Great expectations.
Channels and barriers to university education

Very preliminary draft: please do not quote

September 15, 2012

Abstract

In spite a great expansion in higher education, in the UK access to university is still quite unequal, with the majority of the student population coming from relatively well-off families. Since this gap seems to emerge at or before applying to university, a few initiatives have been introduced insure that children for any socio-economic background hold high expectations towards university education. This paper uses the Longitudinal Study of Young People in England (LSYPE), in combinations with a rich set of administrative data and a novel estimation technique, to investigate the formation of educational expectations for early teen-agers. In particular, we focus on the choice of applying to university and we highlight the role played by the costs of university education in conjunction with expectations on the probability of being admitted for university if apply. We find that there are psychological as well as economic reason which prevent young people from poor backgrounds from applying to university. However, our research concludes that policies aimed at reducing the cost of education for disadvantaged students can be effective in increasing the participation rate among this group.

[1]We thank Emilia del Bono, Elena Fumagalli, Steve Pudney and the participants at the 3^{rd} IWAAE conference, Vibo Valentia (Italy) and the 2012 ID workshop, Collegio Carlo Alberto, Moncalieri (Italy) for useful comments and suggestions

1 Introduction

In spite of a great expansion in higher education in most industrialized countries, access to university education is still quite unequal, with the majority of the student population coming from relatively well-off families (see Blanden and Machin, 2004; Carneiro and Heckman, 2002; Blossfeld and Shavit, 1993; Cameron and Heckman, 2001). For the UK, Anders (2012) shows that the gap emerges at or before the decision to apply, while the probability of attending university given application does not differ much across socio-economic groups. Understanding what leads to this evidence is crucial, and removing the barriers to educational attainment for low-income people is highly desirable both on equity, and on efficiency grounds. Indeed, research shows that poor and rich students differ in the credit constraints they face (Attanasio and Kaufmann, 2009; Kaufmann, 2010), and in a set of 'long run factors' which provide rich students with a better learning environment and more positive educational expectations (Cameron and Heckman, 1999; Carneiro and Heckman, 2002; Cameron and Heckman, 2001; Anders, 2012).

A few initiatives have been put in place to ensure that children from any background hold high expectations towards university education. Examples of these policies are the 'Aim Higher', the 'Social and Emotional Aspects of Learning (SEAL)', and the 'Gifted and Talented' program for the UK, the 'Gear up', and 'I have a dream' program for the USA (see Jerrim, 2011; Duckworth et al., 2009; Cabinet Office, 2008; HM Government, 2009). The novelty of these programs is their having teen-agers or even younger kids as their main target population. In fact, in line with the literature on early investment in human capital (see, for example Cunha and Heckman, 2007, 2008; Cunha and Schennach, 2010; Heckman, 2006; Carneiro and Crawford, 2007), it is believed that targeting 18 years old people might be too late and that early interventions must be designed to successfully break the barriers to university education for all students. In particular, in the UK, great emphasis is given to the transition between year 9 and year 11[1] since, at the end of year 9, English students have to choose the optional subjects they will be studying in years 10 and 11. This is when they are first asked to think about their future and to make choices which could potentially affect their likelihood to go to university.

The literature has provided some evidence in favor of initiatives and stud-

[1] see: http://www.actiononaccess.org/,

ies targeting expectations, since expectations have been found to be strong predictors of educational attainment, even after controlling for demographic characteristics, family background and grades (Jacob and Wilder, 2010 for the USA; Jerrim, 2011: Strand, 2007; Chowdry et al, 2009 and 2010 for the UK; Khoo and Ainley, 2005 for Australia; Attanasio and Kaufmann, 2009 for Mexico; Looker and Thiessen, 2004 for Canada).[2]

However, this literature leaves at least two questions almost unanswered. The first is how to separate educational aspirations from educational expectations, i.e., the desire of pursuing higher education from an objective evaluation of the chances of success in it. The second regards the factors which can prevent highly motivated but poorly economically endowed people from forming high educational expectations. In fact, the literature on credit constraints (Cameron and Heckman, 1999; Carneiro and Heckman, 2002; Cameron and Heckman, 2001) has generally focused more on indicators of educational attainment, rather than on educational expectations, while the literature on educational expectations has almost neglected the role of external factors and financial constraints.

We use data from the Longitudinal Study of Young People in England (LSYPE) to study expectations over the choice of applying to university for young people aged 13-15 (year 9 and 10 at school), because in England this is the main target group for English policies aimed at widening participation in higher education. Early teen-agers are thought to be the youngest possible subjects for studies on expectations, because there is evidence that it is only at that age that educational expectations become quite rational and based on the evaluation of marginal costs and benefits (Morgan, 1998).

The paper contributes to the existing literature in a number of ways. First it studies the formation of expectations over the probability of being admitted for university by focusing on the role played by the perceived likelihood of being accepted to university if apply. In particular, we are interested in studying whether people from different socio-economic background differ in the way their expectations respond to new information on their ability. Being able to study expectations on the likelihood of being admitted for university if apply permits to shed light on a fundamental psychological determinant of the access to higher education which has been almost neglected

[2]However, Looker and Thiessen (2004) focuses on aspirations rather than expectations. On the difference between expectations and aspirations, see: Jacob and Wilder (2010); Madeira (2009); Jerrim (2011).

by the economic literature up to this point. Moreover, this help us investigate whether the educational gap between low and high income families is due to differences in the perceived chances of success if apply or to external factors like direct and indirect costs of university education. Finally, conditioning of application permits to focus on a measure of expectation which is not contaminated by aspirations, which can be hardly achieved otherwise.[3]

The LSYPE collects information on the perceived likelihood of being accepted for university just for those who declare that they are sufficiently likely to apply. The paper corrects for this potentially endogenous selection by applying a new estimation technique (see De Luca and Perotti, 2011). As in standard selection models, exclusion restrictions are desirable. For this purpose, we use the local unemployment rate to proxy for the opportunity costs of being in education (see Reynolds and Pemberton, 2001) and the availability of funding for post-compulsory, non-university education from the Educational Maintenance Allowance. Beside helping with the identification of the models, our exclusion restrictions make a further contribution by providing information on the role of opportunity and monetary costs in shaping people's educational expectations.

The paper is organized as follows. Section 2 reviews the relevant literature. Section 3 presents the problem studied, the characteristics of the English educational system and the available data. Section 4 sketches a simple framework of utility maximization to describe the empirical problem. Section 5 explains the empirical model, discusses the problem of sample selection and presents the specifications used. Section 6 presents and discusses the results. Section 7 concludes.

[3]Attempts to separate the effects on a choice from the pure elicitation of expectations can be found also in Stinebrickner and Stinebrickner (2009) and Delavande and Rohwedder (2011)

2 Literature

The persistent inequality in the educational attainment of low and high income student has been the object of quite a few studies. Two different but not mutually exclusive explanations for such a gap are the existence of credit constraints (Attanasio and Kaufmann, 2009; Kaufmann, 2010), and the role of long-run factors correlated with permanent family income (Cameron and Heckman, 1999; Carneiro and Heckman, 2002; Cameron and Heckman, 2001; Anders, 2012). In fact, high permanent income is usually associated with a better learning environment both inside and outside the family, which fosters the acquisition of cognitive and non-cognitive skills. Indeed, there is evidence showing that long-run factors like ability measured at an early age, parental education or family structure seem to be more important than current income in explaining the gap in attainment (Cameron and Heckman, 1999; Carneiro and Heckman, 2002; Cameron and Heckman, 2001).

Long-run family factors are thought to influence educational attainment through two main channels: ability and expectations. While an extensive literature studies the first channel (see, among others: Cunha and Heckman, 2007, 2008; Cunha and Schennach, 2010; Heckman, 2006; Carneiro and Crawford, 2007), the literature on educational expectations is only at its beginning. Empirical evidence shows that expectations are strongly correlated with a set of individual and familiar characteristics such as: gender, ethnicity, family structure, neighborhood characteristics, socio-economic background, parental expectations (Jacob and Wilder, 2010; Jerrim, 2011; Reynold and Pemberton, 2001; Fowley, 2011; Chowdry et al, 2010; Duckworth et al., 2009). In general, these studies suggest that the most disadvantaged groups are characterized by lower expectations and poorer achievement.

A methodological challenge in the study of educational expectations is unpacking the relationship between low expectations and low achievement. Generally, people assume that expectations respond to grades, which are regarded as signals of students' ability. Most papers address the problem of a possible reverse causation by relying on timing and they assume that if expectations change between two points in time and if new information on grades is released in the same time period, therefore the causation should go from grades to expectations rather than in the opposite directions (see, for example Stinebrickner and Stinebrickner, 2009; Zafar, 2011; Fowley, 2011). Evidence in favor of this is provided by Papay et al., (2011) who exploit a regression discontinuity design on discrete labeling of continuous test scores

to show that getting a more positive label has an effect on the probability of going to college for low-income, urban students. For the UK, Fumagalli (2012) carries out a similar exercise using performance levels in key stage 3 exams and concludes that grades contain new information which is not incorporated in previous expectations.

A growing literature has studied how expectations evolve. Jacob and Wilder (2010) find that expectations do evolve, especially in the case of males, more disadvantaged children and people with low degree of self-concept and locus of control. In general, the literature points out that, when little information on respondents' ability is available, expectations tend to be biased toward high probabilities of attending college (Reynolds and Pemberton, 2001; Jacob and Wilder, 2010) or higher grades (Stinebrickner and Stinebrickner, 2009; Zafar, 2011), especially for blacks (Madeira, 2009) and disadvantaged people (Chowdry et al, 2010; Jacob and Wilder 2010). However, over time, people tend to become more pessimistic in their educational expectations and this is mainly due to low-income people updating their expectations downwards (Chowdry et al, 2010; Jacob and Wilder, 2010).

A few studies have pointed out that people differ in their ways of processing information (Gouret and Hollard, 2011; Dominitz and Manski, 2011; Winswall and Zafar, 2011; Zafar, 2011). Jacob and Wilder (2010) find that ethnic minorities and people from well-off families seem to update their expectations on the basis of the grades they receive, while the most disadvantaged people tend to have quite unresponsive expectations. If this argument holds, the gap in the educational attainment of high and low income people could be due to a different way of updating their expectations when more information is made available. Attanasio and Kaufmann (2009) and Kaufmann (2010) argue that such differential responsiveness could be due to credit constraints, and that low-income people need higher expected returns to education to be induced to attend university. Finally, while recognizing the role of credit constraints, Jensen (2010) suggests that young people living in deprived areas are likely to have little and downward biased information on the expected returns to schooling. In fact, in forming their expectations, they rely on information based on their group of peers and on people living in their local community, which provides role models leading to low achievement.

Indeed, there is contrasting evidence on whether constraints and expected direct and opportunity costs are incorporated into expectations. Jacob and Wilder (2010) conclude that state policies aimed at reducing costs do not seem to be successful in boosting expectations. Along a different line,

Reynolds and Pemberton (2001) link the rise in educational expectations in the 1980s and 1990s to a deterioration of labor market factors and, consequently, to the drop in the opportunity costs of staying in education. Finally, Dinkelman and Martínez (2011) find that exogenous shocks of information on the availability of financial aid for students can have an impact on both expectations and behaviors.

3 Background

3.1 The English Educational System

The English educational system is divided into 'key stages', where each stage reflects 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).

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 college; Further Education (FE)
na	18+	na	ISCED 4-5-6	Further Education (FE) and Higher Education (HE)

[a] Note: the shaded area indicates compulsory education
[b] Source: http://www.unesco.org

All key stages 2, 3, 4 and 5 (from now on: 'K2', 'K3', 'K4', 'K5') end with exams which are graded by external examiners. A certification is awarded after the final exam at the end of key stage 4 and 5, respectively, the 'General Certificate of Secondary Education (GCSE)' and the 'Advanced Level General Certificate of Education (A-level)', with the grades obtained in the exam having practical consequences both in the labor market and in Higher Education. On the contrary, the tests at the end of key stage 2 and 3 are

meant to show whether the children are 'working at, above or below the target level for their age. This helps the school to make plans for their future learning. It also allows schools to see whether they are teaching effectively by comparing their pupils' performance to national results'.[4] For each test they sit,[5] pupils are awarded a score which is summarized into levels (see table 2). The pass score is 21 points for both key stage 2 and key stage 3. This corresponds to level 3. However, most pupils are expected to reach at least level 4 (27 points) at the end of key stage 2 and level 5 (33 points) at the end of key stage 4.

Table 2: Test scores and levels for key stage 2 and key stage 3

levels	scores	Satisfactory for Key Stage 2	Satisfactory for Key Stage 3
E (Exceptional Performance)	57	√	√
8	51	√	√
7	45	√	√
6	39	√	√
5	33	√	√
4	27	√	
3	21		

Source: http://www.education.gov.uk

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. In addition, some measures have been introduced to help people from poor families cover the direct and opportunity cost of being in post-compulsory education. In this respect, one of the main policy initiatives has been the introduction, in year 2004-2005, of the 'Educational Maintenance Allowance' (EMA): a 30£, 20£ or 10£ weekly allowance (The amount of the allowance depends on family income) targeted to young people aged 16-19 and aimed at keeping low-income people in post-compulsory academic education[6] until they reach the age for university education. EMA is means-tested and it is available for pupils whose family income does not exceeds a

[4]see http://www.direct.gov.uk and section 7 of this paper

[5]English, Math and Science for key stage 2, English, Maths, Science, History, Geography, Modern foreign languages, Design and Technology, Information and Communication Technology (ICT), Art and Design, Music, Physical Education, Citizenship, Religious Education for Key Stage 4

[6]starting from school year 2006/2007, it was extended also to those enrolled in E2E and PLA types of courses.

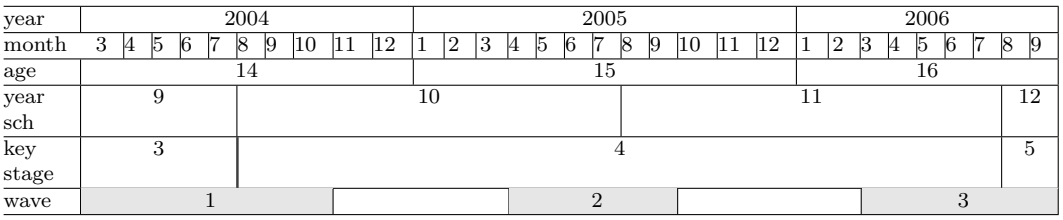
certain threshold (30,000 for school year 2004/2005, 30,810 for the following years), it was introduced at the national level in the academic year 2004-2005 (i.e., immediately after the collection of the first wave of LSYPE) after a 5 year test on 15 randomly selected pilot areas and a 4 year test on additional 26 randomly selected areas.

Free education ends with higher education. The amount of annual tuition fees varies across university and, up to academic year 2011/2012, for British students it could not exceed £3,375. However, starting from the academic year 2012/2013, this cap has been raised to £9,000, with most of the universities charging more than £6,000 a year. All British students are entitled for maintenance students loans which must be repaid after graduation by graduates earning more than £15,000. Non-repayable maintenance mean-tested loans are available for low-income students.

3.2 The LSYPE and its questions on educational expectations

Our analysis is based on the Longitudinal Study of Young People in England (LSYPE), a cohort study containing detailed information for around 15000 pupils living in England. The sample was drawn in 2004 and is composed of young people in year 9 at school who were born between 1^{st} September, 1989 and 31^{st} August, 1990. Figure 1 shows the timing of events experienced by sample members. The shaded cells at the bottom show when the data collection took place and the rows above summarize the median age of respondents, together with the school year and the key stage they are in.

Figure 1: Timeline



The sample design of LSYPE is two-stage probability sample where pri-

mary sampling units are schools. PSUs are stratified by deprivation level and deprived schools are over-sampled. Secondary sampling units are people within school and young people from the main ethnic groups (Indian, Pakistani, Bangladeshi, Black Caribbean, Black African and Mixed) are over-sampled. 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 6 people for the independent sector), boarders and children who are in England just for education purposes. Sample members have been interviewed yearly for six waves. To offset attrition, a boost sample was drawn at wave 4.

LSYPE can be linked to the National Pupil Database (NPD). NPD is an administrative data-set covering all pupils in state schools (from nurseries to secondary schools) in a given year and containing information on pupils' attainment, together with pupils'[7] and schools' characteristics. This information includes pupils' ethnicity, eligibility for free school meals status (a proxy for low income), Special Education Needs (SEN) and history of attendance. Crucially, the NPD contains detailed information on educational attainment for Key stage 2 and Key stage 3 expressed in terms of actual score obtained. Geographical indicators and indices of area characteristics are available. In particular, the dataset includes information on the 2004 Income Deprivation Affecting Children Index (IDACI) at the Lower Layer Super Output Area (LSOA) level.

The LSYPE contains also detailed information on educational expectations. In particular, it contains both a question on the likelihood of applying to university (from now on: question 1) and a question on the likelihood of being admitted given application (from now on: question 2). More specifically, question 1 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'. Let A_{it} be the variable summarizing these four possible outcomes for individual i at time t (year 9 or 10), and $E_{it}(p_i^A)$ a continuous latent variable indicating the expectations over the probability of applying to university in the future. We can write:

[7]Most of the information at pupils' level comes from the Pupil Level Annual School Census (PLASC), which, starting from school year 2006-7 this has been replaced by School Census.

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

The respondents who, in question 1 have answered they are: 'Very likely', 'Fairly likely' or 'Not very likely' to apply to university are asked question 2: 'How likely do you think it is that if you do apply to go to university you will get in?'. Let now $E_{it}(p_i^a)$ be the continuous latent variable measuring the expectations at at time t over the probability of being accepted for university and a_{it} its categorical counterpart, which can be defined:

$$a_{it} = \begin{cases} \text{'very likely' if } \mu_3 < E_{it}(p_i^a) \leq \mu_4. \\ \text{'fairly likely' if } \mu_2 < E_{it}(p_i^a) \leq \mu_3. \\ \text{'not very likely' if } \mu_1 < E_{it}(p_i^a) \leq \mu_2. \\ \text{'not at all likely' if } \mu_0 \leq E_{it}(p_i^a) \leq \mu_1. \end{cases} \tag{2}$$

However, due to the question routing, a_{it} is not observed for those with A_{it} = 'not at all likely'. Let S_{it} be the variable describing the selection into question 2, it can be written:

$$S_{it} = \begin{cases} 0, & \text{if } \alpha_0 \leq E_{it}(p_i^A) \leq \alpha_1. \\ 1, & \text{otherwise.} \end{cases} \tag{3}$$

Therefore, the perceived likelihood of being accepted for university of apply can only be observed as follows:

$$y_{it} = \begin{cases} \text{missing}, & \text{if } \alpha_0 \leq E_{it}(p_i^A) \leq \alpha_1. \\ a_{it}, & \text{otherwise.} \end{cases} \tag{4}$$

This creates a selection problem which will be addressed later on in the paper.

3.3 Educational Expectations on a glance

So far we have argued that the characteristics of the psychological development of the teen agers, together with the features of the English educational system, make it particularly interesting to study the formation of educational expectations between year 9 and 10. The LSYPE permits to have an idea on how persistent these expectations are and when such a persistence arises.

Table 3: Perceived likelihood of applying to university: transition matrices between year 9 and year 10 (top panel), between year 10 and 11 (middle panel), and between year 11 and 12 (bottom panel)

			not interviewed	don't know	very likely	fairly likely	not very likely	not at all likely	cell size
			likelihood of applying to uni, year 10						
likelihood of	applying to	uni, year 9							
		don't know	15.29	13.09	10.74	25.15	18.38	17.35	680
		very likely	14.63	1.64	57.09	21.48	3.62	1.53	5,414
		fairly likely	13.27	3.71	21.64	41.53	14.20	5.64	5,304
		not very likely	14.81	4.45	4.74	20.38	34.86	20.76	2,404
		not at all likely	19.34	3.31	2.95	7.00	22.59	44.81	1,629
		Total	14.72	3.47	28.99	26.84	14.78	11.20	15,431
			likelihood of applying to uni, year 11						
likelihood of	applying to	uni, year 10							
		don't know	8.84	15.47	12.15	28.18	18.97	16	543
		very likely	7	1.33	72.25	15.86	2.03	2	4,526
		fairly likely	7.92	3.25	29.15	41.01	12.87	5.79	4,179
		not very likely	9.69	4.8	5.45	19.52	34.75	25.79	2,311
		not at all likely	12.34	2.84	2.79	6.94	18.32	56.77	1,758
		Total	8.54	3.31	35.51	23.71	13.95	14.97	13,317
			likelihood of applying to uni, year 12						
likelihood of	applying to	uni, year 11							
		don't know	0.99	9.63	18.52	27.9	22.47	20.49	405
		very likely	0.9	0.9	79.22	13.45	3.48	2.06	4,460
		fairly likely	1.05	2.34	37.36	34.47	16.61	8.17	2,950
		not very likely	1.94	2.12	6.13	17.5	37.36	34.94	1,697
		not at all likely	2.76	1.44	2.43	5.91	21.22	66.24	1,810
		Total	1.4	1.85	42.91	18.85	15.49	19.5	11,322

Table 3 shows how the perceived likelihood of applying to university in

the future changes during years 9, 10 and 11 (waves 1 and 2 and 3, see figure 1). All transition matrices show that expectations on the likelihood of applying to university are relatively fluid between year 9 and 10, but become more persistent since year 10. In fact, the share of those who answer they are 'very likely' to apply to university in two subsequent waves is equal to 57.09% (66.88% of the interviewed) between year 9 and 10, 72.25% (77.69% of the interviewed) between year 10 and 11 and equal to 79.22% (79.93% of the interviewed) between years 11 and 12. Similarly, the share of those who, in two subsequent waves, answer they are 'not at all likely to apply to university' is equal to 44.81% (55.56% of the interviewed) between year 9 and 10, while it goes up to 56.77% (64.76% of the respondents) between years 10 and 11 and to 66.24% (68.13% of the interviewed) between years 11 and 12.

As previously stated, the transition between year 9 and 10 is meaningful because at the end of year 9 the young people receive a signal of their ability in the form of key stage 3 data (see tables 1 and figure 1), which is likely to affect the expected likelihood of being admitted for university given application. Table 4 shows the relationship at years 9 and 10 between these two measures of educational expectations. For the cases where both questions are answered, most of the results are on the diagonal, showing that the two events are indeed positively related. However, the expectations on the likelihood of being accepted are somewhat less extreme and this could reflect the uncertainty due to the role of universities in determining acceptance and the fact that this measure is not contaminated by aspirations. Crucially, those who answer they are 'not at all likely to apply to university' (more than 10% at year 9 and about 13% at year 10) do not even get to the question on the probability of being accepted (column 'not applicable'). This suggests that selection is unlikely to be ignorable.

Table 5 shows the changes between year 9 and 10 in the perceived likelihood of being accepted for university given application. This is conditional on being a respondent at year 9. Attrition is close to 15% (see column 'not interviewed'), but it does not seem to differ across groups defined according to their answer at the previous wave. 42.92% of those who, due to the question routing, do not get to the question on the probability of being accepted for university (52.43% among the respondents only), fail to get to the question also in the following year. However, about 39% (47% among the respondents only) of the people who were not asked the question on expectations on the likelihood of being accepted at age 9 do get to answer the question at year

Table 4: likelihood of applying to university and likelihood of being accepted if apply, year 9 (top panel) and 10 (bottom panel)

likelihood of applying to uni, year 9	not applicable	don't know	very likely	fairly likely	not very likely	not at all likely	cell size
	likelihood of being accepted if apply, year 9						
don't know	100	0	0	0	0	0	680
very likely	0	3.77	45.92	48.78	1.39	0.15	5,414
fairly likely	0	7.41	6.13	76.23	9.69	0.55	5,304
not very likely	0	10.57	2.16	35.36	45.92	5.99	2,404
not at all likely	100	0	0	0	0	0	1,629
Total	14.96	5.51	18.55	48.82	10.97	1.17	15,431
	likelihood of being accepted if apply, year 10						
don't know	100	0	0	0	0	0	543
very likely	0	3.25	48.74	46.97	0.93	0.11	4,526
fairly likely	0	7.85	7.27	74.73	9.57	0.57	4,179
not very likely	0	10.39	1.95	32.58	47.77	7.31	2,311
not at all likely	100	0	0	0	0	0	1,758
Total	17.28	5.37	19.19	45.07	11.61	1.49	13,317

10. Moreover, about 17.8% of them (21.75% among the respondents) answer they are at least 'fairly likely' to be admitted if apply, suggesting that the probability of applying to university and the probability of being accepted evolve over time and are partially driven by different factors. Moreover, this suggests that at lest some of the people who answer they are not at all likely to apply to university are not completely discouraged students. As a consequence, failing to control for sample selection would leave out of the analysis most of the 'marginal' students, which is precisely the group policy should target.

The share of those giving the same answer in both year 9 and 10 is equal to 42.65% (50.18% among the interviewed) for those who think it is 'very likely' they get accepted for university, 52.10% (60.15% among the interviewed) for those who think it is 'fairly likely', but only 7.73% (8.97% among the interviewed) for those who think it is 'not at all likely'. Instead, 40.88% (47.44% among the interviewed) of this last group fell into inapplicability at wave 10. In sum, the table suggests that expectations are quite persistent for those who attach a high probability to the event of being accepted for university, however, among those with low levels of expected likelihood of being accepted for university, some updating takes place and it goes together with the decision of not applying for a degree.

Table 5: likelihood of being accepted given application: transition matrices between year 9 and year 10

likelihood of being accepted if apply , year 9	likelihood of being accepted year 10: not interviewed	not applicable	don't know	very likely	fairly likely	not very likely	not at all likely	cell size
not applicable	18.15	42.92	5.24	3.55	14.25	13.17	2.73	2,309
don't know	13.87	19.27	11.75	9.05	33.84	10.81	1.41	851
very likely	15.02	3.18	2.31	42.65	34.65	2.10	0.10	2,863
fairly likely	13.39	7.25	4.35	14.68	52.10	7.59	0.64	7,534
not very likely	15.95	23.57	4.84	1.42	23.69	27.47	3.07	1,693
not at all likely	13.81	40.88	3.87	2.21	12.71	18.78	7.73	181
Total	14.72	14.68	4.56	16.29	38.61	9.90	1.24	15,431

4 The formation of educational expectations

We now describe our empirical problem through a simple framework of utility maximization. This will also help justify the variables and the models used in the empirical specification. To mimic what found in LSYPE data, we make expectations over the choice to apply to university in ($E_t(p^A)$, as defined in (1)) to depend on the perceived likelihood to be admitted for university given application ($E_t(p^a)$, as described in (2)). For simplicity, we assume that everybody applies to university when she or he is 18. For ease of exposition, we first consider the simpler scenario where expectations are formed at age 18 and then we extend it to the case of younger people (years 9 and 10).

4.1 Expectations at age 18

Consider an individual who lives until age T. At age 18, he or she can make two choices: to apply to university ($A = 1$) or not ($A = 0$). The individual applies to university if the expected life-time utility from applying to university exceeds the expected life-time utility from not applying, i.e., if:

$$E_{18}(U(A=1)) > E_{18}(U(A=0)) \tag{5}$$

where:

$$\begin{aligned} E_{18}(U(A=1)) \quad &= E_{18}(p^a)\left(\int_{18}^{T} E_{18}(r_\tau^e - c_\tau^e)\exp^{-\rho(\tau-18)}\, d\tau \right) - c^A + \\ &+(1 - E_{18}(p^a))\left(\int_{18}^{T} E_{18}(r_\tau^{\neg e} - c_\tau^{\neg e})\exp^{-\rho(\tau-18)}\, d\tau \right) \end{aligned} \tag{6}$$

$$E_{i18}(U(A=0)) = \int_{18}^{T} E_{18}(r_\tau^{\neg e} - c_\tau^{\neg e})\exp^{-\rho(\tau-18)}\, d\tau \tag{7}$$

Again, $E_{18}(p^a)$ is the expected probability at age 18 of being accepted for university. r_τ^e and c_τ^e, and $r_\tau^{\neg e}$ and $c_\tau^{\neg e}$ are the returns to and the costs of being and not being in education at time $\tau : 18 \leq \tau \leq T$, $\exp^{-\rho(\tau-18)}$ is a discounting factor and c^A are the costs of applying to university. We model these costs as sunk costs at age 18.

In order to form $E_{18}(p^a)$, the individual relies on signals of his ability (I_{18}). These take the form of grades up to age 18. We can write:

$$E_{18}(p^a) = f(I_{18}, B) \tag{8}$$

where B are behavioral characteristics driving the formation of expectations (i.e. personality)[8]

4.2 Expectations at an earlier age

We now model the formation of $E_t(p^a)$: the expectations over the probability of being accepted for university 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.

Similar to the previous case, at t, in forming his expectations on the future action A_t, the individual compares the expected life-time utility of applying to university and the expected life-time utility of not applying. Crucially, at time t, I_{18} is not observed, but it is possible to form expectations over it based on I_t, i.e. $E_t(I_{18}|I_t)$. As a consequence, $E_t(p^a)$ can be written as:

$$E_t(p^a) = f(E_t(I_{18}|I_t), B) \tag{9}$$

Given that 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, the expected probability at t of applying to university ($E_t(p(A = 1)) = E_t(p^A)$) is the probability attached to the event:

$$\begin{aligned}
& E_t(E_{18}(U(A=1))) - E_t(E_{i18}(U(A=0))) = \\
= \;& E_t(U(A=1)) - E_t(U(A=0)) = \\
= \;& \int_{17}^{18} E_t(r_\tau^{\neg e} - c_\tau^e - r_\tau^{\neg e} + c_\tau^{\neg e}) \exp^{-\rho(\tau - t)} d\tau - \\
- \;& E_t(c^A) \exp^{-\rho(18-t)} + \\
+ \;& E_t(p^a) \int_{19}^{T} E_t(r_\tau^e - c_\tau^e - r_\tau^{\neg e} + c_\tau^{\neg e}) \exp^{-\rho(\tau - t)} d\tau > 0
\end{aligned} \tag{10}$$

Where $\int_{17}^{18} E_t(r_\tau^{\neg e} - c_\tau^e - r_\tau^{\neg e} + c_\tau^{\neg e}) \exp^{-\rho(\tau - t)} d\tau$ is the difference in the expected net returns of being and not being in post compulsory general education; $E_t(c^A) \exp^{-\rho(18-t)}$ are the expected costs of applying to university and $E_t(p^a) \int_{19}^{T} E_t(r_\tau^e - c_\tau^e - r_\tau^{\neg e} + c_\tau^{\neg e}) \exp^{-\rho(\tau - t)} d\tau$ is the expected value of

[8] I assume them time invariant

the difference between expected costs and returns from age 19 onwards in the two scenarios of acquiring and not acquiring university education.

The profile of the net returns of choosing general education is generally steeper than alternative options, because students have to forgone earlier returns to be able to have higher returns in the future. Whether these future returns will be enough to compensate the loss in the short run depends crucially on the probability of being accepted for university given application ($E_t(p^a)$) which is, indeed, the core of our analysis.

5 The empirical model

5.1 The sample selection problem

Equation (9) in section (4.2) shows the determinants of $E_{it}(p_i^a)$. Crucially, just few of its components are observed by the econometrician. While I_i^t, i.e, grades up to time t, are observed, the expectations on grades following t ($E_{it}(I_{i18}|I_{it})$), and B_i are not observed. As a consequence, $E_{it}(p_i^a)$ for individual i can be written, as a function of the observed variables, as follows:

$$E_{it}(p_i^a) = \beta_0 + I_i^t\beta_1 + \mathbf{x}_{it}\boldsymbol{\beta_2} + \nu_{it} \tag{11}$$

Where $\mathbf{x}_{it}$ are individual, family, school and time specific characteristics. Taken together, the variables $\mathbf{x}_{it}$ are meant to capture the effect of $E_{it}(I_{i18}|I_{it})$. Finally, ν_{it} is an error term. Similarly, $E_{it}(p_i^A)$ is defined by (6). It can be written as follows:

$$E_{it}(p_i^A) = f(E_{it}(p_i^a), z_{it}) \tag{12}$$

where z_{it} are variables predicting the returns and the costs, both direct and indirect (opportunity costs), of being in education (in the theoretical framework $c_{i\tau}^e$, c_i^a and $r_{i\tau}^{\neg e}$). In line with (9) and (11), these variables do not enter into the equation for $E_{it}(p_i^a)$, but they do enter into the equation of $E_{it}(p_i^A)$. In fact, $E_{it}(p_i^a)$ is an important but not the unique determinant of $E_{it}(p_i^A)$. Writing (12) in the usual linear form, we have:

$$E_{it}(p_i^A) = \gamma_0 + I_{it}\gamma_1 + \mathbf{x}_{it}\boldsymbol{\gamma_2} + z_{it}\boldsymbol{\gamma_3} + \upsilon_{it} \tag{13}$$

(11) and (13) can be written as:

$$E_{it}(p_i^a) = X_{it}\beta + \nu_{it} \tag{14}$$

and

$$E_{it}(p_i^A) = W_{it}\gamma + \upsilon_{it} \tag{15}$$

with

$$\beta = [\beta_0, \beta_1, \beta_2] \tag{16}$$

$$\gamma = [\gamma_0, \gamma_1, \gamma_2, \gamma_3] \tag{17}$$

$$X_{it} = [I_{it}, \mathbf{x}_{it}] \tag{18}$$

$$W_{it} = [X_{it}, z_{it}] \tag{19}$$

and

$$X_{it} \subset W_{it} \tag{20}$$

Since the question on the probability of being admitted to university given application (question 2) is not asked to those having $\alpha_0 \leq E_{it}(p_i^A) \leq \alpha_1$, the error in (11) will have a non-zero conditional expectation. Formally, given

$$Cov(\upsilon_{it}, \nu_{it}) \neq 0 \tag{21}$$

we have

$$E(\nu_{it}|X_{it}, S_{it} = 1) \neq 0 \tag{22}$$

This means that, if selection is not taken into account, β will be unconsistently estimated.

5.2 Our specification

The equations in section 5.1 can be estimated through the following bivariate threshold crossing model, which permits to overcome the problem of endogenous selection:

$$S_{it} = I(\alpha_0 < E_t(p_i^A) \leq \alpha_1) \tag{23}$$

$$y_{it} = \sum_{l=1}^{4} lI(\mu_l < E_t(p_i^a) \leq \mu_{l+1}) \quad if \quad S_{it} = 1 \tag{24}$$

where $\mu_l < \mu_{l+1}$, $\mu_0 = -\infty$, $\mu_5 = \infty$[9] and the latent variables for the selection process (23) and for the model for the variable of interest (24) are, respectively, those defined in (15) and (14).

To estimate the model, we follow the parametric approach of de Luca and Perotti (2011) (from now on: 'opsel model') which corrects for selection by modeling the correlation between the errors of the two latent variables, namely ν_{it} and υ_{it} in equations (14) and (15). Two restrictions are needed, i.e: i) β_0 is set to 0, given that it can not be separately identified from the thresholds μ.[10] ii) There is at least a variable which is contained in W_{it}, but not in X_{it}. This is in line with (although not equal to) what implied by equation (20).

The model is estimated by maximum likelihood. Being y_{it} and S_{it} discrete, it is possible to write the contribution to the maximum likelihood for all the 5 possible realizations of the two variables, i.e, $(S_{it} = 0)$, $(S_{it} = 1, y_{it} = 1)$, ... , $(S_{it} = 1, y_{it} = 4)$. In order to be able to write such contributions, de Luca and Perotti (2011) adopt a standard approach in literature (see, for example Miranda and Rabe-Hesketh, 2006) and assume that ν_{it} and υ_{it} follow a bivariate gaussian distribution with zero mean, variance equal to 1 and correlation coefficient equal to ρ.

Let $\theta = (\beta, \gamma, \mu, \rho,)$ be the vector of parameters which need to be estimated, the likelihood is:

[9] We can think at the underlying probabilities as normalized variables between $-\infty$ and ∞

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

$$L(\theta) = \sum_{i=1}^{n} \left[(1 - S_{it}) \ln \phi_{0i}(\theta) + \sum_{l=0}^{4} S_{it} I(y_{it} = l) \ln \phi_{1li}(\theta) \right] \tag{25}$$

where $\phi = (\phi_0, \phi_{10}, ..., \phi_{14})$ is the vector of conditional probabilities for the 5 events.

Let Φ and Φ_2 be, respectively, the standard and the bivariate Gaussian distributions, therefore, ϕ can be written:

$$\phi_0 = Pr(S_{it} = 0) = 1 - \Phi(W_{it}\gamma) \tag{26}$$

$$\begin{aligned} & \phi_{1l} = Pr(S_{it} = 1, y_{it} = l) \\ = \; & \Phi_2(W_{it}\gamma; \mu_{l+1} - X_{it}\beta; -\rho) - \Phi_2(W_{it}\gamma; \mu_l - X_{it}\beta; -\rho) \end{aligned} \tag{27}$$

We chose variables in X_{it} in line with both our framework in section 4 and most of the relevant empirical literature. Each model is estimated in three different specifications.

Specification 1 one uses basic demographics, together with some information capturing heterogeneity at the family and at the area level. In particular, we include age, gender, ethnicity[11] which are thought to be correlated with educational expectations (see, for example, Jacob and Wilder, 2010; Duckworth et al., 2009; Strand, 2007; Khoo and Ainley, 2005; Looker and Thiessen, 2004, area fixed effects (Government Office Regions) (Jacob and Wilder, 2010). As proxies for socio-economic status (see Jerrim, 2011; Chowdry et al. 2009; 2010; Jacob and Wilder, 2010; Duckworth et al., 2009), we include a dummy variable indicating free school meal eligibility and the index of deprivation affecting children at the LSOA. Finally, all specifications control for month of interview to take into account the seasonality induced by different stages in the school year and in the data collection.

In specification 2, we add some variables to best control for the socio-economic status of the family, namely whether the family lives in a urban area (Papay et al., 2011; Koo and Ainley, 2005), whether the family lives in a owned house and the number of cars it owns. Moreover, we include education, employment status and age of the mother, whether any of the

[11] As an additional control for the ethnic background, we included whether the respondent was born in the UK.

grandparents has a degree (Jerrim, 2011; Chowdry et al., 2009; 2010; Koo and Ainley, 2005).[12]

Finally, in specification 3, we add measures of behavioral characteristics (B in section 4). Wave 1 contains two batteries of questions about the future (see table 17 in the Appendix). In both cases, the respondent is asked how much he agrees with some statements suggested by the interviewer. We used factor analysis to aggregate these answers into two indicators that we called 'absence of planning' for question A and 'ambition' for question B (see 'factor 1' in the table). The degree of planning can shed some light on people's discount rate (ρ in section 4), while 'ambition' is correlated with how people transform future expected costs and benefits into utility. Before carrying out the factor analysis, chained equation imputation (see Raghunathan et al., 2001) was used to impute the missing sub-items for the item respondents.[13]

For most of the models above, we include as predictors either key stage 2 or key stage 3 grades. Moreover, in a model we add a set of dummy variables indicating whether the respondent got his best K3 score in English, Math or Science. This is meant to capture the heterogeneity in both the type of signals young people get and in the acceptance rate for university (and, as a consequence, in the probability of admission) across different subjects. Finally, in some specifications, grades are interacted with a dummy variable indicating free school meal eligibility. In this, our strategy is similar to the one adopted by Attanasio and Kaufmann (2009) and it is meant to investigate whether poor people are less responsive than rich people to information incorporated in grades. If the difference between rich and poor people is due to credit constraints only, we expect the interaction to have a negative and significant sign only in the case of the expected probability of applying to university. On the contrary, if the coefficient of the interaction is negative

[12]We chose to control for mother's characteristics because usually thought to be more relevant than fathers' in predicting children's outcomes (see Leibowitz, 1974; Heckman and Hotz, 1986; Haveman and Wolfe, 1995; Black, Devereux, and Salvanes, 2005). LSYPE distinguishes between 'main parents', i.e. the parents 'most involved in the young person's education' (NatCen, 2009) and 'second parents' or whom less information is available. The parents are free to choose who want the main parent to be, although in most cases (almost 90%) this role is taken up by mothers. As a consequence, the information we have on mothers is much more complete and precise than the information we have on fathers. A dummy variable is included to take into account this non random selection of main parents and another dummy variable indicates whether the mother is not present in the family

[13]Details on the imputation procedure and on the construction of factors can be found in the appendix

and significant also in the main equation, our model will provide evidence in favor of the hypothesis by Jensen (2010) who claims that poor people have different role models which leads them to underestimate their chances of success or their future returns.

The variables presented so far are used in both equation (14) and (15). In addition, equation (15) includes also the variables z_{it}, which are excluded from equation (14). As first exclusion restriction, we use the unemployment rate of the local authority where the respondent resides, which helps predict $E_t(r^e_{i\tau})$ and $E_t(r^{\neg e}_{i\tau})$ for the local labor market (see Reynolds and Pemberton, 2001). The underlying assumption here is that differences in the local unemployment rate do not affect the probability of being accepted for university given application. As a second exclusion restriction we use a dummy variable indicating whether the respondent is aware of the existence of the Educational Maintainance Allowance (EMA) and he thinks he is eligible for that. In terms of our framework, the variable helps predict $c^e_{it} + \int_{t+1}^{\bar{t}-1} E_{it}(c^e_{i\tau}) \exp^{-\rho\tau} d\tau$: the direct costs of staying in education up to time $\bar{t}$. The question we use is asked since wave 2, which corresponds to the first year of universal application of EMA.

Descriptive statistics of the variables and the sample used are shown in the appendix (tables 16 and 17).

6 Results

Table 6 shows selected coefficients for the model in equations (23)-(24) estimated on wave 1 data (expectations at year 9).[14] For each specification, two columns are presented. The column 'selection' uses as dependent variable the variable S_{it} defined in (3) and focuses on the probability that $S_{it} > 1$, i.e. that the respondent answers that she or he is at least 'not very likely' to apply to university. The column 'accepted' shows the determinants of a_{it} defined in (2), i.e. the ordered variable summarizing the perceived likelihood of being accepted for university if apply and polished from endogenous selection. In fact, selection is taken into account by the opsel models and the perceived likelihood of being admitted for university can be interpreted as representative of the sample as a whole.

White UK-born British, and people from low-educated families[15] are more likely to have less positive expectations on both the likelihood of apply to university and the likelihood of being admitted given application. Coefficients associated to mother's employment status are not significant (see appendix). The other variables have generally a different behavior in the selection and in the main equation. For example, relatively older people and people living in London and in the North (the omitted area variable is 'South West') are found to have higher expectations on the probability of being admitted for university, but they do not seem to be more likely to apply to university. On the contrary, females seem to be more likely to apply, but they do not seem to be more likely to be admitted once applied.

Low-income people and people living in a deprived neighborhood seem to be less likely to have high educational expectations when few controls for familiar wealth and education are included. However, when these controls are added, the negative signs of both variables disappear in the case of the probability of being accepted once application. The estimated coefficient of both variables is always negative and significant in the selection equation, although the coefficient for free school meal eligibility goes towards zero in specifications 2 and 3, while the coefficient for the idaci index remains negative and significant in all specifications. This seems to suggest that the expected likelihood of of being admitted for university is not the main channel through which income affects the perceived likelihood of applying.

[14] The rest of the coefficients can be found in the appendix

[15] For a discussion of the role of parental education, see, for example, Jacob and Wilder (2010)

Table 6: Likelihood of being accepted for university (opsel model, year 9, selected coefficients)

	1		2		3	
	accepted	selection	accepted	selection	accepted	selection
Age	0.111 ***	-0.007	0.110 ***	-0.009	0.099 ***	-0.014
	(0.027)	(0.039)	(0.027)	(0.040)	(0.027)	(0.041)
Female	0.034	0.208 ***	0.041	0.212 ***	0.030	0.194 ***
	(0.026)	(0.037)	(0.026)	(0.038)	(0.026)	(0.039)
Asian	0.479 ***	0.835 ***	0.491 ***	0.924 ***	0.486 ***	0.961 ***
	(0.050)	(0.075)	(0.050)	(0.082)	(0.050)	(0.085)
Black	0.475 ***	0.784 ***	0.452 ***	0.749 ***	0.392 ***	0.740 ***
	(0.070)	(0.133)	(0.067)	(0.138)	(0.067)	(0.142)
Other	0.199 ***	0.397 ***	0.175 ***	0.387 ***	0.152 ***	0.354 ***
	(0.056)	(0.089)	(0.055)	(0.093)	(0.055)	(0.093)
N. East	0.262 ***	-0.123	0.290 ***	-0.126	0.274 ***	-0.123
	(0.073)	(0.106)	(0.075)	(0.108)	(0.075)	(0.110)
N. West	0.204 ***	-0.170 **	0.191 ***	-0.211 ***	0.175 ***	-0.235 ***
	(0.055)	(0.076)	(0.057)	(0.079)	(0.057)	(0.080)
York. & Hum.	0.185 ***	-0.159 *	0.193 ***	-0.179 **	0.203 ***	-0.184 **
	(0.059)	(0.082)	(0.060)	(0.085)	(0.060)	(0.086)
East Mid	0.158 ***	-0.097	0.167 ***	-0.114	0.157 ***	-0.145 *
	(0.059)	(0.084)	(0.060)	(0.086)	(0.060)	(0.087)
West Mid	0.052	-0.006	0.063	-0.030	0.038	-0.048
	(0.056)	(0.085)	(0.057)	(0.088)	(0.058)	(0.090)
East	-0.032	-0.070	-0.022	-0.095	-0.031	-0.118
	(0.057)	(0.079)	(0.058)	(0.082)	(0.058)	(0.083)
London	0.275 ***	0.107	0.254 ***	0.044	0.251 ***	0.009
	(0.059)	(0.102)	(0.059)	(0.105)	(0.060)	(0.107)
S. East	0.041	-0.109	0.036	-0.151 **	0.029	-0.176 **
	(0.053)	(0.073)	(0.054)	(0.076)	(0.054)	(0.077)
Idaci	-0.318 ***	-0.840 ***	-0.065	-0.378 ***	-0.077	-0.379 ***
	(0.088)	(0.128)	(0.092)	(0.143)	(0.093)	(0.146)
Fsm eligibility	-0.131 ***	-0.323 ***	-0.062	-0.140 *	-0.051	-0.122 *
	(0.048)	(0.065)	(0.051)	(0.073)	(0.051)	(0.074)
Mum up to gcse			-0.009	0.348 ***	-0.027	0.323 ***
			(0.041)	(0.055)	(0.041)	(0.055)
Mum above gcse			0.178 ***	0.609 ***	0.161 ***	0.569 ***
			(0.047)	(0.063)	(0.047)	(0.064)
Mum degree			0.482 ***	1.049 ***	0.485 ***	1.028 ***
			(0.061)	(0.098)	(0.060)	(0.100)
Gran. uni			0.141 ***	0.317 ***	0.145 ***	0.339 ***
			(0.039)	(0.072)	(0.039)	(0.073)
Ambition					0.062 ***	0.010
					(0.020)	(0.028)
No plan					-0.290 ***	-0.326 ***
					(0.024)	(0.028)
Un rate		0.044 ***		0.039 ***		0.039 ***
		(0.014)		(0.015)		(0.015)
Observations	9030		9030		9030	

Standard errors in parentheses $^{*}p < 0.1$, $^{**}p < 0.05$,$^{***}p < 0.01$

Absence of planning is strongly and negatively correlated with both the perceived probability of being admitted and the likelihood of applying to

university. This suggests that the variable picks up both heterogeneity in student achievement and expectations formation, and differences in the discount rate, as students with a higher discount rate are willing to forgone the returns derived by an early entrance in the labor market to achieve higher returns in the log run. The coefficient for ambition is significant just in the main equation, suggesting that more ambitious young people are more motivated and positive in forming their educational expectations on the probability of being admitted for university. However, in spite of this, university does not seem the privileged channel for ambitious people to fulfil their aspirations.

Finally, the exclusion restriction, namely the local unemployment rate, has a positive and significant coefficient. This shows that, once the area specific probability of being a low-income family is controlled for via the idaci index (which has a negative sign), high unemployment rates reflect lower opportunity costs of being in education in the form of low expected returns of being in the labor market ($r_{\tau}^{\neg e}$). This is in line with the hypothesis put forward by Reynolds and Pemberton (2001) which links the expansion of university education to the deterioration of labor market conditions.

Table 7 shows the results of the models controlling for average key stage 2 grades, i.e. grades obtained at the end of year 6 when pupils are about 11 (about 3 years before the time when data are collected). This is meant to capture the information on their own ability respondents got early in their life. The coefficient for grades is positive and significant both in the main and in the selection equation, and it remains significant also when other controls are added. When grades are controlled for, the coefficient for age in the main equation becomes insignificant. This means that optimistic expectations over the probability of being admitted for university for relatively older people are at least partially due to their having higher grades. Moreover, the negative and significant coefficient for the variable 'age' suggests that older children think they are less likely to apply to university and that the positive coefficient shown in table 6 could be due to relatively older people having higher grades (see, for example, Crawford et al., 2007). On the contrary, the coefficient for female becomes positive and significant in both equations and the coefficients for ethnic minorities, already positive and significant, become even bigger in magnitude. Both results suggest that high expectations for girls and ethnic minorities are not necessarily due to high grades (see also Duckworth et al., 2009 and Strand, 2007).

Table 7: Likelihood of being accepted for university (opsel model, year 9, K2 grades, selected coefficients)

	1		2		3	
	accepted	selection	accepted	selection	accepted	selection
K2	0.115 ***	0.136 ***	0.106 ***	0.126 ***	0.103 ***	0.120 ***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)
Age	0.041	-0.099 **	0.044	-0.098 **	0.041	-0.097 **
	(0.028)	(0.043)	(0.029)	(0.044)	(0.029)	(0.044)
Female	0.055 **	0.194 ***	0.059 **	0.197 ***	0.065 **	0.198 ***
	(0.027)	(0.040)	(0.027)	(0.041)	(0.027)	(0.042)
Asian	0.730 ***	1.030 ***	0.667 ***	1.039 ***	0.654 ***	1.033 ***
	(0.045)	(0.083)	(0.050)	(0.091)	(0.050)	(0.092)
Black	0.711 ***	0.918 ***	0.657 ***	0.867 ***	0.585 ***	0.846 ***
	(0.070)	(0.150)	(0.071)	(0.153)	(0.071)	(0.156)
Other	0.258 ***	0.394 ***	0.213 ***	0.369 ***	0.185 ***	0.346 ***
	(0.057)	(0.098)	(0.057)	(0.101)	(0.058)	(0.101)
N. East	0.263 ***	-0.052	0.275 ***	-0.068	0.249 ***	-0.068
	(0.076)	(0.116)	(0.077)	(0.118)	(0.078)	(0.119)
N. West	0.152 ***	-0.251 ***	0.144 **	-0.278 ***	0.120 **	-0.286 ***
	(0.058)	(0.083)	(0.059)	(0.085)	(0.059)	(0.086)
York. & Hum.	0.211 ***	-0.126	0.213 ***	-0.143	0.204 ***	-0.142
	(0.061)	(0.090)	(0.062)	(0.092)	(0.062)	(0.092)
East Mid	0.148 **	-0.107	0.158 **	-0.124	0.143 **	-0.145
	(0.062)	(0.091)	(0.062)	(0.092)	(0.062)	(0.093)
West Mid	0.044	-0.041	0.053	-0.052	0.025	-0.058
	(0.059)	(0.093)	(0.059)	(0.095)	(0.060)	(0.096)
East	-0.040	-0.102	-0.027	-0.121	-0.042	-0.128
	(0.060)	(0.086)	(0.060)	(0.088)	(0.060)	(0.088)
London	0.192 ***	0.019	0.182 ***	-0.040	0.180 ***	-0.044
	(0.061)	(0.111)	(0.062)	(0.113)	(0.062)	(0.114)
S. East	0.057	-0.089	0.054	-0.120	0.040	-0.132
	(0.055)	(0.080)	(0.056)	(0.082)	(0.056)	(0.083)
Idaci	0.023	-0.078	0.102	0.053	0.068	0.041
	(0.089)	(0.144)	(0.096)	(0.156)	(0.096)	(0.158)
Fsm eligibility	-0.048	-0.128 *	-0.035	-0.069	-0.032	-0.065
	(0.049)	(0.071)	(0.053)	(0.078)	(0.053)	(0.078)
Mum up to gcse			-0.062	0.236 ***	-0.070	0.229 ***
			(0.043)	(0.058)	(0.043)	(0.059)
Mum above gcse			0.063	0.389 ***	0.068	0.376 ***
			(0.048)	(0.068)	(0.048)	(0.068)
Mum degree			0.313 ***	0.722 ***	0.344 ***	0.731 ***
			(0.059)	(0.108)	(0.059)	(0.109)
Gran. uni			0.104 ***	0.294 ***	0.114 ***	0.306 ***
			(0.040)	(0.080)	(0.040)	(0.080)
Ambition					0.113 ***	0.053 *
					(0.021)	(0.030)
No plan					-0.246 ***	-0.197 ***
					(0.022)	(0.030)
Un rate		0.018		0.015		0.013
		(0.016)		(0.016)		(0.016)
Observations	8753		8753		8753	

Standard errors in parentheses $^{*}p < 0.1$, $^{**}p < 0.05$,$^{***}p < 0.01$

The negative coefficient estimated for the variables proxying for income (idaci index and indicator for free school meal eligibility, **check cars, when models are corrected**) generally disappears. Similarly, also the estimated coefficients for maternal education are reduced if compared with the models in table 6 and they remain significant mainly in the selection equation. This is in line with those arguing that long run factors are incorporated in early attainment and, through that, they are able to influence subsequent outcomes. Moreover, this also suggests that familiar education affects the choice of going to university also through aspirations. Finally, the estimated coefficient for ambition becomes bigger and the one for absence of plan become smaller, especially in the selection equation. This suggests that ambition is a substitute for grades, while ability of planning a complement.

However, these results must be interpreted with caution, given that the coefficient for the exclusion restriction is not significant. We suspect that this is due to the fact that key stage 2 grades are strongly correlated with the variables proxying for household income, which makes it difficult to control for the factors that could be correlated with the impossibility for the families to afford university education. As a consequence, the coefficient for the local unemployment rate picks up an effect which is the combination of the negative effect of being a low-income family (otherwise picked up by the idaci index) and what we meant to capture, i.e. the positive effect of unemployment on opportunity costs.

Table 8 shows selected coefficients for the opsel models estimated at year 10 when the young people are aware of the results of their Key stage 3 exams. K3 scores are made public at the end of year 9 and therefore, unlike K2 score for the models in table (6), they are 'fresh' information on people's ability. For the models at year 10, we use as an additional exclusion restriction the variable indicating whether the respondent thinks he will be eligible for EMA. We expect that those who know they are eligible for a post-compulsory educational allowance are also more likely to expect to stay in education until they reach the age when they can apply to university.

Table 8: Likelihood of being accepted for university (opsel model, year 10, K3 grades, selected coefficients)

	1		2		3	
	accepted	selection	accepted	selection	accepted	selection
k3	0.089 ***	0.114 ***	0.085 ***	0.109 ***	0.083 ***	0.108 ***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Age	0.041	-0.064	0.044	-0.069	0.042	-0.070
	(0.029)	(0.042)	(0.029)	(0.043)	(0.029)	(0.043)
Female	0.043	0.206 ***	0.048 *	0.210 ***	0.054 **	0.213 ***
	(0.027)	(0.040)	(0.027)	(0.040)	(0.027)	(0.041)
Asian	0.743 ***	1.233 ***	0.696 ***	1.227 ***	0.688 ***	1.222 ***
	(0.045)	(0.085)	(0.049)	(0.092)	(0.049)	(0.092)
Black	0.719 ***	0.697 ***	0.703 ***	0.647 ***	0.641 ***	0.628 ***
	(0.069)	(0.120)	(0.069)	(0.122)	(0.070)	(0.123)
Other	0.191 ***	0.147 *	0.179 ***	0.131	0.157 ***	0.123
	(0.057)	(0.088)	(0.058)	(0.090)	(0.058)	(0.090)
N. East	0.231 ***	0.114	0.239 ***	0.108	0.229 ***	0.112
	(0.077)	(0.116)	(0.078)	(0.117)	(0.078)	(0.117)
N. West	0.108 *	-0.036	0.099 *	-0.052	0.087	-0.054
	(0.059)	(0.086)	(0.060)	(0.087)	(0.060)	(0.087)
York. & Hum.	0.174 ***	-0.109	0.168 ***	-0.118	0.164 **	-0.117
	(0.063)	(0.089)	(0.064)	(0.090)	(0.064)	(0.090)
East Mid	0.020	-0.110	0.030	-0.115	0.020	-0.114
	(0.063)	(0.090)	(0.063)	(0.091)	(0.063)	(0.091)
West Mid	0.099	-0.151 *	0.102 *	-0.166 *	0.078	-0.167 *
	(0.060)	(0.088)	(0.061)	(0.089)	(0.061)	(0.090)
East	-0.104 *	-0.101	-0.095	-0.102	-0.101 *	-0.099
	(0.061)	(0.085)	(0.061)	(0.086)	(0.061)	(0.086)
London	0.157 **	0.077	0.150 **	0.043	0.144 **	0.039
	(0.062)	(0.111)	(0.063)	(0.112)	(0.063)	(0.112)
S. East	-0.038	-0.081	-0.041	-0.100	-0.050	-0.103
	(0.057)	(0.079)	(0.058)	(0.080)	(0.058)	(0.080)
Idaci	0.099	-0.208	0.174 *	-0.087	0.150	-0.097
	(0.091)	(0.141)	(0.098)	(0.151)	(0.098)	(0.151)
Fsm eligibility	0.017	-0.042	0.045	0.012	0.047	0.013
	(0.049)	(0.070)	(0.054)	(0.077)	(0.054)	(0.077)
Mum up to gcse			-0.036	0.062	-0.049	0.058
			(0.043)	(0.058)	(0.043)	(0.058)
Mum above gcse			0.057	0.231 ***	0.052	0.227 ***
			(0.047)	(0.067)	(0.047)	(0.067)
Mum degree			0.254 ***	0.549 ***	0.265 ***	0.554 ***
			(0.058)	(0.111)	(0.058)	(0.111)
Gran. uni			0.032	0.142 *	0.038	0.144 *
			(0.040)	(0.074)	(0.041)	(0.074)
Ambition					0.081 ***	0.034
					(0.021)	(0.029)
No plan					-0.191 ***	-0.058 *
					(0.022)	(0.031)
Un rate		0.039 **		0.033 **		0.031 *
		(0.015)		(0.016)		(0.016)
EMA eligibility		0.125 ***		0.157 ***		0.153 ***
		(0.044)		(0.045)		(0.045)
Observations	8774		8774		8774	

Standard errors in parentheses $^{*}p < 0.1$, $^{**}p < 0.05$,$^{***}p < 0.01$

The model shows that, like in table 7, high grades are associated with a higher perceived likelihood of applying to university and with a higher perceived likelihood of being admitted. However, unlike in the models in table 7, now females seem to believe to be more likely to both applying to and be accepted for university. Moreover, people from more deprived areas tend to have higher aspirations on the likelihood of being admitted for university if compared to those from less deprived areas, but they do not seem to have higher expected probabilities of applying to university (if anything, their likelihood of applying to university seems to be lower). The first finding is in line with Chowdry et al., (2010) which suggests that, after controlling for grades, low-income young people are more likely to declare they are good at school than people coming from richer families. The second finding suggests that poor respondents face more binding budget constraints in their choice of applying to university.

People from high-educated families seem to be more likely to have higher expectations if compared to those from low-educated families. This is in line with Cameron and Heckman, (1999; 2001), Carneiro and Heckman (2002) and their theory that 'long run factors' are more important than income itself in determining education. These coefficients look attenuated if compared to those estimated at year 9 (table 7). This is compatible with a scenario in which familiar influence, although persistent, becomes less important with time (see also: Fowley, 2011; Attanasio and Kaufmann, 2009).[16]

The coefficients for both our exclusion restrictions are positive and significant, suggesting that opportunity and monetary costs are likely to affect people's choice to apply to university. Significant coefficients are also a point in favor of our identification strategy, thus reinforcing any conclusion driven from this model.

Table 9 shows the results where grades are interacted with the variable indicating free school meal eligibility. This is meant to study whether different people have a different degree of sensitivity to new information contained in grades (see, for example, Jacob and Wilder, 2010). Our model suggests that this might be the case. In fact, the coefficient for the interaction is negative and significant, suggesting that, compared to higher-income people, students who are eligible for free school meals are less likely to respond to high grades

[16]We found very similar results in a model where we use just the unemployment rate as the unique exclusion restriction (results not shown), suggesting that the results do not depend on the inclusion of the variable on EMA eligibility.

in forming their expectations.

Results in table 9 can shed some light on whether the difference in expectations between low and high income pupils is due to budget constraints only. We could be tempted to interpret the sign of the interaction in the selection equation claiming that low income people are not responsive to information on their ability because of the existence of credit constraints (see Attanasio and Kaufmann, 2009). However, was this the only explanation, the coefficient for the same variable should not be significant in the main equation, because expectations on the probability of being admitted given application do not depend on monetary constraints. This does not seem to be the case, given that the sign of the coefficient is negative and significant in both equations and in all specifications. This last finding is in line with what found by Jensen (2010) and it adds some evidence to the theory that there are psychological reasons on the top of economic reason which determine the gap in attainment between rich and poor people.

However, the sign and the significance of the coefficient on EMA eligibility suggests that the hypothesis of credit constraints can not be ruled out completely. In fact, being able to rely on a grant covering the future cost of education does seem to boost the expected probability of applying to university. Indeed, it adds an additional piece of information, i.e. that policies which make it easier for people to sustain the costs of post compulsory education can have an indirect effect on university attendance, or, at least, on expectations regarding it.

Table 9: Likelihood of being accepted for university (opsel model, year 10, K3 grades and interaction with fsm, selected coefficients)

	1		2		3	
	accepted	selection	accepted	selection	accepted	selection
K3	0.094 ***	0.119 ***	0.089 ***	0.113 ***	0.088 ***	0.112 ***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
K3av*fsm	-0.037 ***	-0.040 ***	-0.034 ***	-0.039 ***	-0.036 ***	-0.040 ***
	(0.008)	(0.011)	(0.008)	(0.011)	(0.008)	(0.011)
Age	0.039	-0.065	0.042	-0.069	0.039	-0.070 *
	(0.029)	(0.042)	(0.029)	(0.043)	(0.029)	(0.043)
Female	0.043	0.206 ***	0.048 *	0.209 ***	0.054 **	0.213 ***
	(0.027)	(0.040)	(0.027)	(0.040)	(0.027)	(0.041)
Asian	0.752 ***	1.224 ***	0.704 ***	1.221 ***	0.696 ***	1.215 ***
	(0.045)	(0.085)	(0.049)	(0.091)	(0.049)	(0.092)
Black	0.722 ***	0.700 ***	0.705 ***	0.649 ***	0.642 ***	0.629 ***
	(0.069)	(0.120)	(0.069)	(0.122)	(0.070)	(0.122)
Other	0.196 ***	0.161 *	0.182 ***	0.144	0.161 ***	0.136
	(0.058)	(0.088)	(0.058)	(0.090)	(0.058)	(0.090)
N. East	0.234 ***	0.124	0.242 ***	0.118	0.232 ***	0.122
	(0.077)	(0.116)	(0.078)	(0.117)	(0.078)	(0.117)
N. West	0.110 *	-0.032	0.101 *	-0.048	0.088	-0.050
	(0.059)	(0.086)	(0.060)	(0.087)	(0.060)	(0.088)
York. & Hum.	0.171 ***	-0.109	0.166 ***	-0.119	0.161 **	-0.118
	(0.063)	(0.089)	(0.064)	(0.090)	(0.064)	(0.090)
East Mid	0.016	-0.114	0.027	-0.118	0.016	-0.117
	(0.063)	(0.090)	(0.063)	(0.091)	(0.063)	(0.091)
West Mid	0.097	-0.153 *	0.100	-0.167 *	0.076	-0.169 *
	(0.061)	(0.089)	(0.061)	(0.090)	(0.061)	(0.090)
East	-0.108 *	-0.107	-0.099	-0.108	-0.106 *	-0.106
	(0.061)	(0.085)	(0.061)	(0.086)	(0.062)	(0.086)
London	0.155 **	0.071	0.148 **	0.037	0.142 **	0.034
	(0.062)	(0.111)	(0.063)	(0.112)	(0.063)	(0.112)
S. East	-0.041	-0.084	-0.043	-0.104	-0.053	-0.107
	(0.057)	(0.079)	(0.058)	(0.081)	(0.058)	(0.081)
Idaci	0.128	-0.191	0.192 **	-0.076	0.169 *	-0.086
	(0.091)	(0.141)	(0.098)	(0.151)	(0.098)	(0.151)
Fsm eligibility	1.198 ***	1.091 ***	1.109 ***	1.115 ***	1.182 ***	1.142 ***
	(0.249)	(0.323)	(0.250)	(0.326)	(0.251)	(0.326)
Mum up to gcse			-0.031	0.068	-0.044	0.064
			(0.043)	(0.058)	(0.043)	(0.058)
Mum above gcse			0.058	0.235 ***	0.053	0.231 ***
			(0.047)	(0.067)	(0.047)	(0.067)
Mum degree			0.246 ***	0.552 ***	0.257 ***	0.558 ***
			(0.058)	(0.111)	(0.058)	(0.111)
Gran. uni			0.032	0.139 *	0.038	0.142 *
			(0.040)	(0.074)	(0.041)	(0.075)
Ambition					0.082 ***	0.035
					(0.021)	(0.029)
No plan					-0.193 ***	-0.060 **
					(0.022)	(0.031)
Un rate		0.038 **		0.033 **		0.030 *
		(0.015)		(0.016)		(0.016)
EMA eligibility		0.136 ***		0.166 ***		0.162 ***
		(0.044)		(0.045)		(0.045)
Observations	8774		8774		8774	

Standard errors in parentheses $^{*}p < 0.1$, $^{**}p < 0.05$,$^{***}p < 0.01$

6.1 Magnitude of the effects

The coefficients estimated in the previous section do not permit to assess the magnitude of the effects nor to fully compare the different role factors play in the selection and in the main equation. To provide some evidence of this, we use specification 3 in table 9 to estimate predicted probabilities for the selection equation and for each outcome in a_{it} (see equations 3 and 2). This strategy permits also to study the differential effect of each factor on each of the levels of a_{it}, which gives some additional information over and above the results estimated for the main equation in the models in table (9).

We focus on male, white British born in the UK, because our analysis so far has shown that this group has significantly lower expectations if compared to females, ethnic minorities and people not born in the UK. Predicted probabilities have been computed separately for people differing in their familiar income, wealth and education. In particular, we simulated low income/wealth families by setting the dummy variables indicating free school meal eligibility to 1, and the dummy variables indicating house ownership and cars owned by the family to 0. Similarly, high income/wealth families are simulated by setting the number of cars to 3 and the dummies indicating free school meal eligibility and house ownership to 0. Finally, high familiar education is simulated by assigning university education to both the mother and at least one of the grandparents, while low-education families are defined as those having mother with no education at all and none of the grandparents with an university degree. Due to the impossibility of presenting all the combinations of familiar income/wealth and education, we decided to focus on 4 profiles, which are summarized in table 10.

Table 10: Characteristics of the profiles created

profile	fsm eligibility	cars	house ownership	familiar education	others
hwe	no	3	1	high	mean
hw	no	3	1	mean	mean
lw	yes	0	0	mean	mean
lwe	yes	0	0	low	mean

We estimated a different probability for each of the three relevant values of the average key stage three score (27, 33 and 39) which are the minimum

required scores to be awarded, respectively, level 4, 5 and 6 (see table 2). Given that level 5 is considered the minimum satisfactory level for key stage 3, being awarded 27 must be considered as a negative signal, being awarded 33 as a neutral signal and being awarded 39 as a positive signal. All the other variables are set at their sample means. Standard errors are computed via delta method.

The estimated probabilities for males are reported in table 24 (the probabilities for females are in the appendix).

Table 11: Predicted probability on the expected likelihood of being accepted for university if apply (males)

	27		33		39		27		33		39	
	not at all likely						not very likely					
hwe	0.040	***	0.012	***	0.003	***	0.304	***	0.170	***	0.076	***
	(0.007)		(0.002)		(0.001)		(0.021)		(0.014)		(0.008)	
hw	0.071	***	0.024	***	0.007	***	0.379	***	0.240	***	0.122	***
	(0.008)		(0.003)		(0.001)		(0.016)		(0.012)		(0.007)	
lw	0.043	***	0.022	***	0.011	***	0.312	***-	0.231	***	0.159	***
	(0.006)		(0.003)		(0.002)		(0.018)		(0.015)		(0.015)	
lwe	0.046	***	0.024	***	0.012	***	0.322	***	0.241	***	0.167	***
	(0.007)		(0.004)		(0.002)		(0.020)		(0.017)		(0.017)	
	fairly likely						very likely					
hwe	0.597	***	0.672	***	0.631	***	0.059	***	0.145	***	0.290	***
	(0.0181)		(0.006)		(0.012)		(0.009)		(0.014)		(0.019)	
hw	0.517	***	0.643	***	0.666	***	0.034	***	0.092	***	0.205	***
	(0.0173)		(0.008)		(0.006)		(0.005)		(0.008)		(0.011)	
lw	0.589	***	0.649	***	0.672	***	0.057	***	0.098	***	0.158	***
	(0.0163)		(0.009)		(0.006)		(0.008)		(0.010)		(0.016)	
lwe	0.580	***	0.644	***	0.673	***	0.052	***	0.091	***	0.148	***
	(0.0185)		(0.011)		(0.006)		(0.008)		(0.011)		(0.017)	

Standard errors (in parentheses) have been computed via the delta method

A comparison between the predicted probabilities estimated for the profiles hw and lw shows that relatively poor people (lw) who obtain 27 are more optimistic than relatively richer people who obtain the same score. In fact, the probabilities estimated for the former are smaller than those estimated for the latter for low levels of likelihood ('not at all likely' and 'not very likely') and bigger for higher levels ('fairly likely' and 'very likely'). Although reduced in magnitude, this pattern can still be found for people getting an average key stage score of 33 and it gets reverted for those getting 39. (perhaps provide significance levels when the definite specification is agreed) Taken together, our results do not suggest that people from low-income families are less likely to apply to university because they have a lower perceived probability of being admitted. In contrast, our estimates suggest that low achievers from relatively poor families tend to have more optimistic expectations on their chances of being admitted given application if compared to their less disadvantaged peers. A comparison between profiles hwe and hw

and between profiles lwe and lw (namely, profiles differing for their level of education only) shows that, for a given obtained mark, education leads to more optimistic predictions and this happens for each of the three key stage 3 levels considered.

The difference between the estimated probabilities computed for different average key stage 3 scores helps quantifying the marginal effects of grades on expectations. Tables 12 report estimates of these effects, together with a measure of their significance. Standard errors have been computed by bootstrap with 50 draws.

Table 12: Marginal effects of grades: males

profile	not at all likely		not very likely		fairly likely		very likely	
Effect of getting an average key stage score of 27								
hwe	0.0279	***	0.1333	***	-0.0753	***	-0.0859	it ***
	(0.0053)		(0.0079)		(0.0184)		(0.0079)	
hw	0.0464	***	0.1383	***	-0.1262	***	-0.0585	***
	(0.0059)		(0.0062)		(0.0122)		(0.0055)	
lw	0.0207	***	0.0808	***	-0.0599	***	-0.0402	***
	(0.0037)		(0.0092)		(0.0107)		(0.0058)	
lwe	0.0220	***	0.0806	***	-0.0645	***	-0.0391	***
	(0.0040)		(0.0091)		(0.0115)		(0.0061)	
Effect of getting an average key stage score of 39								
hwe	-0.0092	***	-0.0939	***	-0.0416	**	0.1447	***
	(0.0019)		(0.0085)		(0.0166)		(0.0088)	
hw	-0.0176	***	-0.1180	***	0.0233	*	0.1123	***
	(0.0025)		(0.0065)		(0.0135)		(0.0082)	
lw	-0.0116	***	-0.0720	***	0.0234	**	0.0602	***
	(0.0019)		(0 .0076)		(0.0097)		(0.0098)	
lwe	-0.0124	***	-0.0739	***	0.0287	***	0.0576	***
	(0.0021)		(0.0080)		(0.0102)		(0.0095)	

Standard errors (in parentheses) have been computed via bootstrap with 50 drawn

Table 12 shows that, for each profile, getting 27 as opposed to getting 33 leads to an increase in the probability associated with the most negative outcomes ('not at all likely' and 'not very likely' of being accepted') and a drop in the probabilities associated with the most positive ones ('fairly likely' and 'very likely' to be accepted). Moreover, getting 39 as opposed to 33 leads to the opposite effect, namely an increase in the probability associated with the positive outcomes and a decrease in the probability associated with the negative ones.

A comparison between profiles hw and lw, namely between profiles which differ in their level of income and wealth only, shows that grades have a bigger effect (in absolute terms) on expectations hold by wealthy people if compared to expectations hold by people from poorer households. For example, getting 27 as opposed to 33 increases the probability that the respondent answer it is 'not at all likely' he will be admitted for university by 4.64% in the case of the profile hw, and by 2.07% in the case of profile lw. The opposite is true when we look at the other tail of the distribution of $E_t(p^a)$. In fact, getting 27 is associated with a 5.85% decrease in the probability of being 'very likely' to be accepted for university for profile hw, and a 4.02% decreases for profile lw. This means that claiming that relative poor people have lower expectations on the probability of being admitted also because of their lower responsiveness if not completely true. In fact, poorer people are less responsive than richer people not only to high grades, but also to low grades.

A comparison between the estimated marginal effects for profiles hw and hwe shows whether the effect of grades changes in the case of pupils differing in their level of familiar education only. The effect of grades on expectations is smaller in the case of people from high-educated families when low levels of expectations are considered (outcomes 'not at all likely' and 'not very likely'), while is higher in case of higher levels of expectations (outcomes 'fairly likely' and 'very likely'). The case of low education is the mirror image: if compared to profile lw, profile lwe shows a smaller effect of grades in the case of low levels and a bigger effect in the case of high levels of expectations. Taken together, the comparisons show that familiar education reduces the effect of grades at the left of the distribution of $E_{it}(p_i^a)$ and increases it at the right.

For example, for profile hwe, getting a negative signal (27 points) increases the probability of answering 'not at all likely' by 2.75%, almost 2 percentage points less than what observed for profile hw, which has everything in common with hwe, but a lower level of familiar education. For profile hwe, the same signals is associated to a 8.59% drop in the probability of answering 'very likely' which is more than 3 points percent bigger than what estimated for profile hw.

If compared to table **??** (in the appendix), table 12 shows that the marginal effects estimated for males for low values of a_{it} are bigger (in absolute value) than those estimated for females, while they are smaller for high values of a_{it}. This means that the more pessimistic educational expectations of males (see table 24) are due to the fact that males are both more prone

Table 13: Estimated selection probabilities (males)

avk3	unrate	EMA	hwe	hw	lw	lwe
27	p50	0	0.7346	0.5432	0.5648	0.5060
			(0.0385)	(0.0196)	(0.0257)	(0.0288)
33	p25	0	0.8901	0.7580	0.7207	0.6711
			(0.0215)	(0.0155)	(0.0258)	(0.0306)
33	p50	0	0.8961	0.7680	0.7314	0.6828
			(0.0205)	(0.0140)	(0.0243)	(0.0291)
33	p50	1	0.9261	0.8206	0.7889	0.7466
			(0.0165)	(0.0141)	(0.0215)	(0.0263)
33	p75	0	0.9018	0.7777	0.7419	0.6943
			(0.0200)	(0.0149)	(0.0243)	(0.0292)
39	p50	0	0.9708	0.9126	0.8581	0.8254
			(0.0076)	(0.0084)	(0.0229)	(0.0283)

Standard errors (in parentheses) been computed via delta methods

than females to incorporate negative information in their expectations and less prone to incorporate positive information.

Table 13 shows the probabilities estimated for males for the selection equation under different scenarios. Results show that for each combination of key stage score, local unemployment rate and EMA eligibility (each row of the table), the probability of expecting of being at least 'not very likely' to apply to university increases together with socio-economic status. A comparison between profiles hwe and hw and lwe and lw shows that such pattern is indeed driven by differences in family education, although a comparison between profiles hw and lw for both males and females shows that wealth also plays a role in shaping people's expectations on applying to university. The latter is an important result given that our previous analysis shows that relatively poor people do not necessarily have lower expectations of being accepted for university if compared to their wealthier peers, this means that there must be other factors preventing poorer people from applying to university, including credit constraints, aspirations, peer effects and role models. Unfortunately, we are aware that our model does not permit to identify them separately.

Table 14 shows selected marginal effects estimates for the selection equation. As a baseline, we chose the probability estimated in the case of an individual with an average key stage 3 score of 33, no EMA eligibility, and a median local unemployment rate. The variables describing socio-economic status are set at the level defined by each profile. Marginal effects are com-

puted by deriving predicted probabilities when we depart from that scenario and then by estimating the difference with the baseline. The estimated effects of getting 27 and 39 as opposed to 33 are smaller (in absolute terms) for relatively poor people if compared to relatively rich ones. In fact, being awarded 27 decreases the probability of attaching a positive probability to the choice of applying to university in the future by 22.47% for boys in high-income-high-wealth families (profile hw), and only by 16.66% for boys from low-income-low-wealth families. Being awarded a 39 still has a smaller effect in the case of low income families, but the difference here is less than 2 percentage points for (14.46% vs 12.66%).

This suggests that the lower responsiveness of poorer people comes mainly from their receiving limited disincentives from low grades. The same pattern has been already observed for the case of the expectations on the probability of being accepted for university, which provides some evidence on the link between the two measures of expectations. Comparisons between profiles hw and hwe, and lw and lwe suggest that a relatively higher familiar education is associated with lower marginal effects and therefore with a lower responsiveness to grades. This has a parallel in the results shown for the main equation, where education seems to induce a lower responsiveness for low levels of the perceived likelihood of being accepted for university given application. The marginal effects estimated for 25% changes in the local unemployment rate are not significant, suggesting that movements in the unemployment rate must be consistently big to be incorporated in expectations. Finally, EMA eligibility does seem to affect the expected probability of applying to university with a slightly bigger effect estimated for relatively poorer people and males.

EMA eligibility increases the probability that $S_{it} > 0$ by 5.75% for boys for low-income families and 5.27% for boys from high-income families. EMA seems to be specifically relevant for males with a low level of familiar education, who are those showing the lowest expected likelihood of applying to university. This suggests that decreasing monetary costs of education can open the doors of university for people who face higher psychological costs of being in education or (and) who have lower expected returns of education due a scarcity, within the family of the group of peers, of examples of success in the labor market.

Table 14: Marginal effects, selection equation

hwe: Non free school meal eligible, above average wealth and family education									
-0.1615	***	0.0747	***	-0.0060	ns	0.0058	ns	0.0301	***
(0.0203)		(0.0157)		(0.0038)		(0.0036)		(0.0069)	
hw: Non free school meal eligible, above average wealth									
-0.2247	***	0.1446	***	-0.0099	ns	0.0097	ns	0.0527	***
(0.0103)		(0.0137)		(0.0063)		(0.0060)		(0.0096)	
lw: Free school meal eligible, below average wealth									
-0.1666	***	0.1266	***	-0.0107	ns	0.0105	ns	0.0575	***
(0.0185)		(0.0133)		(0.0068)		(0.0066)		(0.0105)	
lwe: Free school meal eligible, below average wealth									
-0.1769	***	0.1426	***	-0.0117	ns	0.0115	ns	0.0638	***
(0.0183)		(0.0134)		(0.0075)		(0.0073)		(0.0118)	

Standard errors (in parentheses) have been computed via bootstrap with 50 drawns

6.2 Robustness checks

One might argue that the probability of being accepted for university depends strongly on the degree the respondent plans to apply to, which a piece of information that we do not observe. Under the assumptions that ii) students know already with certainty the degree they want to apply to ii) acceptance rate differ significantly across subjects and iii) students are aware of these differences in acceptance rates, there will be an element of heterogeneity we are not fully able to control for.

To study this, we add to the models in table 15 a set of dummy variables indicating the subject in which the respondent got his or her best K3 score. This is meant to capture a difference in attitudes which could lead to the choice of different degrees.

Results show that getting the best score in English, as opposed to getting the best score in science or in math (reference category) is associated with a higher perceived likelihood of both applying to and being accepted for university. However, in both equations the coefficients indicating the best subject become insignificant with the introduction of behavioral variables, suggesting that heterogeity in the distribution of best key stage 3 result reflects heterogeneity in aspirations and in formation of expectations (B in our framework), rather than sharp subject-specific differences in the probability of being accepted for university given application.

Table 15: Likelihood of being accepted for university (opsel model, year 10, K3 grades and subjects, selected coefficients)

	1		2		3	
	accepted	selection	accepted	selection	accepted	selection
k3	0.091 ***	0.116 ***	0.087 ***	0.111 ***	0.085 ***	0.109 ***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Best english	0.066 *	0.089 *	0.061 *	0.079	0.049	0.077
	(0.035)	(0.050)	(0.035)	(0.050)	(0.035)	(0.050)
Best science	0.067	0.025	0.060	0.011	0.053	0.010
	(0.044)	(0.057)	(0.044)	(0.058)	(0.044)	(0.058)
Age	0.043	-0.063	0.045	-0.067	0.042	-0.069
	(0.029)	(0.042)	(0.029)	(0.043)	(0.029)	(0.043)
Female	0.034	0.189 ***	0.040	0.194 ***	0.048 *	0.198 ***
	(0.027)	(0.041)	(0.028)	(0.041)	(0.028)	(0.042)
Asian	0.745 ***	1.227 ***	0.697 ***	1.220 ***	0.689 ***	1.215 ***
	(0.045)	(0.085)	(0.049)	(0.092)	(0.049)	(0.093)
Black	0.716 ***	0.696 ***	0.701 ***	0.645 ***	0.640 ***	0.627 ***
	(0.069)	(0.121)	(0.070)	(0.123)	(0.070)	(0.123)
Other	0.192 ***	0.145	0.180 ***	0.129	0.159 ***	0.121
	(0.057)	(0.088)	(0.058)	(0.090)	(0.058)	(0.090)
N. East	0.234 ***	0.118	0.242 ***	0.112	0.231 ***	0.116
	(0.077)	(0.116)	(0.078)	(0.117)	(0.078)	(0.117)
N. West	0.111 *	-0.033	0.102 *	-0.049	0.089	-0.050
	(0.059)	(0.086)	(0.060)	(0.087)	(0.060)	(0.087)
York. & Hum.	0.176 ***	-0.106	0.170 ***	-0.114	0.165 ***	-0.114
	(0.063)	(0.089)	(0.064)	(0.090)	(0.064)	(0.090)
East Mid	0.024	-0.105	0.034	-0.109	0.023	-0.108
	(0.063)	(0.090)	(0.063)	(0.091)	(0.063)	(0.091)
West Mid	0.101 *	-0.146 *	0.104 *	-0.161 *	0.080	-0.162 *
	(0.060)	(0.088)	(0.061)	(0.090)	(0.061)	(0.090)
East	-0.101 *	-0.100	-0.093	-0.101	-0.100	-0.099
	(0.061)	(0.085)	(0.061)	(0.086)	(0.061)	(0.086)
London	0.159 **	0.075	0.151 **	0.042	0.145 **	0.038
	(0.062)	(0.111)	(0.063)	(0.112)	(0.063)	(0.112)
S. East	-0.039	-0.083	-0.042	-0.102	-0.051	-0.105
	(0.057)	(0.079)	(0.058)	(0.080)	(0.058)	(0.080)
Idaci	0.100	-0.199	0.175 *	-0.080	0.151	-0.090
	(0.091)	(0.142)	(0.098)	(0.151)	(0.098)	(0.151)
Fsm eligibility	0.017	-0.042	0.046	0.011	0.047	0.012
	(0.049)	(0.070)	(0.054)	(0.077)	(0.054)	(0.077)
Mum up to gcse			-0.038	0.059	-0.050	0.055
			(0.043)	(0.058)	(0.043)	(0.058)
Mum above gcse			0.054	0.228 ***	0.050	0.225 ***
			(0.047)	(0.067)	(0.047)	(0.067)
Mum degree			0.249 ***	0.545 ***	0.260 ***	0.550 ***
			(0.058)	(0.111)	(0.058)	(0.111)
Gran. uni			0.031	0.140 *	0.037	0.142 *
			(0.040)	(0.074)	(0.041)	(0.074)
Ambition					0.081 ***	0.034
					(0.021)	(0.029)
No plan					-0.189 ***	-0.056 *
					(0.022)	(0.031)
Un rate		0.038 **		0.033 **		0.031 *
		(0.015)		(0.016)		(0.016)
EMA eligibility		0.124 ***		0.156 ***		0.152 ***
		(0.044)		(0.045)		(0.045)
Observations	8770		8770		8770	

Standard errors in parentheses $^{*}p < 0.1$, $^{**}p < 0.05$,$^{***}p < 0.01$

7 Conclusions

The paper analyzes the formation of expectations of young people in year 9 and 10 at school over their perceived likelihood of applying to university in the future. By doing this, we aim at providing some evidence to guide the implementation of those policies aimed at widening the access to university education which, in England, have early teen-agers as their main targets and beneficiaries. A descriptive analysis of data from the Longitudinal Study of Young People in England shows that the transition between year 9 and 10 is indeed when educational expectations evolve the most, perhaps as a consequence of the flow of new information contained in key stage 3 test scores.

To explain why this information should play a role in the formation of educational expectations, while exploiting the information provided by the LSYPE data, we study the choice of apply to university as depending on the expected probability of being admitted if apply. The latter is clearly influenced by grades and provides a measure of expectations which is polished from the effect of aspirations.

Our results show that White-British-UK born males from low socio-economic background are those having the lowest educational expectations. As a consequence, this should be the main target sub-group for future policies. In particular, a specific attention should go to young from low-income-low education families who are heavily under-represented in the university population.

Our study helps show the mechanism leading young from low socio-economic background to be stuck in a low expectations-low achievement trap. Contrary to what is commonly believed, our analysis does not give evidence to conclude that relatively poor people hold lower expectations on the probability of being admitted to university given application. If anything, our estimates suggest that young people from free school meal eligible families have more positive expectations, even when grades are controlled for. However, expectations over the probability of being admitted for university do seem to be influenced by familiar education, although such an effect seems to be more prominent in the early ages and then decreases with time. In contrast, both social and economic background seem to affect the likelihood of applying to university with people from both low-income and low-education families declaring low likelihood of trying and access to university.

Crucially, young people from economically disadvantaged families display

a smaller responsiveness to grades if compared to wealthier peers. This happens for both the likelihood of being admitted for university (regardless of the level of probability they attach to the event) and the likelihood of applying. Therefore, the results show that there are psychological reasons on the top of economic reasons which lead to such a low responsiveness. The main consequence of this is that providing low-income people with early interventions aimed at filling the gap in attainment with high-income people could not be enough to persuade them that it is worth trying to apply for university. Moreover, our result seem to suggest that policies like the EMA, as well as anything affecting the direct and indirect costs of education, could be successful in opening the doors of university, especially for males from low-income-low-education families.

Finally, we find that students from high education families display high responsiveness to grades (both upwards and downwards) in the case of people having high expectations on the probability of being admitted to university given application. The opposite is true for people from high-educated families having a low expected likelihood of being admitted. A reason for this can be that students whose parents and grandparents went to university themselves find it easier to understand how the selection takes place. Therefore, by the age of 15, those who are really interested in higher education and who perceive they have positive chances to get in pay more attention to the signals contained in grades, because they know that grades have a great impact on the probability of admission. This implies that a policy who wants to increase the responsiveness to grades for talented people from low-educated families has to provide them with the tools to be able to interpret the signals contained in high test scores.

8 Appendix

8.1 Descriptive statistics

Table 16: Descriptive Statistics: individual variables

	pooled		wave 1		wave 2		wave 3	
Variable	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean
general								
age	41744	15.245	15766	14.326	13539	15.325	12439	16.323
female	39292	0.492	13441	0.493	13437	0.492	12414	0.492
not born Uk	41354	0.049	15538	0.051	13451	0.049	12365	0.048
ethnicity								
white brit	41016	0.844	15412	0.843	13331	0.844	12273	0.844
asian	41016	0.068	15412	0.069	13331	0.068	12273	0.068
black	41016	0.033	15412	0.033	13331	0.033	12273	0.033
other	41016	0.055	15412	0.055	13331	0.055	12273	0.055

Weighted data.

Table 17: Descriptive Statistics: familiar background

	pooled		wave 1		wave 2		wave 3	
Variable	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean
				general				
house owned	41419	0.719	15582	0.713	13469	0.722	12368	0.724
house rent	41419	0.281	15582	0.287	13469	0.278	12368	0.276
number cars	34632	1.728	12851	1.679	11288	1.738	10493	1.780
both parents	41122	0.750	15638	0.748	13281	0.760	12203	0.743
urban	39495	0.803	13532	0.799	13531	0.805	12432	0.804
hh income	29594	27714.4	11734	23857	7661	30699.5	10199	29865.6
gran uni	41476	0.129	15628	0.127	13467	0.129	12381	0.130
				gor				
north east	39494	0.051	13532	0.053	13531	0.051	12431	0.050
north west	39494	0.155	13532	0.153	13531	0.156	12431	0.156
york-humber	39494	0.104	13532	0.106	13531	0.103	12431	0.104
e. midlands	39494	0.084	13532	0.084	13531	0.083	12431	0.083
w. midlands	39494	0.115	13532	0.116	13531	0.115	12431	0.114
east	39494	0.106	13532	0.106	13531	0.106	12431	0.106
london	39494	0.129	13532	0.124	13531	0.132	12431	0.132
south east	39494	0.162	13532	0.164	13531	0.161	12431	0.161
south west	39494	0.094	13532	0.095	13531	0.093	12431	0.093
				mother				
full time	41267	0.348	15588	0.328	13354	0.345	12325	0.376
part time	41267	0.341	15588	0.353	13354	0.343	12325	0.324
not working	41267	0.278	15588	0.284	13354	0.282	12325	0.266
no education	41221	0.195	15468	0.195	13419	0.195	12334	0.193
gcse or below	41221	0.402	15468	0.403	13419	0.402	12334	0.401
above gcse	41221	0.257	15468	0.255	13419	0.258	12334	0.260
degree	41221	0.113	15468	0.112	13419	0.113	12334	0.115
age 0-35	41578	0.071	15656	0.100	13504	0.067	12418	0.040
age 35-50	41578	0.825	15656	0.820	13504	0.832	12418	0.826
age 50 plus	41578	0.083	15656	0.060	13504	0.083	12418	0.113
main parent	41534	0.856	15632	0.854	13510	0.859	12392	0.854
no mother	41122	0.033	15638	0.034	13281	0.031	12203	0.034

Weighted data.

8.2 Behavioral variables

Behavioral variables have been created based on answers to three batteries of questions asking the respondents to state how much they agree with a few statements and how important they consider a few life goals (see table 18). Between 2 and 3 percent of the cases go lost due to unit non response, in our definition, when the respondents do not get to answer the relevant section of the interview. Moreover, between 0.33 and 13.18 percent of the cases go lost due to item non response, when the respondent fails to answer to a specific question of the battery. Given that item non response varies across questions in the same battery, an available cases analysis is likely to be based on a small and possibly selected sample. For this reason, we decided to use chained equation imputation (Raghunathan et al. 2001) to impute the missing subitems for the item-respondents.

The chained-equations method was chosen because it permits to take into account the correlation among different variables, while allowing for missingness in the covariates. The missing items have been imputed just for the unit-respondents. Since answers of the questions contained in each battery are likely to be strongly correlated, this insures that no imputations have been done based primarily on information taken from other respondents. A first imputation was carried out for the question asked at wave 1, i.e. batteries A and B in table 18. Each item has been imputed through ordered logit using as covariates gender, year of birth and the other items of batteries A and B. After being imputed, these questions have been used, together with gender and year of birth, to impute answers to question C. Each estimation was carried out using 10 cycles.

After imputing the missing items, we used factor analysis to aggregate the items into a few factors (see table 18). We retained two factors for each question. We named 'locus of control' the first factor derived from battery A , 'absence of planning' the first factor derived from battery B and 'ambition' the first factor derived from battery C. The literature on survey methods has claimed that questions about beliefs and values can be subject to both 'acquiescence bias' and 'social desirability bias', i.e. the tendency of agreeing with the interviewer (see Paulhus, 1991; Saris, Revilla, Krosnick, and Shaeffer, 2010; Bentler and Jackson, 1971) and the tendency of giving answers that are socially acceptable (see Tourangeau, Rips, and Rasinski, 2000; Kreuter, Presser, and Tourangeau, 2008). Given the wording of the questions used and the results of the factor loading, we think that these bias

can be captured by factors 2. In particular, we think that factor 2 in batteries A and C are more likely to capture acquiescence, while factor 2 in battery B is more likely to be correlated with social desirability bias.

Table 18: Descriptive Statistics: behavioral variables

	non response			factor 1			factor 2	
	unit	item	name	eigen-value	factor loading	name	eigen-value	factor loading
A) And thinking more generally about the future, how much do you agree or disagree that:								
Having a job or career in the future is important to me	2.15	0.51		0.742	-0.304		0.123	0.192
raising a family in the future is important to me.	2.15	4			-0.150			0.238
I don't really think much about what I might be doing in a few years time	2.15	2.85	no planning		0.53	acquiescence		0.035
When I think about what I might do after Year 11, if I had to choose, it's more important to me that I'm doing something I enjoy rather than how it might help me me get a job later on	2.15	1.37			0.237			0.151
ll just wait and see where I end up.	2.15	1.41			0.540			0.073
B) I'm going to ask about things you might want to do in the future. Please tell me for each whether it matters a lot to you, a little, or not at all?								
To have a job where I help other people	2.15	0.93		0.841	0.119		0.062	0.196
To have a job which pays well	2.15	0.33			0.423			-0.096
To be my own boss/have my own business	2.15	1.43	ambition		0.453	social desiderability		-0.047
To have a job that's interesting and not routine	2.15	0.92			0.301			0.023
To have a job where I can get promoted and get ahead	2.15	1.07			0.509			-0.001
To have a job with regular hours	2.15	1.35			0.305			0.106
	N=15770				n=15431			

8.3 Results

Table 19: Likelihood of being accepted at university (opsel model, year 9, selected coefficients)

	1		2		3	
	accepted	selection	accepted	selection	accepted	selection
Month int	0.021 **	-0.011	0.019 *	-0.013	0.018 *	-0.014
	(0.010)	(0.014)	(0.010)	(0.015)	(0.010)	(0.015)
Not born uk	0.140 **	0.270 **	0.149 **	0.327 ***	0.158 ***	0.300 **
	(0.058)	(0.120)	(0.059)	(0.124)	(0.059)	(0.126)
House owned			0.078 **	0.107 **	0.075 **	0.090 *
			(0.038)	(0.052)	(0.038)	(0.053)
Number of cars			0.005	-0.030	0.003	-0.033
			(0.019)	(0.028)	(0.019)	(0.028)
Both parents			0.087 **	0.086	0.086 **	0.085
			(0.040)	(0.056)	(0.040)	(0.057)
Urban			0.001	0.081	-0.001	0.087 *
			(0.036)	(0.050)	(0.036)	(0.051)
M. part time			-0.009	0.064	0.001	0.069
			(0.031)	(0.045)	(0.031)	(0.046)
M. not working			0.018	0.018	0.031	0.034
			(0.038)	(0.055)	(0.038)	(0.056)
No mother			0.079	-0.007	0.090	-0.002
			(0.103)	(0.131)	(0.103)	(0.133)
Age m. 35 50			0.022	0.105 *	0.013	0.087
			(0.046)	(0.062)	(0.047)	(0.063)
Age m 50 plus			0.143 **	0.181 *	0.140 **	0.144
			(0.070)	(0.101)	(0.070)	(0.102)
Mum main parent			-0.064 *	-0.198 ***	-0.069 *	-0.187 ***
			(0.037)	(0.061)	(0.038)	(0.062)
Constant		1.221 **		0.661		0.809
		(0.571)		(0.597)		(0.605)
Observations	9030		9030		9030	

Standard errors in parentheses *$p < 0.1$, **$p < 0.05$,***$p < 0.01$

Table 20: Likelihood of being accepted at university (opsel model, year 9, K2 grades, selected coefficients)

	1		2		3	
	accepted	selection	accepted	selection	accepted	selection
Month int	0.015	-0.015	0.014	-0.016	0.013	-0.017
	(0.010)	(0.015)	(0.010)	(0.016)	(0.010)	(0.016)
Not born uk	0.187 ***	0.326 **	0.169 **	0.347 **	0.178 ***	0.319 **
	(0.065)	(0.143)	(0.066)	(0.145)	(0.067)	(0.145)
House owned			0.029	-0.009	0.029	-0.013
			(0.039)	(0.056)	(0.039)	(0.056)
Number of cars			-0.012	-0.044	-0.014	-0.045
			(0.020)	(0.029)	(0.020)	(0.029)
Both parents			0.098 **	0.109 *	0.097 **	0.107 *
			(0.041)	(0.060)	(0.041)	(0.060)
Urban			0.018	0.097 *	0.014	0.099 *
			(0.037)	(0.054)	(0.037)	(0.054)
M. part time			-0.044	0.033	-0.033	0.035
			(0.032)	(0.049)	(0.032)	(0.049)
M. not working			0.028	0.059	0.040	0.068
			(0.040)	(0.060)	(0.040)	(0.060)
No mother			0.037	-0.049	0.046	-0.048
			(0.110)	(0.141)	(0.110)	(0.142)
Age m. 35 50			-0.037	0.059	-0.035	0.053
			(0.049)	(0.066)	(0.049)	(0.067)
Age m 50 plus			0.105	0.130	0.116	0.111
			(0.073)	(0.110)	(0.073)	(0.110)
Mum main parent			-0.082 **	-0.200 ***	-0.089 **	-0.200 ***
			(0.039)	(0.067)	(0.039)	(0.068)
Constant		-1.064 *		-1.090 *		-0.877
		(0.630)		(0.647)		(0.652)
Observations	8753		8753		8753	

Standard errors in parentheses *$p < 0.1$, **$p < 0.05$,***$p < 0.01$

Table 21: Likelihood of being accepted at university (opsel model, year 10, K3 grades, selected coefficients)

	1		2		3	
	accepted	selection	accepted	selection	accepted	selection
Month int	0.011	-0.026	0.009	-0.026	0.010	-0.026
	(0.012)	(0.017)	(0.012)	(0.017)	(0.012)	(0.017)
Not born uk	0.262 ***	0.657 ***	0.251 ***	0.673 ***	0.265 ***	0.661 ***
	(0.061)	(0.136)	(0.062)	(0.138)	(0.062)	(0.138)
House owned			0.039	0.086	0.037	0.084
			(0.041)	(0.055)	(0.041)	(0.055)
Number of cars			-0.024	-0.020	-0.025	-0.021
			(0.018)	(0.026)	(0.018)	(0.026)
Both parents			0.128 ***	0.043	0.123 ***	0.042
			(0.042)	(0.059)	(0.042)	(0.059)
Urban			0.008	0.061	0.004	0.063
			(0.038)	(0.054)	(0.038)	(0.054)
M. part time			0.010	-0.028	0.014	-0.025
			(0.032)	(0.048)	(0.032)	(0.048)
M. not working			0.019	-0.003	0.021	0.001
			(0.040)	(0.059)	(0.040)	(0.059)
No mother			0.068	0.108	0.063	0.116
			(0.116)	(0.153)	(0.117)	(0.154)
Age m. 35 50			0.043	-0.036	0.044	-0.034
			(0.060)	(0.078)	(0.060)	(0.078)
Age m 50 plus			0.116	0.175	0.123 *	0.174
			(0.075)	(0.112)	(0.075)	(0.112)
Mum main parent			-0.113 ***	-0.046	-0.117 ***	-0.045
			(0.040)	(0.065)	(0.040)	(0.065)
Constant		-1.897 ***		-1.836 ***		-1.755 **
		(0.666)		(0.682)		(0.684)
Observations	8774		8774		8774	

Standard errors in parentheses *$p < 0.1$, **$p < 0.05$,***$p < 0.01$

Table 22: Likelihood of being accepted at university (opsel model, year 10, K3 grades and interaction with fsm, selected coefficients)

	1		2		3	
	accepted	selection	accepted	selection	accepted	selection
Month int	0.010	-0.026	0.008	-0.026	0.009	-0.026
	(0.012)	(0.017)	(0.012)	(0.017)	(0.012)	(0.017)
Not born uk	0.252 ***	0.637 ***	0.242 ***	0.653 ***	0.256 ***	0.640 ***
	(0.061)	(0.136)	(0.062)	(0.137)	(0.062)	(0.137)
House owned			0.033	0.080	0.031	0.077
			(0.041)	(0.055)	(0.041)	(0.055)
Number of cars			-0.025	-0.020	-0.026	-0.021
			(0.018)	(0.027)	(0.018)	(0.027)
Both parents			0.123 ***	0.039	0.117 ***	0.037
			(0.042)	(0.059)	(0.042)	(0.059)
Urban			0.010	0.065	0.007	0.066
			(0.038)	(0.054)	(0.038)	(0.054)
M. part time			0.009	-0.029	0.013	-0.026
			(0.032)	(0.048)	(0.032)	(0.048)
M. not working			0.022	0.002	0.025	0.006
			(0.040)	(0.059)	(0.040)	(0.059)
No mother			0.070	0.121	0.066	0.128
			(0.116)	(0.153)	(0.117)	(0.153)
Age m. 35 50			0.041	-0.034	0.042	-0.033
			(0.060)	(0.078)	(0.060)	(0.078)
Age m 50 plus			0.117	0.178	0.124 *	0.178
			(0.075)	(0.112)	(0.075)	(0.112)
Mum main parent			-0.117 ***	-0.050	-0.121 ***	-0.050
			(0.040)	(0.065)	(0.040)	(0.065)
Constant		-2.021 ***		-1.955 ***		-1.874 ***
		(0.669)		(0.684)		(0.686)
Observations	8774		8774		8774	

Standard errors in parentheses $*p < 0.1$, $**p < 0.05$,$***p < 0.01$

Table 23: Likelihood of being accepted at university (opsel model, year 10, K3 grades and subjects, selected coefficients)

	1		2		3	
	accepted	selection	accepted	selection	accepted	selection
Month int	0.011	-0.026	0.009	-0.026	0.010	-0.026
	(0.012)	(0.017)	(0.012)	(0.017)	(0.012)	(0.017)
Not born uk	0.259 ***	0.657 ***	0.249 ***	0.671 ***	0.263 ***	0.660 ***
	(0.061)	(0.137)	(0.062)	(0.139)	(0.062)	(0.139)
House owned			0.039	0.084	0.037	0.082
			(0.041)	(0.055)	(0.041)	(0.055)
Number of cars			-0.023	-0.018	-0.024	-0.019
			(0.018)	(0.026)	(0.018)	(0.027)
Both parents			0.128 ***	0.042	0.123 ***	0.041
			(0.042)	(0.059)	(0.042)	(0.060)
Urban			0.008	0.060	0.005	0.061
			(0.038)	(0.054)	(0.038)	(0.054)
M. part time			0.010	-0.027	0.014	-0.024
			(0.032)	(0.048)	(0.032)	(0.048)
M. not working			0.017	-0.002	0.020	0.001
			(0.040)	(0.059)	(0.040)	(0.059)
No mother			0.066	0.105	0.063	0.113
			(0.117)	(0.154)	(0.117)	(0.154)
Age m. 35 50			0.044	-0.034	0.045	-0.032
			(0.060)	(0.078)	(0.060)	(0.078)
Age m 50 plus			0.115	0.175	0.122	0.174
			(0.075)	(0.112)	(0.075)	(0.112)
Mum main parent			-0.115 ***	-0.051	-0.118 ***	-0.051
			(0.040)	(0.065)	(0.040)	(0.065)
Constant		-2.003 ***		-1.918 ***		-1.836 ***
		(0.670)		(0.685)		(0.687)
Observations	8770		8770		8770	

Standard errors in parentheses $*p < 0.1$, $**p < 0.05$,$***p < 0.01$

Table 24: Predicted probability on the expected likelihood of being accepted for university if apply (females)

	27		33		39		27		33		39	
	not at all likely						not very likely					
hwe	0.036	***	0.011	***	0.002	***	0.289	***	0.158	***	0.069	***
	(0.006)		(0.002)		(0.001)		(0.020)		(0.014)		(0.007)	
hw	0.064	***	0.021	***	0.006	***	0.365	***	0.226	***	0.112	***
	(0.007)		(0.003)		(0.001)		(0.015)		(0.011)		(0.007)	
lw	0.038	***	0.020	***	0.009	***	0.297	***	0.217	***	0.147	***
	(0.006)		(0.003)		(0.002)		(0.018)		(0.014)		(0.014)	
lwe	0.041	***	0.021	***	0.010	***	0.307	***	0.227	***	0.155	***
	(0.006)		(0.003)		(0.002)		(0.019)		(0.016)		(0.016)	
	fairly likely						very likely					
hwe	0.610	***	0.674	***	0.620	***	0.066	***	0.158	***	0.308	***
	(0.0166)		(0.006)		(0.013)		(0.010)		(0.015)		(0.020)	
hw	0.534	***	0.651	***	0.662	***	0.038	***	0.101	***	0.220	***
	(0.0162)		(0.007)		(0.007)		(0.005)		(0.008)		(0.011)	
lw	0.602	***	0.656	***	0.672	***	0.063	***	0.108	***	0.172	***
	(0.0149)		(0.008)		(0.006)		(0.008)		(0.010)		(0.017)	
lwe	0.594	***	0.652	***	0.674	***	0.058	***	0.100	***	0.161	***
	(0.0169)		(0.010)		(0.006)		(0.008)		(0.011)		(0.017)	

Standard errors (in parentheses) have been computed via the delta method

Table 25: Marginal effects of grades: females

profile	not at all likely		not very likely		fairly likely		very likely	
Effect of getting an average key stage score of 27								
hwe	0.0251	***	0.1305	***	-0.0637	***	-0.0919	***
	(0.0049)		(0.0083)		(0.0189)		(0.0083)	
hw	0.0423	***	0.1385	***	-0.1174	***	-0.0634	***
	(0.0056)		(0.0063)		(0.0128)		(0.0059)	
lw	0.0187	***	0.0797	***	-0.0538	***	-0.0447	***
	(0.0034)		(0.0092)		(0.0106)		(0.0064)	
lwe	0.0200	***	0.0808	***	-0.0586	***	-0.0422	***
	(0.0038)		(0.0093)		(0.0114)		(0.0062)	
Effect of getting an average key stage score of 39								
hwe	-0.0080	***	-0.0889	***	-0.0538	***	0.1507	***
	(0.0018)		(0.0085)		(0.0166)		(0.0091)	
hw	-0.0156	***	-0.1137	***	0.0107	ns	0.1186	***
	(0.0024)		(0.0065)		(0.0141)		(0.0087)	
lw	-0.0103	***	-0.0697	***	0.0161	*	0.0639	***
	(0.0017)		(0.0074)		(0.0096)		(0.0099)	
lwe	-0.0111	***	-0.0716	***	0.0215	***	0.0612	***
	(0.0019)		(0.0079)		(0.0102)		(0.0097)	

??

Standard errors (in parentheses) have been computed via bootstrap with 50 drawn

Table 26: Estimated selection probabilities (females)

avk3	unrate	EMA	hwe	hw	lw	lwe
27	p50	0	0.8010	0.6270	0.6475	0.5923
			(0.0332)	(0.0194)	(0.0243)	(0.0279)
33	p25	0	0.9258	0.8201	0.7883	0.7459
			(0.0162)	(0.0134)	(0.0222)	(0.0270)
33	p50	0	0.9303	0.8284	0.7975	0.7562
			(0.0154)	(0.0123)	(0.0208)	(0.0256)
33	p50	1	0.9522	0.8715	0.8458	0.8112
			(0.0119)	(0.0118)	(0.0178)	(0.0223)
33	p75	0	0.9346	0.8365	0.8065	0.7663
			(0.0150)	(0.0130)	(0.0208)	(0.0256)
39	p50	0	0.9826	0.9421	0.9010	0.8760
			(0.0050)	(0.0063)	(0.0178)	

Standard errors (in parentheses) been computed via delta methods

Table 27: Marginal effects, selection equation (females)

27		39		p25		p75		EMA1	
hwe: Non free school meal eligible, above average wealth and family education									
-0.1293	***	0.0523	***	-0.0045	ns	0.0043	ns	0.0219	***
(0.0201)		(0.0126)		(0.0030)		(0.0027)		(0.0056)	
hw: Non free school meal eligible, above average wealth									
-0.2014	***	0.1137	***	-0.0083	ns	0.0081	ns	0.0431	***
(0.0123)		(0.0127)		(0.0053)		(0.0050)		(0.0084)	
lw: Free school meal eligible, below average wealth									
-0.1500	***	0.1035	***	-0.0092	ns	0.0089	ns	0.0483	***
(0.0177)		(0.0117)		(0.0059)		(0.0057)		(0.0093)	
lwe: Free school meal eligible, below average wealth and education									
-0.1639	***	0.1197	***	-0.0103	ns	0.0101	ns	0.0550	***
(0.0176)		(0.0118)		(0.0067)		(0.0064)		(0.0107)	

Standard errors (in parentheses) have been computed via bootstrap with 50 drawns

References

Anders, j. (2012): "What's the link between household income and going to university?," Discussion paper, Institute of Education, DoQSS Working Paper No. 12-01, march 2012.

Attanasio, O., and K. Kaufmann (2009): "Educational Choices, Subjective Expectations, and Credit Constraints," Discussion paper, NBER Working Papers 15087, National Bureau of Economic Research.

Bentler, P., and D. Jackson (1971): "Identification of Content and Style: a Two-dimensional Interpretation of Acquiescience.," *Psychological Bulletin*, Vol. 76, No. 3, 186–204.

Black, S., P. Devereux, and K. Salvanes (2005): "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital," *American Economic Review*, 95, No. 1, 437–449.

Blanden, J., and S. Machin (2004): "Educational Inequality and the Expansion of UK Higher Education," *Scottish Journal of Political Economy*, 51(2), 230–249.

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

Cameron, S., and J. Heckman (1999): *Financing College Tuition*chap. Can Tuition Policy Combat Rising Wage Inequality? Washington, D.C.: AEI Press.

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

Carneiro, P., and A. Crawford, C.and Goodman (2007): "The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes," CEE Discussion Papers 0092, Centre for the Economics of Education, LSE.

Carneiro, P., and J. Heckman (2002): "The Evidence on Credit Constraints in Post-secondary Schooling," *The Economic Journal*, 112, 705–734.

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

Chowdry, H., C. Crawford, and A. Goodman (2009): "Drivers and Barriers to Educational Success Evidence from the Longitudinal Study of Young People in England," Discussion paper, DCSF research Report RR102.

Crawford, C., L. Dearden, and C. Meghir (2007): "When You Are Born Matters: The Imapct of Date of Birth on Child Cognitive Outcomes in England," CEE Discussion Papers 0093, Centre for the Economics of Education, LSE.

Cunha, F., and J. J. Heckman (2007): "The Technology of Skill Formation," *American Economic Review*, 97(2), 31–47.

——— (2008): "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, 43(4), 738–782.

Cunha, F., H. J., and S. Schennach (2010): "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 78(3), 883–931.

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

Delavande, A., and S. Rohwedder (2011): "Individuals' Uncertainty about their Future Social Security Benefits and Portfolio Choice.," *Journal of Applied Econometrics*, Vol. 26, 3.

Dinkelman, T., and A. Martínez (2011): "Investing in Schooling in Chile: The Role of Information About Financial Aid for Higher Education," Discussion paper, CEPR Discussion Paper No. DP8375.

Dominitz, J., and C. F. Manski (2011): "Measuring and interpreting expectations of equity returns.," *Journal of Applied Econometrics*, 26, 352–370.

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

Fowley (2011): "Merit Goods and the Role that Parents Play in their Children's Educational Expectations and Outcomes," Discussion paper, mimeo.

Fumagalli, L. (2012): "New information, confidence and educational expectations," mimeo.

Gouret, F., and G. Hollard (2011): "When Kahneman meets Manski: Using dual systems of reasoning to interpret subjective expectations of equity returns.," *Journal of Applied Econometrics*, 26, 371–392.

Government, H. (2009): "New Opportunities: Fair chances for the future.," Discussion paper, Norwich: The Stationery Office.

Haveman, R., and B. Wolfe (1995): "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature*, 33, No. 4, 1829–1878.

Heckman, J., and V. J. Hotz (1986): "An Investigation of the Labor Market Earnings of Panamian Males Evaluating the Sources of Inequality," *Journal of Human Resources*, Vol.21, No. 4, pp. 507–542.

Heckman, J. J. (2006): "Skill Formation and the Economics of Investing in Disadvantaged Children," *Science,*, 312(5782), 1900–1902.

Jacob, B. A., and T. Wilder (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," Discussion paper, IGIER, working paper n. 362.

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

Kreuter, F., S. Presser, and R. Tourangeau (2008): "Social Desirability Bias in CATI, IVR, and web Surveys. The Effect of Mode and Question Sensitivity," *Public Opinion Quarterly*, 72, 5, 847–865.

Leibowitz, A. (1974): "Home Investments in Children," *Journal of Political Economy*, Vol. 82, No. 2, S111–S131.

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

Madeira, C. (2009): "Estimating the beliefs of economic agents," Ph.D. thesis, Northwestern University.

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

Morgan, S. (1998): "Adolescent Educational Expectations. Rationalized, Fantasized, or Both?," *Rationality and Society*, 10, 131–162.

NatCen (2009): "Longitudinal Study of Young People in Rngland: wave one documentation. Study Number 5545.," Discussion paper, UK Data Archieve.

Office, C. (2008): "Aspiration and attainment amongst young people in deprived, communities.," Discussion paper, London: Cabinet Office.

Papay, J., R. Murname, and J. Willet (2011): "How performance information affects human-capital investment decisions: the impact of test score labels on educational outcomes," Discussion paper, NBER working paper 17120.

Paulhus, D. L. (1991): *Measures of Personality and Social Psychological Attitudes,*chap. Acquiescence. London: Academic Press.

Raghunathan, T., J. Lepkowski, J. Van Hoewyk, and P. Solenberger (2001): "A multivariate technique for multiply imputing missing values using a sequence of regression models.," *Survey Methodology*, 27 (1), 85–95.

Reynolds, J., and J. Pemberton (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.

Saris, W., M. Revilla, J. Krosnick, and E. M. Shaeffer (2010): "Comparing Questions with Agree/Disagree Response Options to Questions with Item-Specific Response Options," *Survey Research Methods*, Vol.4, No.1, 61–79.

Stinebrickner, T. R., and R. Stinebrickner (2009): "Learning about Academic Ability and the College Drop-out Decision," Working Paper 14810, National Bureau of Economic Research.

Strand, S. (2007): "Minority Ethnic Pupils in the Longitudinal Study of Young People in England.," Discussion paper, London. DCSF.

Tourangeau, R., L. Rips, and K. Rasinski (2000): *The Psychology of Survey response*. Cambridge University Press.

Winswall, M., and B. Zafar (2011): "Belief Updating among College Students: Evidence from Experimental Variation in Information," Discussion paper, Federal Reserve Bank of New York , Staff report no 516.

Zafar, B. (2011): "How Do College Students Form Expectations?," *Journal of Labor Economics*, 29(2), 301 – 348.