional on application, which is interesting as it should not be affected by short-term monetary barriers. $E_t(c^A) \exp^{-\rho(18-t)}$ are the expected costs of applying to university. These are short-term barriers that have been found to influence university application (see, for example: Bettinger et al., 2012; Oreopoulos and Dunn, 2013; Castleman and Page, 2015). Finally, $\int_{16}^{18} E_t(r_\tau^{\neg e} - c_\tau^{e} - r_\tau^{\neg e} + c_\tau^{\neg e}) \exp^{-\rho(\tau - t)} d\tau$ is the difference in the expected net returns of being and not being in post-compulsory, not university education. Examples of factors influencing these net expected returns are the direct and opportunity costs of being in post-compulsory non-university education. Availability of the EMA or availability of information about its existence is an example of these direct costs, the local unemployment rate is an example of these opportunity costs. These costs are short-term barriers uncorrelated with the perceived probability of being admitted to university conditional on applying and thus can be used as exclusion restrictions to tackle the problem of endogenous selection induced by the LSYPE question wording.

5 The empirical model

We estimate two sets of regressions. In the first set we study whether short-term monetary barriers affect students' expectations on university application. We do this by looking at the impact of living in a former EMA pilot area on the perceived likelihood of applying to university measured at year 9 and 10. In the second set of regressions we look into the determinants of the perceived likelihood of being admitted to university conditional on a hypothetical application: a measures of expectations that should not be affected by short-term barriers. To address the problem of endogenous selection into the question on the perceived likelihood of being admitted to university induced by the LSYPE question wording, we use variation in the areas subject to the piloting of the EMA and variation in the local unemployment rate.

Let start looking at the perceived likelihood of being admitted to university, which is a crucial determinant of the decision to apply to university (see section 4) We model it as a function of observed characteristics, as follows:[14]

$$E_t(p_i^I) = \beta_0 + \beta_1 FSM_{it} + \beta_2 R_{it} + \beta_3 FSM_{it} R_{it} + \mathbf{x}_{it} \boldsymbol{\beta_4} + \nu_{it} \quad (7)$$

FSM_t is an indicator for free school meals eligibility, R_{it} is academic performance at the time of the survey, $FSM_t R_i$ is the interaction between FSM eligibility and academic performance, $\mathbf{x}_{it}$ is a set of variables at the individual, family and area level, and ν_{it} is an error term.

[14]Time subscripts are only used here to remind that expectations are measured previous to age 18, but the data are used cross-sectionally.

FSM eligibility status is used as an indicator for low SES. Academic performance at the time of the survey (R_{it}) is measured through the average KS2 score (taken at the end of year 6) at year 9, and through the average KS3 score (taken at the end of year 9) at year 10. R_{it} is a key determinant of expectations at year 9 or 10 of success in a future university application. R_{it} predicts academic performance at age 18: an important requirement for admission into HE. Moreover, R_{it} measures cognitive and non-cognitive skills, which determine success in other requirements for university application, such as personal statements, cover letters or interviews. R_{it} does not necessarily capture causality, as academic performance is likely to be correlated with unobserved factors affecting educational expectations. For example, students with high cognitive and non-cognitive skills are likely to have both high test scores and high expectations, and this does not necessarily imply that the causation goes from the former to the latter. However, controlling for R_{it} is important to show how much of the SES gap in expectations disappears when academic performance- a factor strongly determined by long-term barriers, such as students' lifetime quality of the learning environment- is controlled for.

We are particularly interested in $FSM_{it}R_{it}$, that is the interaction between free school meal eligibility status and academic performance. This helps investigate whether students from different SES differ in how academic performance is linked to expectations and can help explain why low-SES high-achievers underestimate their suitability for selective colleges compared to similarly bright richer peers (e.g., Hoxby and Avery, 2012; Hoxby and Turner, 2014; Goodman, 2016).

There is not a clear reason why the coefficient for FSM_tR_i should be significantly different from zero in the case of the perceived likelihood of being admitted to university given a hypothetical application. In a world of equal access to education and to information about university admission, this should not happen. Academic performance is meant to measure whether the students are working towards the targets defined for their age and therefore should be an indication of students' future chances of success. This means that two students with the same academic performance at year 9 or 10 should not differ in their expectations on future attainment, and thus in their expected likelihood of being admitted to university.

If the coefficient for FSM_tR_i is found to be different from zero, there must be log-term factors correlated with the permanent family income influencing how academic performance is translated into expectations. For example, students from different SES may differ in the degree of precision they attach to R_{it} as a predictor of future academic performance, both due to different available information on successful academic trajectories and due to different levels of self-perception and self-beliefs.

Finally, controls in $\mathbf{x}_{it}$ are chosen in line with the relevant empirical literature (see, for example Looker and Thiessen, 2004; Khoo and Ainley, 2005; Strand, 2007; Duckworth et al., 2009; Jacob and Wilder, 2010). These are: age, gender, ethnicity, number of siblings, the index of deprivation affecting children (IDACI) and Government Office Regions fixed effects.

The perceived likelihood of being admitted to university given a hypothetical application is an important determinant of the perceived likelihood of applying to university. Therefore, the perceived likelihood of applying to university can be written as:

$$E_{it}(p_i^A) = f(X_{it}, E_{it}(p_i^I), \mathbf{z}_{it}) \tag{8}$$

where X_{it} are the predictors of $E_{it}(p_i^I)$, or $X_{it} = [1, R_{it}, FSM_{it}, FSM_{it}R_{it}, \mathbf{x}_{it}]$, which also affect $E_{it}(p_i^A)$, while $\mathbf{z}_{it}$ are variables predicting $E_{it}(p_i^A)$ but not $E_{it}(p_i^I)$. Section 4 suggests $\mathbf{z}_{it}$ are, for example, variables predicting the returns and the costs, both direct and indirect (opportunity costs), of being in education (c_i^A, $c_{i\tau}^e$, and $r_{i\tau}^{\neg e}$). In our case, the variables in $\mathbf{z}_{it}$ are a dichotomous indicator for residence in a former EMA pilot area (plus its interaction with having an older sibling) and the local unemployment rate. The local unemployment rate is used to support identification especially at wave two when the effect of living in a EMA pilot area disappears.

Writing (8) in the usual linear form, we have:

$$E_{it}(p_i^A) = \gamma_0 + \gamma_1 R_{it} + \gamma_2 FSM_{it} + \gamma_3 FSM_{it}R_{it} + \mathbf{x}_{it}\gamma_\mathbf{4} + \gamma_5 E_{it}(p_i^I) + \mathbf{z}_{it}\delta_\mathbf{5} + \eta_{it} \tag{9}$$

or

$$E_{it}(p_i^A) = \delta_0 + \delta_1 R_{it} + \delta_2 FSM_{it} + \delta_3 FSM_{it} R_{it} + \mathbf{x}_{it}\boldsymbol{\delta_4} + \mathbf{z}_{it}\boldsymbol{\delta_5} + \upsilon_{it} \tag{10}$$

where $\delta_0 = \gamma_0 + \gamma_5\beta_0$, $\delta_1 = \gamma_1 + \gamma_5\beta_1$, $\delta_2 = \gamma_2 + \gamma_5\beta_2$, $\delta_3 = \gamma_3 + \gamma_5\beta_3$, $\delta_4 = \gamma_4 + \gamma_5\beta_4$, $\upsilon_{it} = \gamma_5\nu_{it} + \eta_{it}$

Therefore, (7) and (10) can be written, in matrix form, as:

$$E_{it}(p_i^I) = X_{it}\beta + \nu_{it} \tag{11}$$

$$E_{it}(p_i^A) = W_{it}\gamma + \upsilon_{it} \tag{12}$$

With $W_{it} = [X_{it}, \mathbf{z}_{it}]$. Equation (10) means that the variables in X_{it} affect $E_{it}(p_i^A)$ both through $E_{it}(p_i^I)$, and directly, and in fact the coefficients δs are a combination of γs and βs.

We estimate (11), i.e., the perceived likelihood of applying to university via ordered probit, with standard errors clustered at the Local Education Level (LEA), as this is the level at which the EMA was piloted. We are particularly interested in evaluating the effects of the short-term monetary factors affecting the decision to apply, that we operationalize using information on whether the respondent lived in a EMA pilot are. We estimate the regressions separately by year (9 or 10). For each year, we present eight specifications. Specification (1) contains only an indicator of whether the respondent lives in a former EMA pilot area. Specification (2) adds the IDACI index. The IDACI controls for deprivation at the area level and thus accounts for the fact that the EMA might have been piloted in more deprived areas.[15] Specification (3) adds FSM eligibility indicator, demographics and household characteristics. Specification (4) adds GOR fixed effects. Specification (5) adds an interaction between the indicator of residence in a former EMA pilot area and the FSM eligibility indicator. Specification (6) adds an indicator for whether the respondents has any old sibling and its interaction with the indicator for residence in a former EMA pilot area. Specification (6) adds grades, and specification (7) adds the interaction between the grades and the FSM eligibility status

[15] Eventual endogeneity left in the error term after controlling for the IDACI will bias our results downward, and thus our estimates of the effect of living in a EMA pilot area on the perceived likelihood of applying to university must be considered lower bounds.

The estimation of (7), i.e., the perceived likelihood of being admitted given a hypothetical application is complicated by the fact that the question on the perceived likelihood of being admitted to university given a hypothetical application is not asked to students not at all likely to apply to university, that is those having $\alpha_0 \leq E_{it}(p_i^A) \leq \alpha_1$. These cases are around 10% (14%) of the total among FSM (non-FSM) students at wave 1, and about 13% (16%) among FSM (non-FSM) students at wave 2 (see Figure A.1).

Expression (9) shows that, for $\gamma_5 \neq 0$, $Cov(\upsilon_{it}, \nu_{it}) \neq 0$ even when $Cov(\upsilon_{it}, \eta_{it}) = 0$, therefore the error in (7) will have a non-zero conditional expectation, i.e., $E(\nu_{it}|X_{it}, S_{it} = 1) \neq 0$. This means that, if selection into answering the question about the likelihood of being admitted to university is not taken into account, β will be inconsistently estimated. To overcome this problem, we use the following bivariate threshold crossing model:

$$S_{it} = \mathbb{1}(\alpha_1 < E_t(p_i^A) \leq \alpha_4) \tag{13}$$

$$I_{it} = \sum_{l=0}^{3} l\mathbb{1}(\mu_l < E_t(p_i^I) \leq \mu_{l+1}) \text{ if } S_{it} = 1 \tag{14}$$

where $\mu_l < \mu_{l+1}$, $\mu_1 = -\infty$, $\mu_4 = \infty$ and $E_t(p_i^A)$ and $E_t(p_i^I)$ are the latent variables defined, respectively, in (9) and (7).

We estimate the above bivariate threshold crossing model parametrically, by maximum likelihood (De Luca and Perotti, 2011), with standard errors clustered at the LEA level.[16] As I_{it} and S_{it} are discrete, it is possible to write the contribution to the maximum likelihood for all the 5 possible realizations of the two variables, i.e, $(S_{it} = 0)$, $(S_{it} = 1, I_{it} =$ 'very likely'), ... , $(S_{it} = 1, I_{it} =$ 'not at all likely'). Writing those contributions requires making parametric assumptions and thus ν_{it} and υ_{it} in equations (7) and (9) are assumed to follow a bivariate Gaussian distribution with zero mean, variance equal to 1 and correlation coefficient equal to ρ (see also: Miranda and Rabe-Hesketh, 2006). In addition, two restrictions are made, i.e: i) β_0 is set to 0, as it cannot be separately identified from the thresholds μ_1.[17] ii) There is at

[16] This is available as a stata routine called 'heckprobit'.

[17] This is the common restriction which is also needed to estimate standard ordered probit and logit.

least an exclusion restriction, that is a variable contained in W_{it}, but not in X_{it}, as implied by (8).[18]

We use a number of exclusion restrictions, namely: a dichotomous variable indicating whether the respondent lives in a former EMA pilot area, its interaction with whether the respondent has older siblings, and the unemployment rate, measured in the interview year.[19] Exclusion restrictions must satisfy two conditions. First, it must affect the perceived likelihood of applying to university. Second, it should not affect the perceived likelihood of being admitted to university given a hypothetical application (nor its perception), once indicators of family and neighborhood socio-economic deprivation are controlled for.

As previously discussed, living in a former EMA pilot area is likely to increase awareness of the existence of the EMA, reduce the perceived costs of post-compulsory non-university education, and thus boost the perceived likelihood of applying to university. The effect of local unemployment on education participation is theoretically ambiguous (see: Micklewright et al., 1990; Hubbard, 2018). On the one hand, high unemployment can reduce the expected returns from job-search, and the opportunity costs of staying in education. This will lead to a positive correlation between the local unemployment rate and the perceived likelihood of applying to university. On the other hand, pursuing further education may not be viable due to credit constraints. Moreover, high current unemployment can be interpreted as a signal of high future (i.e., post-university) unemployment, which reduces the expected future returns to university education. This will lead to a null or negative correlation between the local unemployment rate and the perceived likelihood of applying to university. Empirically, for the UK, high unemployment has been generally found to be associated with higher participation in post-compulsory education (Pissarides, 1981; Rice, 1987, 1999; Whitfield and Wilson,

[18]In principle, this restrictions is not required as identification can be achieved through functional form only. However, the use of exclusion restrictions is highly desirable (see a discussion in De Luca and Perotti, 2011).

[19]Data on the local unemployment rate are provided by the Office for National Statistics (ONS). Unemployment rates are calculated as the ratio at the local authority level between people aged 16 and over who are unemployed (those not in employment but available for or actively seeking work) and people aged 16 and over who are economically active (the sum of those in employment and the unemployed). Unemployment rates are model-based estimates, derived using the Annual Population Survey/Labour Force Survey (APS/LFS) and administrative data from the claimant count. The reference period for the unemployment data used in conjunction with wave 1 of LSYPE (year 9) is January-December 2004, while the reference period for the unemployment data used in conjunction with wave 2 of LSYPE (year 10) is January-December 2005.

1991; McVicar and Rice, 2001; Clark, 2011; Meschi et al., 2011; Tumino, 2013), and higher educational aspirations (Taylor and Rampino, 2014).[20]

Exogeneity of the exclusion restrictions with respect to the perceived likelihood of being admitted to university given a hypothetical application cannot be tested and thus must be assumed. For the case of the local unemployment rate, a potential problem could arise, by observing high unemployment rates, students infer an increase in university applications and therefore a drop in the probability of acceptance. In this scenario, the perceived likelihood of being admitted to university is affected by the local unemployment rate if i) students think strategically and take into account other students' application behavior AND ii) students apply only to universities near home, or apply to national/international universities but misinterpret the observed local unemployment rate as a proxy for the national/international unemployment rate. We believe that the contemporaneous occurrence of both conditions is unlikely, as the former condition requires the 14/15 years old respondents to be extremely familiar with the university application process and fully rational in they expectations, while the latter requires them to be extremely myopic.[21]

We estimate separate models for year 9 and year 10. For each model we present seven specifications. Specification (1) only uses the IDACI index in the main equation (probability of being admitted to university given a hypothetical application), and the IDACI index and the local unemployment rate in the selection equation. Controlling for the IDACI index is needed to make sure the local unemployment rate does not pick up the area's economic deprivation, and only picks up a measure of the opportunity costs of being in education. Specification (2) adds the FSM indicator. The coefficient for the FSM indicator in this specification describes the SES gap in the perceived likelihood of being admitted to university given a hypothetical application, once the problem of endogenous question routing is controlled for. Specification

[20] The only notable exception for the UK is Micklewright et al. (1990), which is unable to conclude unemployment leads to higher school participation. Some evidence that high unemployment leads to more/more academic education is also available for countries other than the UK, namely: the USA (Betts and McFarland, 1995; Hubbard, 2018), Spain (Petrongolo and Segundo, 2002), Italy (Mocetti, 2012; Tumino, 2013), and Germany (Tumino, 2013). Evidence that high unemployment is associated with higher expectations of university participation have been found, for the USA, by Reynolds and Pemberton (2001).

[21] In this respect, Morgan (1998) observes that young people's educational expectations are neither completely rational nor completely fantasized, while Czermak et al. (2016) estimate around 40% of adolescents think strategically.

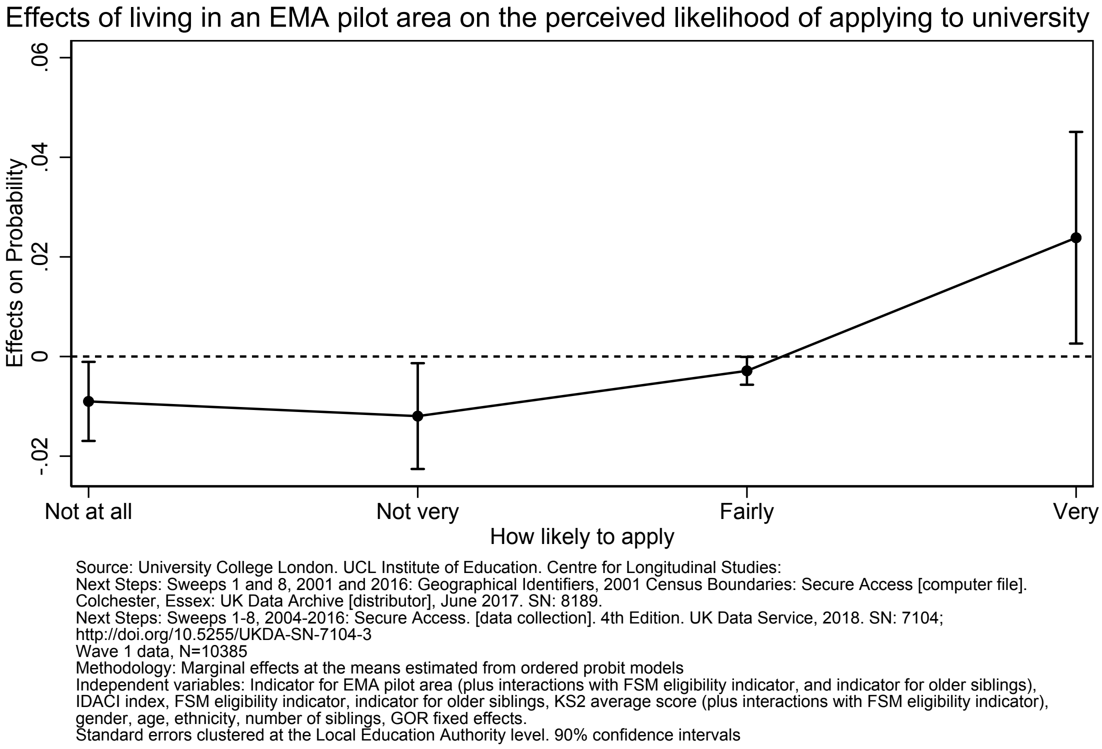
(3) adds demographics and household characteristics. Specification (4) adds GOR fixed effects. Specification (5) adds grades, which depend on long-term factors correlated with the permanent family income. We are interested to see what happens to the coefficient for FSM eligibility when academic performance -and thus the effect of long-term factors- is controlled for. Specification (6) adds the interaction between the FSM indicator and the grades. It is used to test the hypothesis that the link between academic performance and the perceived likelihood of being admitted to university conditional on a hypothetical application is weaker for low-SES students. Finally, specification (7) adds the indicator of residence in a former EMA pilot area, plus its interaction with an indicator for the presence of old siblings to strengthen the identification of the selection equation.

6 Results

We find evidence that short-term monetary barriers affect the perceived likelihood of applying to university. Figure 3 shows marginal effects at the means for the variable identifying former EMA pilot areas on the perceived likelihood of applying to university measured at age 14 (the first year of implementation of the EMA) computed from our full specification (specification 8). The figure shows that living in a former EMA pilot area decreases the probability of perceiving to be not at all likely and not very likely to go to university by around one percentage point, and increases the probability of perceiving to be very likely to apply to university by over two percentage points. Note that around 10 (17) percent of the sample perceive to be not at all likely (not very likely) to apply to university, and around 35 percent of the sample declare to perceive to be very likely to apply to university. Therefore, these effects are not negligible.

Figure 4 shows this effect is driven by FSM eligible students. Moreover, the figure shows the positive effect of the EMA on the perceived likelihood of applying to university is weaker in the case of students with older siblings. The left column of figure 4 shows living in a EMA pilot area increases the likelihood of perceiving to be very likely to apply to university by around five percentage points for non-FSM eligible students with no older sibling, and by almost 10 percentage points for FSM with no older siblings. The right column of figure

Figure 3: Marginal effects of living in an EMA pilot area on the perceived likelihood of applying to university (wave one, age 14)

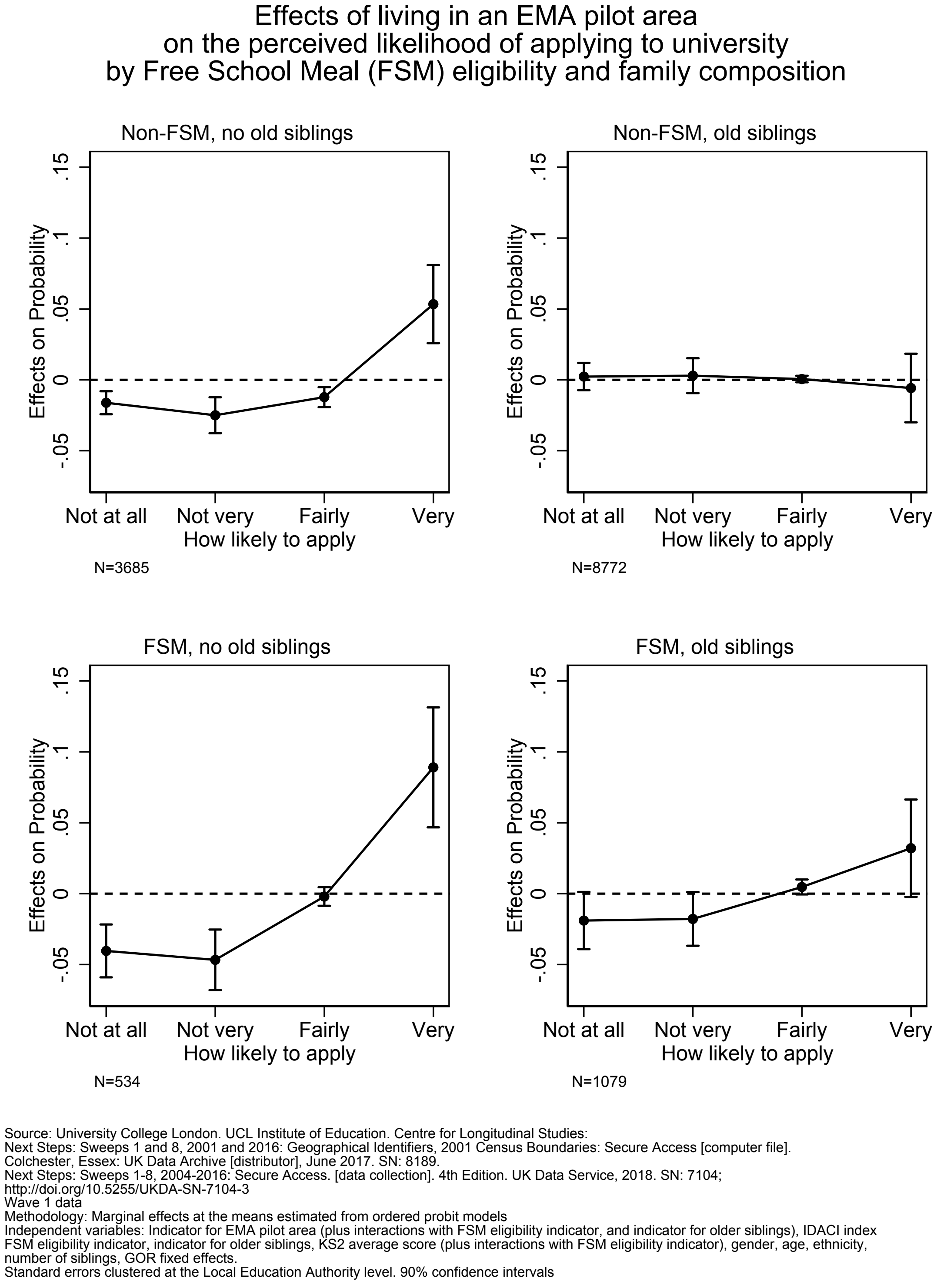


4 shows living in a EMA pilot area has no effect of the perceived likelihood of applying to university for non FSM eligible students with older siblings, and a have limited positive effect on FSM eligible students with older siblings.

The early availability of EMA in the pilot areas is likely to have increased awareness of the existence of financial aid to attend post-compulsory education. This is likely to have boosted the perceived likelihood of applying to university. Availability of the EMA is particularly salient for FSM eligible students who are both more in need of financial aid and more likely to be eligible for the EMA. Early availability of EMA in the pilot areas is also likely to have induced older siblings to attend university themselves. This may have led to a shift of family resources towards the older siblings (and away from the younger siblings), thus reducing the positive impact of the awareness of the EMA on the perceived likelihood of applying to university for younger siblings.

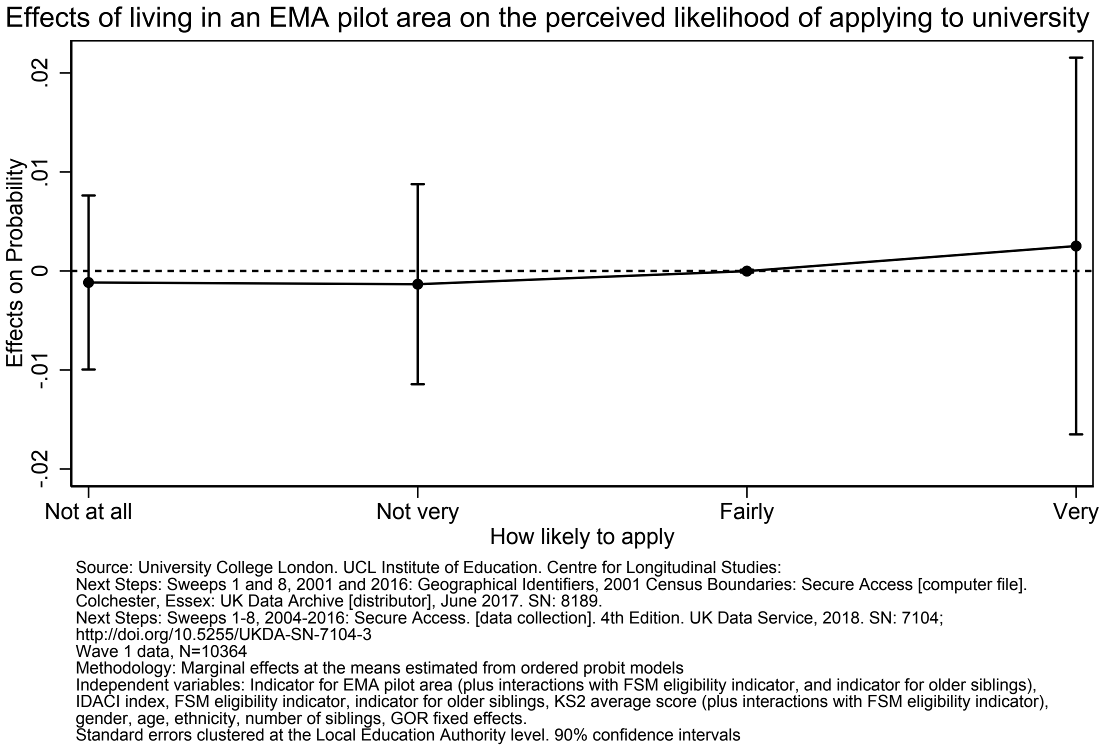
Figure 5 looks at expectations at age 15, when the EMA was fully in place in the whole country. It shows there is no positive impact of living in a EMA pilot areas on the perceived

Figure 4: Marginal effects of living in an EMA pilot area on the perceived likelihood of applying to university, by FSM eligibility and family composition (wave one, age 14)



likelihood of applying to university. However, for one particular subgroup, that is FSM eligible students without older siblings, we still find that living in an EMA pilot area increases the

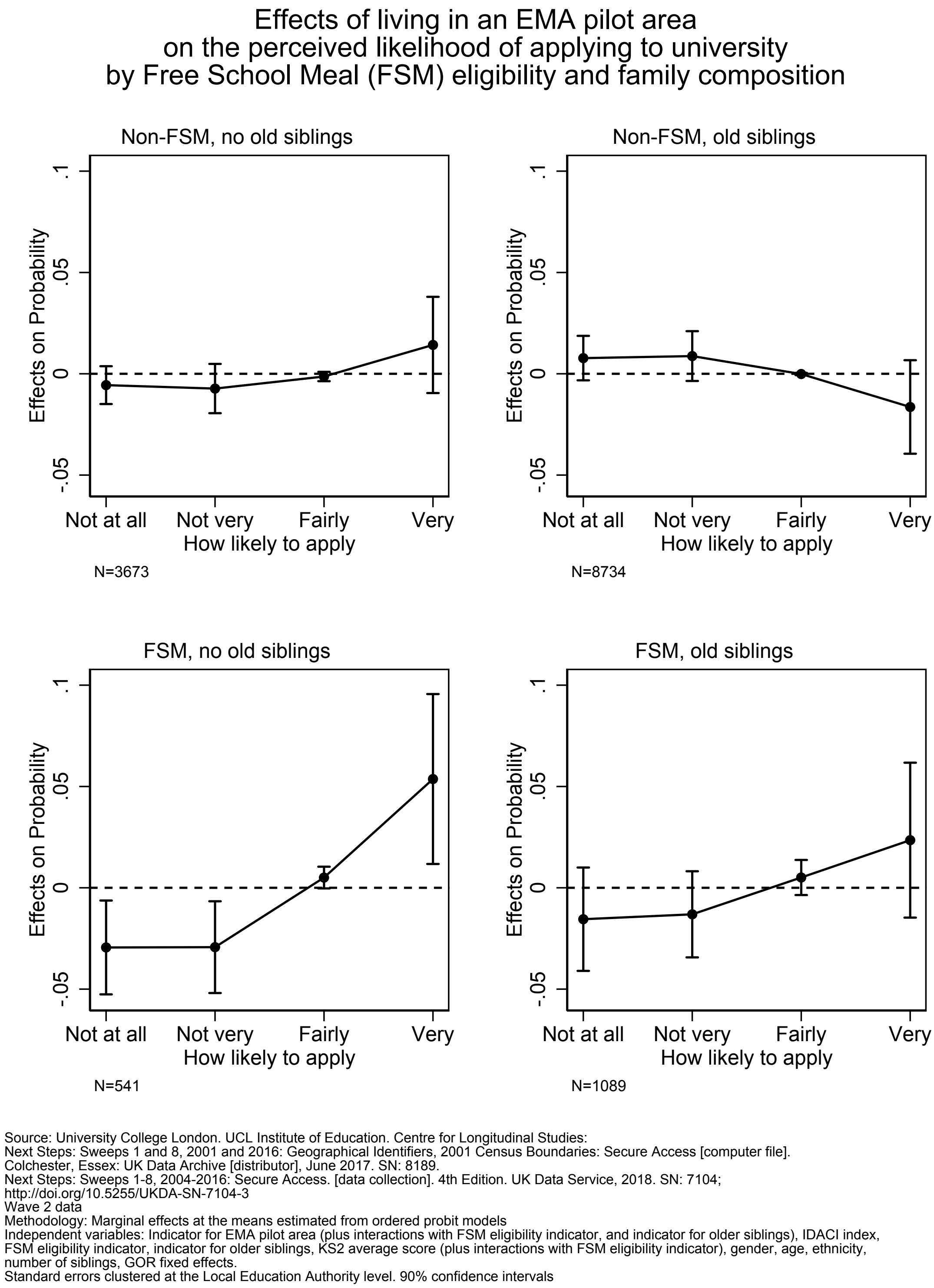
Figure 5: Marginal effects of living in an EMA pilot area on the perceived likelihood of applying to university (wave TWO, age 15)



perceived likelihood of applying to university. Figure 6 (bottom left quarter) shows the effect is around half of the size of the same effect estimated at wave one (and shown in Figure 4): living in a EMA pilot area decreases by two/three percentage points the probability of perceiving to be not at all likely and not very likely of applying to university, and increases by around five percentage points the probability of perceiving to be very likely to apply to university.

Coefficient for the ordered probit are shown in tables A.1 and A.2. They are of limited interests to show the impact of the EMA, which is better captured by marginal effects. However, three observations are worth noting on the controls used. First, the IDACI index has a negative coefficient, and its introduction leads to an increase in the the estimated positive coefficient for living in a former EMA pilot area. This suggests the EMA was in fact piloted in more deprived areas and confirms that our estimates on the impact of the positive EMA are conservative. Second, the coefficient for FSM eligibility is negative and it remains negative even when controls - including GOR fixed effects and measures of academic performance- are

Figure 6: Marginal effects of living in an EMA pilot area on the perceived likelihood of applying to university, by FSM eligibility and family composition (wave two, age 15)



added. This suggest a persistent SES gap in the perceived likelihood of applying to university. Third, the link between academic performance and perceived likelihood of applying to university is weaker for FSM eligible students. Both the lower subjective likelihood of applying to university and the weaker link between academic performance and subjective likelihood of applying to university estimated for FSM eligible students can be due to a combination of short-term and long-term barriers dis-proportionally affecting FSM eligible students.[22]

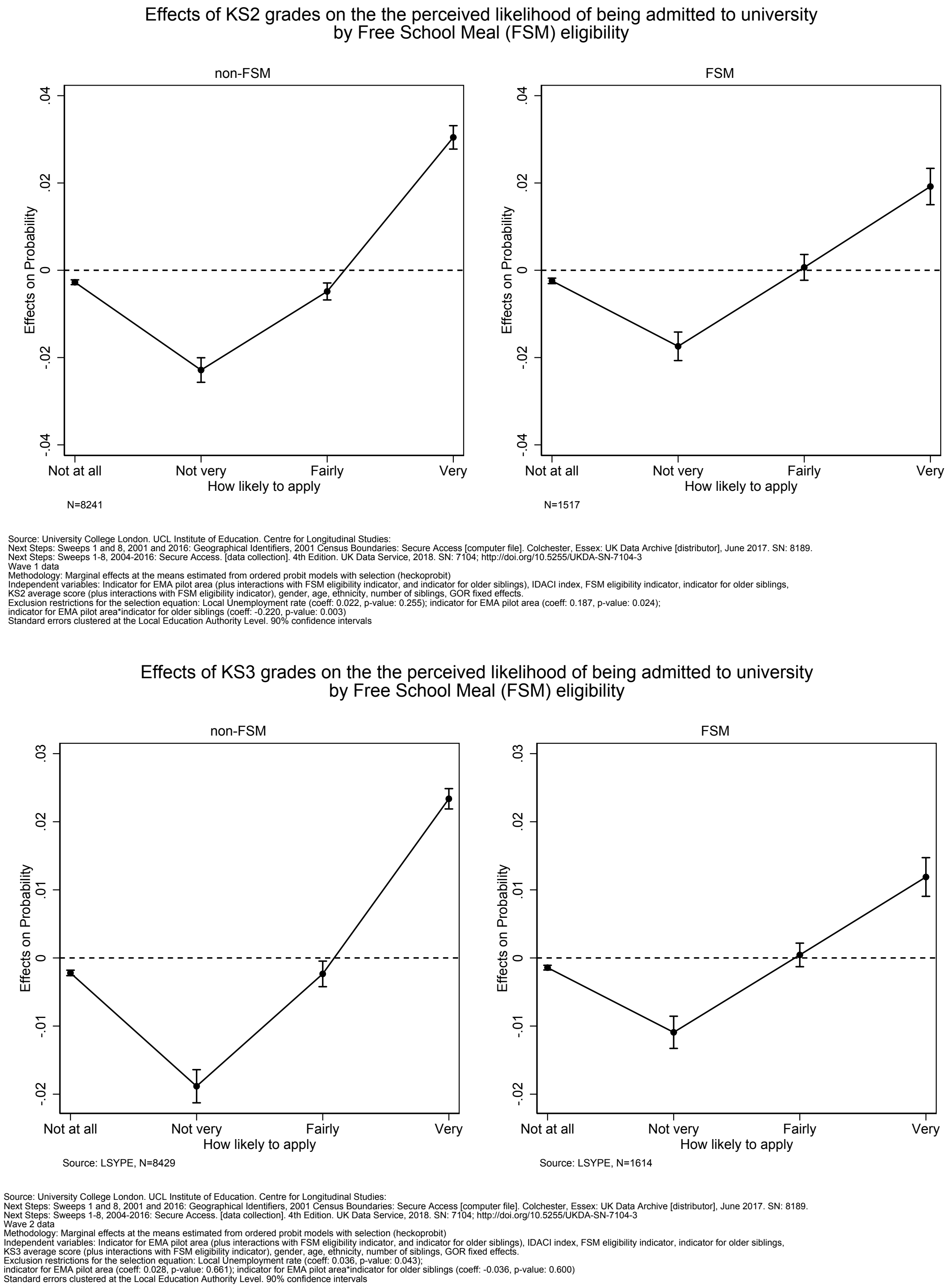
Let us now look at the perceived likelihood of being admitted conditional on a hypothetical application. It can be used to disentangle the role of long-term barriers in determining the SES gap in university participation. Tables A.3, A.4 and A.5 show the estimated coefficients. Higher local unemployment rate and and earlier provision of EMA are associated to increased likelihood of being at least not very likely to apply to university, and hence to be asked the question about the perceived likelihood of being admitted. The estimated coefficient for the indicator for FSM eligibility in specifications (2)-(5) is generally negative, confirming FSM-eligible students have a lower perceived likelihood of being admitted to university conditional on applying. However, when the respondent's grades are controlled for in specifications (6) and (7), the SES gap in the perceived likelihood of being admitted disappears. Note that, in the case of the perceived likelihood of applying to university (tables A.1 and A.2), controlling for the respondent's grades did not close the SES gap in expectations, suggesting part of the SES gap in the perceived likelihood of applying to university is due to short-term monetary constraints.

We are mainly interested in testing whether the link between the perceived likelihood of being admitted to university and students' academic performance is weaker for low-SES students. The marginal effects in figure 7, derived from specification (7) shows this is the case. For example, the bottom panel of figure 7, which uses wave two data, shows that a six percentage point increase in the KS3 grades increases the probability of perceiving to be very likely to be admitted to university conditional on applying by just over two percentage points for non-FSM eligible students (left column) and by just over one percentage point for FSM eligible students (right column). When we use KS2 instead of KS3, for comparability with wave 1, the results are qualitatively the same: the effects of grades are around double the size in the case of non-FSM eligible students, as shown in Figure A.2. This is evidence supporting

[22]Marginal effects for these variables lead to the same conclusions as coefficients and thus are not presented.

the hypothesis that there is less of an association between perceived likelihood of admission and academic performance for FSM eligible students.

Figure 7: Marginal effects of KS2 grades (top panel) and KS3 grades (bottom panel) on the perceived likelihood of being admitted to university given a hypothetical application (wave 1, age 14 in the top panel; wave 2, age 15 in the bottom panel), by FSM status.



Our results, derived on a type of expectations unaffected by short-term monetary barriers, suggest the miss-alignment between academic performance and (expectations over) university enrollment for low SES students can not only be the result of short-term monetary constraints. Other factors, for example misunderstanding in the requirements for university admission must be invoked to explain the results. This explanation is in line with existing evidence showing that, compared to more affluent peers, low-income students have noisier expectations on future returns to education (Jensen, 2010), and are less aware of the determinants of academic success, both in terms of test scores (Fryer, 2011) and in terms of admission probabilities in educational institutions (Hastings et al., 2007). We speculate that this might be due to students from low-SES households having more limited access to information over the requirements for university admission. This explanation is similar to what put forward by Hoxby and Avery (2012) and Hoxby and Turner (2014) to shed light on why, in the USA, low-income-high achievers generally apply to less prestigious universities if compared to richer peers with similar attainment.

7 Conclusions

This paper uses unique English data on teenagers' expectations at age 14/15 on higher education participation to shed light on why low SES young people are underrepresented among university students. In particular, we use information on the perceived likelihood of applying to university and the perceived likelihood of being admitted given a hypothetical application. Under the assumption that short-term monetary barriers only affect the former, while long-term factors correlated with permanent family income should affect both, we can shed light on the importance of short-term versus long-term barriers in shaping educational expectations.

Short-term monetary barriers do seem to affect the perceived likelihood of applying to university. Using variation in the length of the exposure to the EMA, a weekly allowance aimed at keeping low-SES students in post-compulsory education until they reach university, we provide evidence that information about the available financial aid increases the perceived likelihood of applying to university, especially for low-SES students. We interpret this as evidence of the role of short-term monetary barriers.

Short-term monetary barriers do not seem the only force in place to explain why low-SES are underrepresented in higher education. Long-term barriers correlated with the family permanent income, such as the quality of the learning environment and the access to information about the requirements for admission to university also matter. The SES gap in expectations over university application is strongly reduced by controlling for academic performance, and it is even reverted in the case of the subjective likelihood of being admitted to university. This suggests that the quality of the home learning environment does matter in the formation of the SES gap in expectations. Moreover, the link between the perceived likelihood of being admitted to university and academic performance is weaker for low-SES students: we interpret this as evidence that low-SES students may not be aware of the requirements for university admission.

Our results suggest there are multiple ways to reduce the SES gap in university participation. Increasing awareness of available financial support offered to low-income students to participate in higher education is important to increase low-SES students' expectations towards university education. However, it may not be enough to close the SES gap in university participation (see also: Carneiro and Heckman, 2002; Stinebrickner and Stinebrickner, 2003). Long term barriers need also to be removed. For example, investing in low-SES children's skills by offering them early support to compensate the effect of a poor home learning environment could reduce the SES gap in attainment. Better information about the requirements for university admission may also be required to make sure better attainment for low SES students is translated into higher participation into HE.

References

Anders, J. and Micklewright, J. (2013). Teenagers' expectations of applying to university; how do they change? Technical report, Working Paper No 13-13. Institute of Education.

Ashworth, K., Hardman, J., Liu, W., maquire, S., Middleton, S., Dearden, L., Emmerson, C., Frayne, C., Goodman, A., Ichimura, H., and Meghir, C. (2001). Education maintenance allowance: The first year a quantitative evaluation. Technical report, Department for Education and Employment. Research Report RR257.

Attanasio, O. and Kaufmann, K. (2009). Educational choices, subjective expectations, and credit constraints. Technical report, NBER Working Papers 15087, National Bureau of Economic Research.

Auspurg, K., Burton, J., Cullinane, C., Delavande, A., Fumagalli, L., Iacovou, M., Jäckle, A., Kaminska, O., Lynn, P., Mathews, P., Nicolaas, G., Nicoletti, C., Cong, Y., and Zafar, B. (2013). Understanding society innovation panel wave 5: Results from methodological experiments. Understanding Society Working Paper Series 2013-06, Understanding Society at the Institute for Social and Economic Research.

Barr, A. (2019). Fighting for education: Financial aid and degree attainment. *Journal of Labor Economics*, forthcoming.

Bettinger, E., Long, B., Oreopoulos, P., and Sanbonmatsu, L. (2012). The role of application assistance and information in college decisions: Results from the h&r block fafsa experiment. *Quarterly Journal of Economics*, 127(3):1205–1242.

Betts, J. R. and McFarland, L. L. (1995). Safe port in a storm: The impact of labor market conditions on community college enrollments. *The Journal of Human Resources*, 30(4):pp. 741–765.

Black, S. E., Devereux, P. J., and Salvanes, K. G. (2005). The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, 120(2):669–700.

Black, S. E., Grönqvist, E., and Öckert, B. (2018). Born to lead? the effect of birth order on noncognitive abilities. *Review of Economics and Statistics*, 100(2):274–286.

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

Blossfeld, H. and Shavit, Y. (1993). *Persistent Inequality: Changing Educational Attainment in Thirteen Countries.* Boulder, Co: Westview Press.

Bolton, P. (2008a). Student loan statistics. Briefing Paper 1079, House of Commons.

Bolton, P. (2008b). Tuition fees statistics. Briefing Paper 917, House of Commons.

Bolton, P. (2011). Education maintenance allowance (ema) statistics. Standard Note SNSG/5778, House of Commons Library, London, UK.

Cameron, S. and Heckman, J. (1999). Can tuition policy combat rising wage inequality? In Kosters, M., editor, *Financing College Tuition*, chapter Can Tuition Policy Combat Rising Wage Inequality? Washington, D.C.: AEI Press.

Cameron, S. and Heckman, J. (2001). The dynamics of educational attainment for black, hispanic and white males. *Journal of Political Economy*, 109:455–99.

Carneiro, P. and Heckman, J. (2002). The evidence on credit constraints in post-secondary schooling. *The Economic Journal*, 112:705–734.

Castleman, B. L. and Long, B. T. (2016). Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation. *Journal of Labor Economics*, 34(4):1023–1073.

Castleman, B. L. and Page, L. C. (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior & Organization*, 115:144–160.

Chevalier, A., Harmon, C., O'Sullivan, V., and Walker, I. (2010). The impact of parental income and education on the schooling of their children. Working Papers 201032, Geary Institute, University College Dublin.

Chowdry, H., Crawford, C., Dearden, L., Goodman, A., and 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., Crawford, C., Dearden, L., Joyce, R., Sibieta, L., Sylva, K., and Washbrook, E. (2010). Poorer children's education attainment: how important are attitudes and behaviour? Technical report, Joseph Rowtree Foundation.

Chowdry, H., Crawford, C., and Goodman, A. (2009). Drivers and barriers to educational success. evidence from the longitudinal study of young people in england. Technical report, DCSF research Report RR102.

Chowdry, H., Dearden, L., and Emmerson, C. (2008). Education maintenance allowance evaluation with administrative data: the impact of the ema pilots on participation and attainment in post-compulsory education.

Chowdry, H. and Emmerson, C. (2010). An efficient maintenance allowance? Available at www.ifs.org.uk/publications/5370.

Clark, D. (2011). Do recessions keep students in school? the impact of youth unemployunem on enrolment in post-compulsory education in england. *Economica*, 78:523–545.

Czermak, S., Feri, F., Glätzle-Rützler, D., and Sutter, M. (2016). How strategic are children and adolescents? experimental evidence from normal-form games. *Journal of Economic Behavior & Organization*, 128:265–285.

De Luca, G. and Perotti, V. (2011). Estimation of ordered response models with sample selection. *Stata Journal*, 11(2):213–239.

Dearden, L., Emmerson, C., and Meghir, C. (2009). Conditional cash transfers and school dropout rates. *Journal of Human Resources*, 44(4).

Deming, D. and Dynarski, S. (2009). Into college, out of poverty? policies to increase the postsecondary attainment of the poor. Working Paper 15387, National Bureau of Economic Research.

Dilnot, C. (2017). The relationship between a-level subject choice and league table score of university attended: the 'facilitating', the 'less suitable' and the counter-intuitive. Technical report, Centre for Longitudinal Studies Working paper 2017/7.

Dinkelman, T. and Martínez, A. (2011). Investing in schooling in chile: The role of information about financial aid for higher education. Technical report, CEPR Discussion Paper No. DP8375.

Duckworth, K., Akerman, R., Morrison Gutman, L., and Vorhaus, J. (2009). Influences and leverages on low levels of attainment: a review of literature and policy initiatives. Technical report, Centre for Research on the Wider Benefits of Learning, Institute of Education, University of London.

Dynarski, S. (2008). Building the stock of college-educated labor. *Journal of human resources*, 43(3):576–610.

Fowley (2011). Merit goods and the role that parents play in their children's educational expectations and outcomes. Technical report, mimeo.

Fryer, R. G. (2011). Financial incentives and student achievement: Evidence from randomized trials. *The Quarterly Journal of Economics*, 126(4):1755–1798.

Goodman, S. (2016). Learning from the test: Raising selective college enrollment by providing information. *Review of Economics and Statistics*, 98(4):671–684.

Hastings, J. S., Weelden, R. V., and Weinstein, J. (2007). Preferences, information, and parental choice behavior in public school choice. NBER Working Papers 12995, National Bureau of Economic Research, Inc.

Hobbs, G. and Vignoles, A. (2007). Is Free School Meal Status a Valid Proxy for Socio-Economic Status (in Schools Research)? CEE Discussion Papers, Centre for the Economics of Education, LSE 0084, Centre for the Economics of Education, LSE.

Hoxby, C. and Avery, C. (2012). The missing 'one-offs': The hidden supply of high-achieving, low income students. Working Paper 18586, National Bureau of Economic Research.

Hoxby, C. and Turner, S. (2014). Expanding college opportunities for high-achieving, low income students. Technical report, SIEPR Discussion Paper No. 12-014.

Hubbard, D. (2018). The impact of local labor market shocks on college choice: Evidence from plant closings in michigan. mimeo.

Jacob, B. A. and Wilder, T. (2010). Educational expectations and attainment. Working Paper 15683, National Bureau of Economic Research.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, 125 (2):515–548.

Jerrim, J. (2011). Disadvantaged childrens 'low' educational expectations: Are the us and uk really so different to other industrialized nations? DoQSS Working Papers 11-04, Department of Quantitative Social Science - Institute of Education, University of London.

Kaufmann, K. (2010). Understanding the income gradient in college attendance in mexico: The role of heterogeneity in expected returns. Technical report, IGIER, working paper n. 362.

Khoo, S. and Ainley, J. (2005). Attitudes, intentions and participation. Technical report, LSAY Research Reports. Longitudinal surveys of Australian youth research report n.41.

Kounali, D., Robinson, T., Goldstein, H., and Lauder, H. (2008). The probity of free school meals as a proxy measure for disadvantage. Technical report, Education Department, University of Bath.

Looker, D. and Thiessen, V. (2004). Aspirations of canadian youth for higher education: final report. Technical report, Learning Policy Directorate, Human Resources and Skills Development Canada.

Lord, A., Easby, J., and Evans, H. (2013). Pupils not claiming free school meals-2013. research report. Technical report, Department For Education and Skills.

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

Manski, C. (1993). Adolescent econometricians: How do youth infer the returns to schooling? In *Studies of Supply and Demand in Higher Education*, NBER Chapters, pages 43–60. National Bureau of Economic Research, Inc.

McVicar, D. and Rice, P. (2001). Participation in further eduction in england and wales: An analysis of post-war trends. *Oxford Economic Papers*, 53(1):47–66.

Meschi, E., Swaffield, J. K., and Vignoles, A. (2011). The Relative Importance of Local Labour Market Conditions and Pupil Attainment on Post-Compulsory Schooling Decisions. IZA Discussion Papers 6143, Institute for the Study of Labor (IZA).

Micklewright, J., Pearson, M., and Smith, S. (1990). Unemployment and early school leaving. *Economic Journal*, Vol. 100(400, Conference Papers (1990)):163–169.

Miranda, A. and Rabe-Hesketh, S. (2006). Maximum likelihood estimation of endogenous switching and sample selection models for binary, ordinal, and count variables. *Stata Journal*, 6(3):285–308.

Mocetti, S. (2012). Educational choices and the selection process: before and after compulsory schooling. *Education Economics*, 20(2):189–209.

Morgan, S. (1998). Adolescent educational expectations. rationalized, fantasized, or both? *Rationality and Society*, 10:131–162.

Nguyen, T. (2010). Information, role models and perceived returns to education: Experimental evidence from madagascar. Technical report, MIT Working Paper.

Oreopoulos, P. and Dunn, R. (2013). Information and college access: Evidence from a randomized field experiment. *The Scandinavian Journal of Economics*, 115(1):3–26.

Petrongolo, B. and Segundo, M. J. S. (2002). Staying-on at school at 16: the impact of labor market conditions in spain. *Economics of Education Review*, 21(4):353 – 365.

Pissarides, C. A. (1981). Staying-on at school in england and wales. *Economica*, 48(192):345–63.

Price, J. (2008). Parent-child quality time does birth order matter? *Journal of human resources*, 43(1):240–265.

Reynolds, J. and Pemberton, J. (2001). Rising college expectations among youth in the united states: A comparison of 15 and 16 year olds in the 1979 and 1997 nlsy. *Journal of Human Resources*, 36:703–72.

Rice, P. (1999). The impact of local labour markets on investment in further education: Evidence from the england and wales youth cohort studies. *Journal of Population Economics*, 12(2):pp. 287–312.

Rice, P. G. (1987). The Demand for Post-compulsory Education in the UK and the Effects of Educational Maintenance Allowances. *Economica*, 54(216):465–75.

Sjoquist, D. L. and Winters, J. V. (2015). State merit-based financial aid programs and college attainment. *Journal of Regional Science*, 55(3):364–390.

Stinebrickner, R. and Stinebrickner, T. (2003). Understanding educational outcomes of students from low-income families: Evidence from a liberal arts college with a full tuition subsidy program. *Journal of Human Resources*, 38(3).

Stinebrickner, T. and Stinebrickner, R. (2012). Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, 30(4):707 – 748.

Strand, S. (2007). Minority ethnic pupils in the longitudinal study of young people in england. Technical report, London. DCSF.

Taylor, M. and Rampino, T. (2014). Educational aspirations and attitudes over the business cycle. *Economica*, 81(324):649–673.

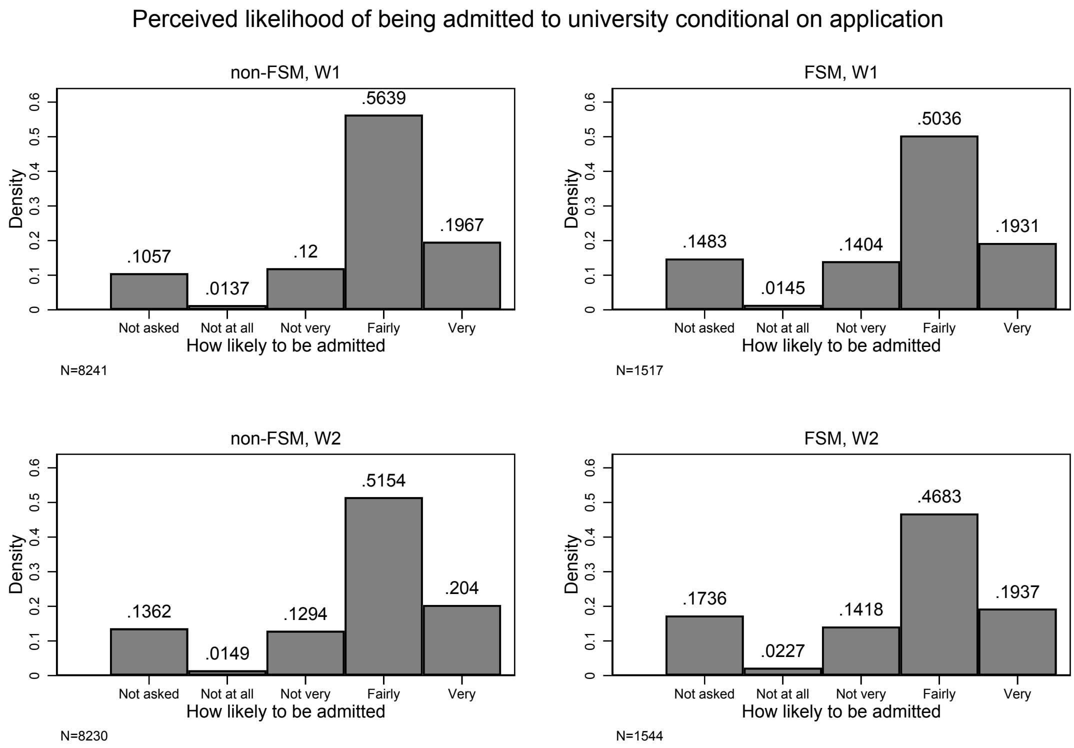
Tumino, A. (2013). The effect of local labour market conditions on educational choices: a cross country comparison. Technical report, University of Antwerp: Herman Deleeck Centre for Social Policy. ImPRovE Working Papers 13/6.

Whitfield, K. and Wilson, R. (1991). Staying on in full-time education: The educational participation rate of 16-year-olds. *Economica*, 58(231):391–404.

Zafar, B. (2011). How do college students form expectations? *Journal of Labor Economics*, 29(2):301 – 348.

A Additional tables and figures

Figure A.1: Marginal effects of KS3 grades on the perceived likelihood of being admitted to university given a hypothetical application (wave 2, age 15), by FSM status.



Source: University College London. UCL Institute of Education. Centre for Longitudinal Studies:
Next Steps: Sweeps 1 and 8, 2001 and 2016: Geographical Identifiers, 2001 Census Boundaries: Secure Access [computer file]. Colchester, Essex: UK Data Archive [distributor], June 2017. SN: 8189.
Next Steps: Sweeps 1-8, 2004-2016: Secure Access. [data collection]. 4th Edition. UK Data Service, 2018. SN: 7104; http://doi.org/10.5255/UKDA-SN-7104-3
Wave 1 and 2 data

Table A.1: Perceived likelihood of applying to university, wave 1.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EMA	0.122**	0.165***	0.076*	0.063	0.033	0.130**	0.141***	0.142***
	(0.051)	(0.057)	(0.040)	(0.040)	(0.041)	(0.051)	(0.044)	(0.044)
IDACI		-0.317***	-0.551***	-0.618***	-0.620***	-0.643***	-0.209***	-0.174**
		(0.094)	(0.082)	(0.083)	(0.083)	(0.084)	(0.080)	(0.079)
FSM			-0.226***	-0.237***	-0.331***	-0.345***	-0.162***	1.156***
			(0.028)	(0.029)	(0.041)	(0.042)	(0.045)	(0.247)
EMA*FSM					0.179***	0.201***	0.107	0.118*
					(0.056)	(0.056)	(0.068)	(0.063)
Old Siblings						-0.090***	-0.062**	-0.061**
						(0.023)	(0.025)	(0.024)
EMA*Old Siblings						-0.164***	-0.160***	-0.158***
						(0.043)	(0.045)	(0.045)
KS2							0.123***	0.133***
							(0.005)	(0.005)
FSM*KS2								-0.052***
								(0.009)
Demographics	No	No	Yes	Yes	Yes	Yes	Yes	Yes
GOR fixed effects	No	No	No	Yes	Yes	Yes	Yes	Yes
N	14737	14735	10783	10783	10783	10778	10385	10385

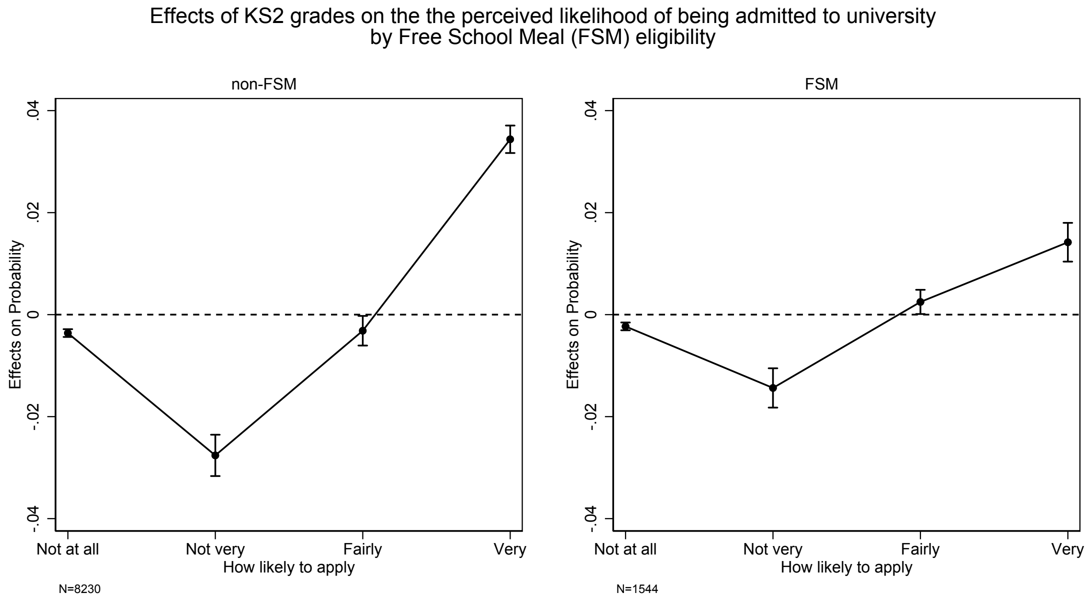
Source: University College London. UCL Institute of Education. Centre for Longitudinal Studies: Next Steps: Sweeps 1 and 8, 2001 and 2016: Geographical Identifiers, 2001 Census Boundaries: Secure Access [computer file], Colchester, Essex: UK Data Archive [distributor], June 2017.
SN: 8189. Next Steps: Sweeps 1-8, 2004-2016: Secure Access. [data collection]. 4th Edition. UK Data Service, 2018. SN: 7104. http://doi.org/10.5255/UKDA-SN-7104-3
Coefficients from ordered probit models.
Demographics are: ethnicity, gender, age, number of siblings.
Standard errors clustered at the Local Education Authority level.
*** 1 percent, *** 5 percent, * 10 percent

Table A.2: Perceived likelihood of applying to university, wave 2.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EMA	0.080	0.142**	0.055	0.034	0.004	0.057	0.039	0.039
	(0.053)	(0.059)	(0.039)	(0.040)	(0.040)	(0.047)	(0.040)	(0.040)
IDACI		-0.463***	-0.849***	-0.925***	-0.926***	-0.950***	-0.478***	-0.433***
		(0.115)	(0.091)	(0.084)	(0.084)	(0.084)	(0.083)	(0.082)
FSM			-0.185***	-0.197***	-0.292***	-0.302***	-0.104*	1.608***
			(0.036)	(0.036)	(0.059)	(0.059)	(0.058)	(0.203)
EMA*FSM					0.177**	0.200***	0.112	0.124*
					(0.070)	(0.070)	(0.073)	(0.071)
Old Siblings						-0.098***	-0.056*	-0.055*
						(0.027)	(0.030)	(0.029)
EMA*Old Siblings						-0.097**	-0.088*	-0.087*
						(0.041)	(0.047)	(0.046)
KS2							0.137***	0.151***
							(0.005)	(0.005)
FSM*KS2								-0.068***
								(0.007)
Demographics	No	No	Yes	Yes	Yes	Yes	Yes	Yes
GOR fixed effects	No	No	No	Yes	Yes	Yes	Yes	Yes
N	12762	12760	10947	10947	10947	10752	10364	10364

Source: University College London. UCL Institute of Education. Centre for Longitudinal Studies: Next Steps: Sweeps 1 and 8, 2001 and 2016: Geographical Identifiers, 2001 Census Boundaries: Secure Access [computer file], Colchester, Essex: UK Data Archive [distributor], June 2017.
SN: 8189. Next Steps: Sweeps 1-8, 2004-2016: Secure Access. [data collection]. 4th Edition. UK Data Service, 2018. SN: 7104. http://doi.org/10.5255/UKDA-SN-7104-3
Coefficients from ordered probit models.
Demographics are: ethnicity, gender, age, number of siblings.
Standard errors clustered at the Local Education Authority level.
*** 1 percent, *** 5 percent, * 10 percent

Figure A.2: Marginal effects of KS3 grades on the perceived likelihood of being admitted to university given a hypothetical application (wave 2, age 15), by FSM status.



Source: University College London. UCL Institute of Education. Centre for Longitudinal Studies:
Next Steps: Sweeps 1 and 8, 2001 and 2016: Geographical Identifiers, 2001 Census Boundaries: Secure Access [computer file]. Colchester, Essex: UK Data Archive [distributor], June 2017. SN: 8189.
Next Steps: Sweeps 1-8, 2004-2016: Secure Access. [data collection]. 4th Edition. UK Data Service, 2018. SN: 7104; http://doi.org/10.5255/UKDA-SN-7104-3
Wave 2 data
Methodology: Marginal effects at the means estimated from ordered probit models with selection (heckoprobit)
Independent variables: Indicator for EMA pilot area (plus interactions with FSM eligibility indicator, and indicator for older siblings), IDACI index, FSM eligibility indicator, indicator for older siblings,
KS2 average score (plus interactions with FSM eligibility indicator), gender, age, ethnicity, number of siblings, GOR fixed effects.
Exclusion restrictions for the selection equation: Local Unemployment rate (coeff: 0.018, p-value: 0.348);
indicator for EMA pilot area (coeff: 0.031, p-value: 0.622); indicator for EMA pilot area*indicator for older siblings (coeff: -0.027, p-value: 0.701)
Standard errors clustered at the Local Education Authority Level. 90% confidence intervals

Table A.3: Perceived likelihood of being admitted to university, given a hypothetical application, wave 1.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Main equation			
IDACI	0.206***	0.184**	-0.149*	-0.287***	-0.004	0.021	0.015
	(0.063)	(0.078)	(0.080)	(0.082)	(0.084)	(0.084)	(0.084)
FSM		-0.007	-0.101***	-0.120***	-0.053	0.874***	0.857***
		(0.032)	(0.035)	(0.035)	(0.040)	(0.285)	(0.285)
KS2					0.111***	0.118***	0.116***
					(0.007)	(0.007)	(0.007)
FSM*KS2						-0.036***	-0.035***
						(0.011)	(0.011)
Old Siblings							-0.057**
							(0.025)
				Selection equation			
Local Unemployment rate	0.127***	0.136***	0.066***	0.051***	0.024	0.024	0.022
	(0.013)	(0.013)	(0.015)	(0.017)	(0.017)	(0.017)	(0.019)
IDACI	-1.020***	-0.771***	-0.868***	-0.775***	-0.121	-0.095	-0.125
	(0.136)	(0.147)	(0.131)	(0.127)	(0.134)	(0.135)	(0.138)
EMA							0.187**
							(0.083)
EMA*Old Siblings							-0.220***
							(0.074)
FSM		-0.203***	-0.327***	-0.325***	-0.180***	0.670*	0.635*
		(0.048)	(0.048)	(0.049)	(0.052)	(0.346)	(0.354)
KS2					0.128***	0.135***	0.133***
					(0.006)	(0.006)	(0.006)
FSM*KS2						-0.035**	-0.034**
						(0.014)	(0.014)
Old Siblings							-0.071*
							(0.041)
Demographics	No	No	Yes	Yes	Yes	Yes	Yes
GOR fixed effects	No	No	No	No	Yes	Yes	Yes
N	13884	10493	10136	10136	9763	9763	9758

Source: University College London. UCL Institute of Education. Centre for Longitudinal Studies: Next Steps: Sweeps 1 and 8, 2001 and 2016: Geographical Identifiers, 2001 Census Boundaries: Secure Access [computer file], Colchester, Essex: UK Data Archive [distributor], June 2017.
SN: 8189. Next Steps: Sweeps 1-8, 2004-2016: Secure Access. [data collection]. 4th Edition. UK Data Service, 2018.
SN: 7104. http://doi.org/10.5255/UKDA-SN-7104-3
Coefficients from ordered probit models with sample selection (heckprobit).
Demographics are: ethnicity, gender, age, number of siblings.
Standard errors clustered at the Local Education Authority level.
*** 1 percent, *** 5 percent, * 10 percent

Table A.4: Perceived likelihood of being admitted to university, given a hypothetical application, wave 2.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Main equation							
IDACI	0.085	0.111	-0.355***	-0.503***	-0.193**	-0.145*	-0.158*
	(0.069)	(0.075)	(0.095)	(0.089)	(0.086)	(0.087)	(0.087)
FSM		-0.011	-0.110***	-0.127***	-0.034	1.676***	1.723***
		(0.032)	(0.036)	(0.035)	(0.045)	(0.318)	(0.323)
KS2					0.117***	0.131***	0.129***
					(0.008)	(0.008)	(0.009)
FSM*KS2						-0.066***	-0.068***
						(0.012)	(0.012)
Old Siblings							-0.133***
							(0.028)
Selection equation							
Local Unemployment rate	0.124***	0.132***	0.057***	0.037*	0.014	0.014	0.018
	(0.018)	(0.018)	(0.018)	(0.019)	(0.019)	(0.019)	(0.019)
IDACI	-1.037***	-0.863***	-1.033***	-0.981***	-0.358***	-0.329***	-0.370***
	(0.145)	(0.138)	(0.124)	(0.120)	(0.127)	(0.127)	(0.129)
EMA							0.031
							(0.063)
EMA*Old Siblings							-0.027
							(0.071)
FSM		-0.147**	-0.228***	-0.234***	-0.077	1.036***	1.030***
		(0.061)	(0.058)	(0.057)	(0.060)	(0.268)	(0.270)
KS2					0.139***	0.148***	0.146***
					(0.005)	(0.006)	(0.005)
FSM*KS2						-0.046***	-0.046***
						(0.011)	(0.011)
Old Siblings							-0.090*
							(0.046)
Demographics	No	No	Yes	Yes	Yes	Yes	Yes
GOR fixed effects	No	No	No	No	Yes	Yes	Yes
N	12045	10548	10326	10326	9952	9952	9774

Source: University College London. UCL Institute of Education. Centre for Longitudinal Studies: Next Steps: Sweeps 1 and 8, 2001 and 2016: Geographical Identifiers, 2001 Census Boundaries: Secure Access [computer file], Colchester, Essex: UK Data Archive [distributor], June 2017.
SN: 8189. Next Steps: Sweeps 1-8, 2004-2016: Secure Access. [data collection]. 4th Edition. UK Data Service, 2018. SN: 7104. http://doi.org/10.5255/UKDA-SN-7104-3
Coefficients from ordered probit models with sample selection (heckprobit).
Demographics are: ethnicity, gender, age, number of siblings.
Standard errors clustered at the Local Education Authority level.
*** 1 percent, *** 5 percent, * 10 percent

Table A.5: Perceived likelihood of being admitted to university, given a hypothetical application, wave 2, KS3 grades.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Main equation			
IDACI	0.085	0.111	-0.355***	-0.503***	0.045	0.089	0.084
	(0.069)	(0.075)	(0.095)	(0.089)	(0.083)	(0.085)	(0.084)
FSM		-0.011	-0.110***	-0.127***	0.040	1.270***	1.322***
		(0.032)	(0.036)	(0.035)	(0.042)	(0.233)	(0.241)
KS3					0.082***	0.090***	0.090***
					(0.005)	(0.005)	(0.005)
FSM*KS3						-0.039***	-0.041***
						(0.007)	(0.007)
Old Siblings							-0.101***
							(0.027)
				Selection equation			
Local Unemployment rate	0.124***	0.132***	0.057***	0.037*	0.031*	0.032*	0.037**
	(0.018)	(0.018)	(0.018)	(0.019)	(0.018)	(0.018)	(0.018)
IDACI	-1.037***	-0.863***	-1.033***	-0.981***	-0.097	-0.076	-0.115
	(0.145)	(0.138)	(0.124)	(0.120)	(0.141)	(0.139)	(0.141)
EMA							0.028
							(0.064)
EMA*Old Siblings							-0.036
							(0.069)
FSM		-0.147**	-0.228***	-0.234***	0.007	0.830***	0.828***
		(0.061)	(0.058)	(0.057)	(0.063)	(0.199)	(0.202)
KS3					0.108***	0.113***	0.112***
					(0.003)	(0.003)	(0.003)
FSM*KS3						-0.029***	-0.029***
						(0.007)	(0.007)
Old Siblings							-0.076*
							(0.046)
Demographics	No	No	Yes	Yes	Yes	Yes	Yes
GOR fixed effects	No	No	No	No	Yes	Yes	Yes
N	12045	10548	10326	10326	10231	10231	10043

Source: University College London. UCL Institute of Education. Centre for Longitudinal Studies: Next Steps: Sweeps 1 and 8, 2001 and 2016: Geographical Identifiers, 2001 Census Boundaries: Secure Access [computer file], Colchester, Essex: UK Data Archive [distributor], June 2017.
SN: 8189. Next Steps: Sweeps 1-8, 2004-2016: Secure Access. [data collection]. 4th Edition. UK Data Service, 2018. SN: 7104. http://doi.org/10.5255/UKDA-SN-7104-3
Coefficients from ordered probit models with sample selection (heckprobit).
Demographics are: ethnicity, gender, age, number of siblings.
Standard errors clustered at the Local Education Authority level.
*** 1 percent, *** 5 percent, * 10 percent