



The influence of socioeconomic status on changes in young people's expectations of applying to university

Jake Anders

To cite this article: Jake Anders (2017) The influence of socioeconomic status on changes in young people's expectations of applying to university, Oxford Review of Education, 43:4, 381-401, DOI: [10.1080/03054985.2017.1329722](https://doi.org/10.1080/03054985.2017.1329722)

To link to this article: <http://dx.doi.org/10.1080/03054985.2017.1329722>



© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 22 Jun 2017.



Submit your article to this journal [↗](#)



Article views: 629



View related articles [↗](#)



View Crossmark data [↗](#)

The influence of socioeconomic status on changes in young people's expectations of applying to university

Jake Anders 

UCL Institute of Education, UK

ABSTRACT

A much larger proportion of English 14-year-olds expect to apply to university than ultimately make an application by age 21, but the proportion expecting to apply falls from age 14 onwards. In order to assess the role of socioeconomic status in explaining changes in expectations, this paper applies duration modelling techniques to the Longitudinal Study of Young People in England, analysing transitions in young people's expectations both from being 'likely to apply' to being 'unlikely to apply' and vice versa. Young people's socioeconomic background has a significant association with changes in expectations, even after controlling for prior academic attainment and other potentially confounding factors; in addition, young people's backgrounds affect their responsiveness to new evidence on academic attainment at age 16. This suggests more could usefully be done to maintain the educational expectations of academically able young people from less advantaged families, especially providing guidance on how to view new academic results.

KEYWORDS

Educational expectations;
socioeconomic status;
higher education; duration
modelling

1. Introduction

There is a large socioeconomic gradient in university attendance in England. Much of this gap can be explained by differences in academic achievement that emerge long before the point at which young people apply to university (Chowdry et al., 2013). However, there remains a socioeconomic gradient in university application (Anders, 2012a), despite the fact that a larger proportion of English 14-year-olds from disadvantaged backgrounds expect to apply to university than the overall proportion who have ultimately done so by age 21 (Anders & Micklewright, 2015, pp. 42–43).

This raises the question of when and why young people from less advantaged families change their minds about making an application to university. Are their changes in expectations explicable by other factors, such as academic attainment, or does socioeconomic status continue to have an influence? Given the previous evidence that much of the socioeconomic gap in university attendance opens at or before the point of application, a better understanding of the dynamics of expectations is of significant importance to the formulation of policy on reducing inequality in access to higher education.

CONTACT Jake Anders  jake@jakeanders.uk

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Rather than following previous authors in using expectations data as an explanatory factor for later outcomes, this paper takes a step back, addressing the issue by directly analysing the influence of socioeconomic status on the changes in young people's expectations of applying to university between ages 14 and 17. Using rich panel data from a recent English cohort, it takes the novel approach of using duration modelling to analyse the dynamics of young people's expectations.

The research question and data lend themselves naturally to this approach. Duration modelling allows the flexibility to make use of all available information on the timing of events. The technique also allows separate analysis of transitions from being 'likely to apply' to being 'unlikely to apply' and vice versa. This is important, since it is quite possible that the factors which cause young people to start thinking that they are likely to apply to university are quite different from the causes of movement in the other direction. Despite this, duration modelling is not regularly used in such settings and has not been used before to model changes in young people's educational expectations over time.

This paper makes an important contribution to the literature on higher education access. It provides non-parametric estimates of changes in young people's expectations between the ages of 14 and 17, quantifying the extent of change during this period. Making minimal assumptions, this technique is also used to examine whether young people from less advantaged backgrounds are more likely to stop, and less likely to start, thinking they will apply to university than their more advantaged peers. Furthermore, taking advantage of the rich survey data and retaining the flexibility of duration modelling, this paper provides estimates of the continued influence of socioeconomic status, after controlling for potentially confounding factors, including prior attainment. Finally, it explores the interplay between SES and new information on academic attainment at age 16.

The paper proceeds as follows. Section 2 reviews the literature on the socioeconomic patterning of educational expectations and lays out the empirical strategy for identifying the influence of socioeconomic status on changes in expectations. Section 3 describes the dataset and measures used in this paper, and highlights the advantages of using duration modelling for this context. Non-parametric duration modelling methods are then applied in Section 4 to explore how young people's expectations change during their teenage years and the association with socioeconomic status. This initial analysis is extended through use of regression models, introduced in Section 5, and with the results reported in Section 6. Finally, Section 7 concludes.

2. Background

Studying expectations is not worthwhile if they are just individuals' whims. However, Morgan (1998) argues that 'educational expectations are not "flights of fancy" or "vague preferences" [but rather,] because they can be explained by a reasonable theory of rational behavior, should be considered rational' (p. 157) and hence, presumably, informative. Certainly, previous work has shown a correlation between educational expectations and later outcomes in several developed countries around the world (Anders & Micklewright, 2015; Chowdry et al., 2011; Khoo & Ainley, 2005; Reynolds & Pemberton, 2001). Why do these associations exist?

Causal explanations focus on narratives such as that young people who 'expect less of themselves [...] may not fully develop their academic potential because they see little hope

for ever being able to complete college or use their schooling in any effective way' (Cameron & Heckman, 1999, p. 86), or that 'entering high school a student with college expectations may be placed in the college-prep high school program as opposed to the general program' (Jacob & Wilder, 2010, p. 5). However, others are highly critical of the jump from these plausible explanations and observed correlations to treating the relationship as truly causal (Cummings et al., 2012). Gorard (2012) argues that formulating policy on this basis, when evidence of causation is so weak, is misguided.

This paper, rather than attempting to identify the effect of young people's expectations on university attendance, takes a step back. It explores the role of socioeconomic status (SES) in determining the paths of young people's expectations in the first place; expectations are, thus, an outcome. There is no need to take a view on whether or not expectations have a causal impact on academic attainment and progression. Instead, it is enough to be convinced that expectations are symptomatic of the underlying social processes leading from SES, prior attainment, and other background characteristics to the ultimate decision as to whether or not to apply to university. Expectations are indicators of young people's likelihood of going to university, which cannot necessarily be manipulated causally, but provide timely information on whether an individual remains on track to attend university.

There exists a previous literature on the formation and correlates of young people's educational expectations and aspirations (Anders & Micklewright, 2015; Baker et al., 2014; Chowdry, Crawford, & Goodman, 2010; Rampino & Taylor, 2013). Previous work has always found a role for SES. Kao and Tienda (1998, p. 370) find that SES 'exerts a strong influence on educational aspirations and is vital to their maintenance'. Fumagalli (2012) finds that young people from higher SES families are more responsive to new information about their academic attainment in updating their expectations of applying to university. Strand (2011) reports a correlation between aspirations and both attainment at age 14 and academic progress between ages 11 and 14.

This paper develops the literature in two important respects. First, through use of duration modelling, this paper analyses the dynamic relationship between SES and expectations in a flexible way. Importantly, it allows for different relationships between characteristics of interest and whether young people make a transition depending on direction of the transition. Second, while both Kao and Tienda and Rampino and Taylor focus on aspirations rather than expectations, and Fumagalli focuses on formation of young people's expectations of being admitted to university, the focus of this paper is on expectations of applying to university.

The empirical strategy to isolate the role of SES is to control for a rich set of other characteristics, including young people's age, academic ability, demographic characteristics, school characteristics, traumatic experiences, and local labour market conditions. However, there are several challenges to achieving this. Most fundamentally, one cannot be sure that other unobserved or unobservable characteristics are not leading to omitted variable bias. In the absence of exogenous variation in SES (which is conceptually, let alone practically, challenging) one cannot be certain that this problem has been dealt with. One strategy is to use individual-level random effects to deal with unobserved heterogeneity. The results obtained when applying this approach (allowing the random effects to have either normal or discrete mixing distributions) do not substantively alter the findings, giving additional confidence to the conclusions.

3. Data and duration modelling

This paper applies duration modelling to analyse changes in young people's expectations. This technique is not commonly applied in this setting (Alcott, 2013, pp. 50–51). It is, however, extremely well suited to analysing changes in young people's expectations, since this is a setting where the timing of changes is central to the question.

This section has two intertwined roles, introducing both the data and the duration modelling approach. First, it discusses the measurement of young people's expectations of applying to university (the outcome variable). Second, it describes measurement of the main explanatory variable of interest (young people's SES). In order to isolate the influence of SES it is also important to be able to control for other factors that may influence expectations; the dataset includes a rich set of participant characteristics and experiences but these are discussed as they are introduced to the model in Section 5. Finally, particular features of duration modelling, their relevance to this analysis, and to these data, are highlighted.

The Longitudinal Study of Young People in England (LSYPE) is a major panel survey, tracking the experiences of one cohort of young people over seven years, from approximately age 14 (in 2004) to age 20 (in 2010), including annual interviews with young people themselves (throughout) and their parents (up to age 17). It includes a wide variety of data on participants, including details of their SES, educational attainment, and educational expectations. A more in-depth description is conducted by Anders (2012b).

Only individuals present in both Waves 1 and 2 are included to ensure that at least one potential transition is observed for all individuals included in this analysis. As such, the data are weighted using the LSYPE-provided attrition and non-response weights for Wave 2. The findings are robust to using the sub-sample still responding in Wave 4 applying the relevant weights.

The LSYPE records university application expectations from age 14, the age at which previous literature has argued aspirations are refined into expectations (Gottfredson, 2002; Gutman and Akerman, 2008). As such, periods of reporting expectations are treated as starting at this point at the earliest. However, selection into these initial states is non-random. Modelling this 'initial conditions problem' (Hsiao, 1986), as part of a dynamic random effects probit model, suggests this does not undermine our results, although this cannot, of course, account for unobservables.

Young people's expectations of applying to university are measured through one question, 'How likely do you think it is that you will apply to university?', with four response options: 'very likely', 'fairly likely', 'not very likely', and 'not at all likely'. This is dichotomised into a distinction between 'likely' ('very likely' or 'fairly likely') or 'unlikely' ('not very likely' or 'not at all likely') to apply to university.

Duration modelling is typically employed to analyse transitions between defined states such as employment and unemployment. Use of a 'stated preference' measure as an outcome variable in this way is innovative, although some precedent is provided by studies of the dynamics of poverty (e.g. Bane & Ellwood, 1986) where measurement of income may affect movement in or out of poverty, but increased measurement error is likely to bias overall transition rates upwards. Another challenge when analysing expectations, rather than observed behaviour, is that cognitive biases, such as social desirability bias, may affect responses. However, young people's reports do seem informative: 64% of those who say

they think it is 'likely' that they will apply to university at age 14 have done so by the last point of observation, while only 22% of those saying it is 'unlikely' have done so.

For an initial impression of the evolution of young people's expectations during this period, Figure 1 shows for each wave, 1–7, the percentages of young people who report being 'very likely', 'fairly likely', 'not very likely', and 'not at all likely' to apply to university. From Wave 5 onwards there is an additional category for those who have actually applied. In Wave 7, only this measure of having actually applied is available. The overall percentage who are 'likely' (or who have already applied) can be calculated by adding the percentages above the 'very likely', 'fairly likely', and 'have applied' blocks in Figure 1.

Overall, this declines from 68% in Wave 1 to 57% in Wave 4 (the end of the first year following compulsory education), followed by essentially no change in Wave 5 (when actual applications begin to be included), and a small rise in the following year. As the aim of this paper is to understand changes in young people's expectations in the period leading up to making an application, the analysis is deliberately curtailed at the last wave in which individuals have not yet started applying to university (Wave 4, roughly age 17).

Figure 2 presents the 10 most common sequences of individuals' expectations between ages 14 and 17 observed in the dataset (approximately 85% of the sample). The most frequent sequence of expectations (40% of the sample) is for individuals to report being 'likely to apply' at every interview; the second most frequent (17% of the sample) is reporting being 'unlikely to apply' at every interview. Note that the absence of a line is important in itself so that, for example, there is a difference between sequence 4 and 7, since in the former expectations are observed to change from 'likely' to 'unlikely' at age 17, while in the latter the response at age 17 is missing.

Table 1 reports summary statistics for individuals who have the sequences in Figure 2, plus a category for all remaining sequences. Individuals who always report being likely to

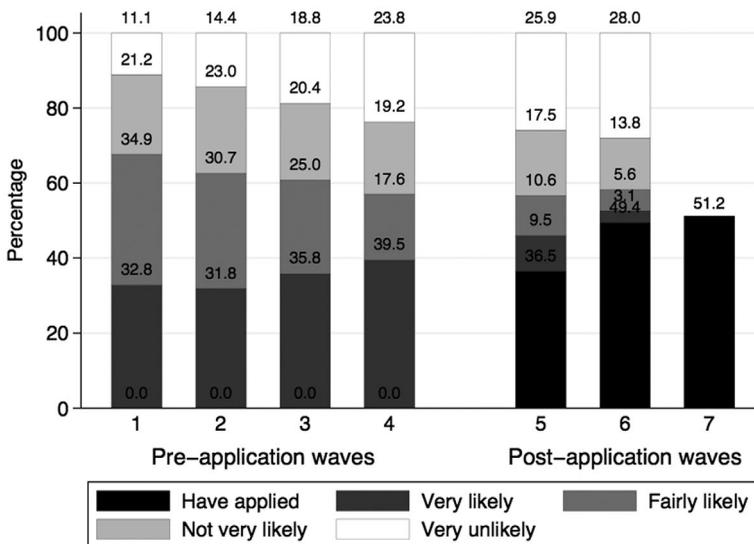


Figure 1. Young people's expectations of university application, Wave 1 (age 13–14) to Wave 7 (age 19–20). Notes: Sample: Wave 7 respondents with non-missing data on university expectations and university application at each wave (complete case analysis). 'Don't know' (4.4% of weighted Wave 1 respondents) treated as 'not very likely'. Wave 7 attrition and non-response weights applied. Unweighted sample size = 8029. Data labels show cumulative percentages.

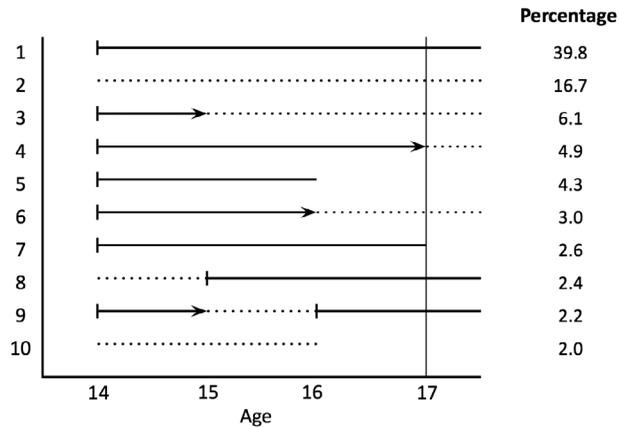


Figure 2. Most common sequences of individuals' expectations from age 14 to 17 in the Longitudinal Study of Young People in England.

Notes: Solid line indicates 'likely to apply', dotted line indicates 'unlikely to apply' and absence of a line indicates no report at the most recent wave. Transitions are highlighted: arrow tail highlights a negative outcome in the previous wave; arrow head highlights a negative outcome in the following wave. Vertical line at age 17 highlights the final point of observation and hence data beyond this point only provide information on whether the spell was censored (whether by no change or missing data) at this point. Frequency of spell types weighted using LSYPE Wave 2 attrition and non-response weights.

Table 1. Summary statistics about sequences of expectations.

Group	N	Percentage	SES index
1	4503	40.2	0.45
2	1857	16.6	-0.49
3	673	6.0	-0.35
4	547	4.9	-0.07
5	478	4.3	-0.23
6	342	3.1	0.04
7	279	2.5	-0.04
8	269	2.4	-0.53
9	249	2.2	0.05
10	225	2.0	-0.27
Other	1828	15.9	-0.30
All	11,249	100.0	0.00

Notes: Adjusted using LSYPE-provided Wave 2 survey design, attrition and non-response weights. Individuals with missing data in either of Waves 1 or 2 are excluded.

apply to university (type 1) are, on average, half a standard deviation more advantaged than the sample as a whole. Conversely, those who always report being unlikely to apply (type 2) are roughly the same amount less advantaged than the sample as a whole.

While an individual's changes in expectations are best thought of as a continuous underlying process, they are only reported once a year in the data. This is illustrated in Figure 2: spells are only observed to start or end at exact ages, never somewhere in between, even though the reality is, of course, different. It is, therefore most appropriate to think of this as a continuous-time duration model, but to apply discrete-time methods in order to take into account the structure of the data (Allison, 1982, p. 63). An important limitation of this data structure is that, as some transitions back and forth may occur between observation points (interval censoring), overall transition rates could be biased downwards.

The LSYPE includes a rich set of data with which to measure young people's socioeconomic status (SES), including household income, parental education, and parental occupational status, all of which are important in measuring SES (Hauser, 1994).

Household income is measured at each wave between 1 and 4. As previous studies have highlighted the particular role of permanent, rather than transitory, income on educational outcomes (Jenkins & Schluter, 2002, p. 2), an approximation of the household's equivalised 'permanent' income is made by averaging across these four measures and dividing by the square root of household size.

Parental education is likely to play a role in the formation of young people's educational expectations (Ganzach, 2000), not least because young people whose parents went to university are more likely to see it as a natural next step in their education. Indeed, the descriptive differences reported in Table 2 bear this out.

Social class is also a key element of an individual's SES (Goldthorpe & McKnight, 2004). In particular, as 'young people (and their families) have, as their major educational goal, the acquisition of a level of education that will allow them to attain a class position at least as good as that of their family of origin' (Breen & Yaish, 2006, p. 232), individuals from different class backgrounds will have, on average, different educational expectations. Parents' occupational status is recorded in the LSYPE using the National Statistics Socio-Economic Classification (NS-SEC), designed to capture social class differences between occupational types (Rose & Pevalin, 2001).

The above measures are combined using principal component analysis with a polychoric correlation matrix (Kolenikov & Angeles, 2009; Olsson, 1979) to construct a single SES index. This explains roughly three quarters of the variation in the three individual measures, but provides a broader measure of family circumstances than any one measure would provide. Table 3 reports characteristics of the median member of quintile groups of the index.

Regression-based duration models allow for individuals to switch back and forth between 'likely' and 'unlikely', which simple discrete-choice modelling would ignore. Dynamic discrete

Table 2. Summary statistics of sample by whether young person reports being likely or unlikely to apply to university at age 14.

Variable	Mean of 'Unlikely'	Mean of 'Likely'	Mean of whole sample	Standard deviation
SES index (Z-score)	-0.40 (0.02)	0.20 (0.02)	0.00 (0.02)	1.00
Equivalised family permanent income (£)	12,464.07 (209.35)	18,029.33 (256.24)	16,199.21 (208.44)	12,220.12
Parent has higher education	0.06 (0.00)	0.25 (0.01)	0.19 (0.01)	0.39
Parent has higher occupational status	0.26 (0.01)	0.49 (0.01)	0.41 (0.01)	0.49
Lone parent	0.28 (0.01)	0.20 (0.01)	0.22 (0.00)	0.42
Gender: male	0.55 (0.01)	0.48 (0.01)	0.51 (0.01)	0.50
Ethnicity: non-white	0.07 (0.00)	0.16 (0.01)	0.13 (0.01)	0.34
Age 11 attainment Z-score	-0.48 (0.02)	0.23 (0.02)	-0.00 (0.02)	0.97
Age 16 attainment Z-score	-0.60 (0.03)	0.29 (0.02)	-0.00 (0.02)	1.00
Attend independent school	0.02 (0.01)	0.10 (0.01)	0.07 (0.01)	0.26
Attend grammar school	0.01 (0.00)	0.05 (0.01)	0.04 (0.01)	0.19
Attend school with sixth form	0.52 (0.02)	0.56 (0.02)	0.55 (0.02)	0.50
Local unemployment rate (%) at age 14	4.61 (0.07)	4.80 (0.07)	4.74 (0.06)	2.14
N	3686	7523	11,209	

Notes: Weighted using LSYPE Wave 2 sample design and non-response weighted weights. Standard errors, clustered by school, in parentheses. Household income is equivalised by dividing by the square root of household size.

Table 3. Median family characteristics by quintile group of socioeconomic status index.

Quintile group	Q1	Q2	Q3	Q4	Q5
Parental education	<A*-C GCSE	A*-C GCSE	A-level	HE<Degree	Degree
Occupational status	Routine occupations	Routine occupations	Intermediate occupations	Higher occupations	Higher occupations
Family income (£)	5699	9549	12,992	16,433	29,941
N	2585	2221	2171	2201	2071

Notes: Adjusted using LSYPE-provided Wave 2 survey design, attrition, and non-response weights. Standard errors, clustered by school, in parentheses. Family income is equivalised by dividing by the square root of household size. Sample: Wave 2 respondents with non-missing data on university expectations ('don't know' treated as 'not very likely') and university applications.

choice modelling would also capture this feature of the relationship, but wouldn't allow for the possibility of asymmetric relationships between independent variables and transitions. (In any event, such models produce results that tell a very similar story to that emerging from the duration models reported in this paper.) Furthermore, a duration modelling approach is able to take into account 'censoring', where the start and/or end points of a spell are not observed, meaning that the true length of the spell is unknown.

Not observing the end of a spell is referred to as 'right censoring', which occurs in the final report for all individuals, whether this is due to the end of the period under analysis (at age 17) or earlier as a result of attrition. As with any longitudinal survey, the LSYPE suffers from attrition. Treating attrition as 'censoring' is preferable to the alternative of restricting attention to a complete case sample.

All of these features are important in fitting the most appropriate model to understand changes to young people's expectations during these critical years for their education.

4. Analysis of changes in expectations

Non-parametric estimates of the probability that spells have not ended with a transition by a given age are a useful way to begin exploring transitions in young people's expectations. These are first estimated for the sample as a whole, then by sub-samples defined by SES in order to provide a simple way of assessing transition rates by family background: for ease of interpretation, while maintaining some continuity with the quintile groups used in later models, SES is dichotomised into 'high' (comprising the top 40% of the distribution of the SES index) and 'low' (comprising the bottom 60% of the distribution), although the results are robust to splitting at different points.

The spells under analysis are restricted to those beginning at age 14 (the start of the dataset) in order to produce estimates using this method. By definition, this also means concentrating on an individual's first spell. Among the benefits of the regression-based analysis introduced in Section 5, these restrictions will be relaxed. Inference testing is performed using Cox-regression based tests, which makes the proportional hazards assumption; the non-parametric log-rank test is 'not appropriate' with sampling weights (StataCorp, 2013, p. 446).

We first consider the transition from reporting being 'likely to apply' to reporting being 'unlikely to apply'. Relating this to the sequences of expectations shown in Figure 2, this means including the first (or only) spell of individuals of type 1, 3, 4, 5, 6, 7, or 9 (amongst others not shown in the diagram), but not the spell that type 8 spends reporting being 'likely

to apply'. Nevertheless, this includes over 70% of the individuals in the data, with much of the remainder being individuals who never report being 'likely to apply' rather than individuals who are excluded simply because of this restriction.

The solid line on Figure 3 shows that roughly a third of the observed periods of being 'likely to apply' end by age 17. There are evidently a significant number of transitions during this stage of life. However, this sheds no light on the reasons for these changes, other than ageing. Comparing the dashed (lower SES) and dotted (higher SES) lines shows that individuals from lower SES households are more likely to make a transition to reporting 'unlikely to apply' than their richer counterparts throughout the period under analysis: almost half of those from lower SES backgrounds have made a transition from 'likely to unlikely' by age 17, whereas only around 25% of those from high SES backgrounds have done so. A Cox-regression based test suggests rejecting the null hypothesis of no difference between the two estimated survivor functions.

It is possible that the relationship between SES and young people raising their expectations is quite different from that associated with movement in the opposite direction. The analysis of this transition includes the first (or only) spell from individuals of types 2, 8, and 10 in Figure 2, but not the spell that types 3, 4, 6, and 9 spend reporting being 'unlikely to apply'. This represents over 20% of the overall sample, but much of the remainder again comprises individuals who never report being 'unlikely to apply', rather than exclusions because of restricting to spells that start at age 14.

The proportion of spells of 'unlikely to apply' that do not end in transition to being 'likely to apply' by a given age is reported by the solid line in Figure 4. Around half of spells have ended in transition by the last point of observation at age 17. These are higher rates of

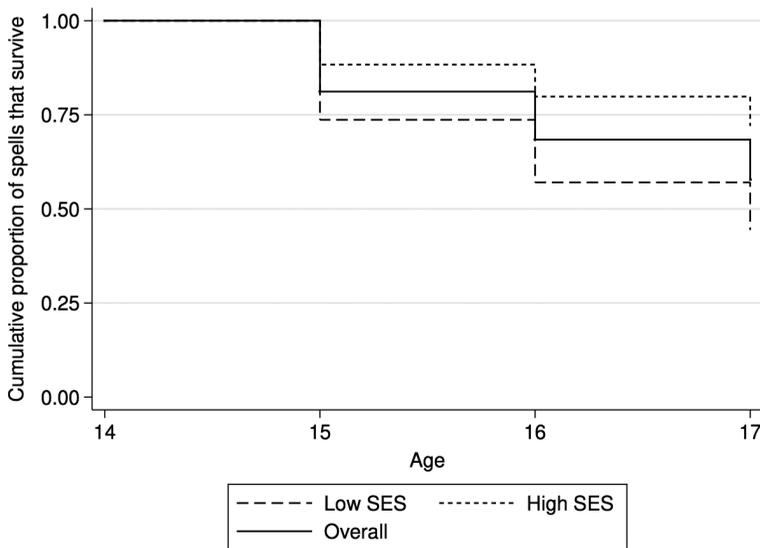


Figure 3. Probability that an individual who reports being 'likely to apply' at age 14 has not moved to reporting that they are 'unlikely to apply', by age and household SES.

Notes: Kaplan-Meier estimated survivor function. Excludes spells beginning after age 14. Analysis weighted using Wave 2 sample design and non-response weights. 'High SES' denotes individuals in the top two quintiles of SES, while 'low SES' refers to all other individuals. Unweighted number of subjects: 6129; weighted number of subjects: 6009. Cox regression-based test for equality of survivor functions rejects the null hypothesis of no difference ($p < 0.01$).

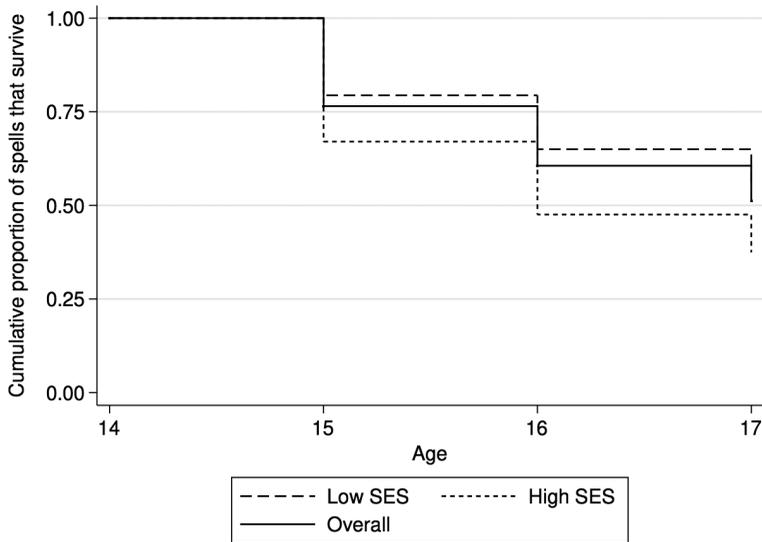


Figure 4. Probability that an individual who reports being ‘unlikely to apply’ at age 14 has not moved to reporting that they are ‘likely to apply’, by age and SES.

Notes: Kaplan-Meier estimated survivor function. Excludes spells beginning after age 14. Analysis weighted using Wave 2 sample design and non-response weights. ‘High SES’ denotes individuals in the top two quintiles of SES, while ‘low SES’ refers to all other individuals. Unweighted number of subjects: 2556; weighted number of subjects: 2946. Cox regression-based test for equality of survivor functions rejects the null hypothesis of no difference ($p < 0.01$).

transition than those seen for the same time points in the analysis of the transition from ‘likely to unlikely’ above, despite a larger overall shift in the opposite direction. Comparing the two dashed (lower SES) and dotted (higher SES) lines shows that there are clear socio-economic differences in the expected proportion of transitions from ‘unlikely’ to ‘likely’. However, in this case those from the less advantaged groups are less likely to make a transition out of being ‘unlikely’ than their more advantaged peers. Again, a Cox regression-based test suggests rejecting the null hypothesis of no difference between the two survivor functions.

Comparing Figure 4 with Figure 3, it is clear that the differences in rates of transition from being ‘unlikely’ to being ‘likely’ by SES are markedly smaller than for the transition in the opposite direction: by age 17 around 35% of those from lower SES backgrounds have made a transition from ‘unlikely to likely’, while just over 50% of those from more advantaged backgrounds have done so. This suggests that more of the inequality in expectations builds from less advantaged individuals having a higher probability of switching to reporting being ‘unlikely’, than from movements in the other direction.

However, this analysis has limitations: it cannot accommodate spells that started after age 14 or, hence, multiple spells from one individual. Furthermore, it does not allow me to control for additional covariates. Regression-based duration modelling techniques relax these limitations.

5. Regression modelling

A regression-based duration model may be estimated as a binary choice model applied to a dataset organised such that there is one observation for each time point that each individual is 'at risk' of making the relevant transition (Jenkins, 1995). This exposition concentrates on the transition from 'likely to apply' to 'unlikely to apply' to avoid unnecessary duplication; it is easy to see how the model is modified for the transition from 'unlikely to apply' to 'likely to apply'.

The outcome variable is an indicator of whether an individual reports being 'unlikely to apply' to university:

$$Y_{it} = 1 \text{ if young person } i \text{ is unlikely to apply to university at time } t$$

$$= 0 \text{ if young person } i \text{ is likely to apply to university at time } t$$

It only makes sense to include in modelling individuals who are 'at risk' of making a transition from 'likely to unlikely'. Variable d_{it} is defined as an indicator of whether an individual makes the transition at a given time point, given that the individual was at risk of making the transition:

$$d_{it} = 1 \text{ if } Y_{it} = 1 \cap Y_{it-1} = 0$$

$$= 0 \text{ if } Y_{it} = 0 \cap Y_{it-1} = 0$$

Including a piecewise constant function of age with age 15 the baseline category minimises the need to make functional form restrictions on the association between age and transition. As a large proportion of spells start at the same point in time (age 14), age and duration are highly collinear, meaning that modelling the effect of the length of time individuals have spent in their current state is not possible here.

As this is continuous-time modelling applied to discrete-time data it is most appropriate to use complementary log–log regression models (Allison, 1982, p. 72), although using logistic regression makes little difference. School-level cluster-robust standard errors are calculated.

Using these variables and \mathbf{x} , a vector of time-invariant and time-varying control variables (discussed further below), models of the following form are estimated:

$$\log(-\log(1 - d_{it})) = \alpha(A_{it}) + \beta\mathbf{x}_{it}$$

M0 is a baseline model, including month of birth and month of interview (to adjust for differences in age at time of interview) and the age function. This places the survivor functions from Section 4 into this regression framework, allowing for multiple spells from one individual and spells that begin later than age 14, and allows inspection of the raw coefficients on age, providing insight into when adjustment of expectations most often occurs.

The first model of substantive interest (M1) attempts to capture the 'total' association between SES and the probability of transition between being 'likely' and 'unlikely' to apply. Compared to M0, M1 adds dummy variables indicating an individual's SES quintile group, measured using the index described in Section 3.

The second model (M2) estimates the 'conditional' association between SES and transition. The following controls (predominantly measured at age 14) are added: demographic

characteristics (gender, ethnicity, government office region, number of siblings, and number of older siblings), school characteristics (independent school, grammar school, presence of sixth form), traumatic experiences (time-varying measures of worklessness and family separation), and local labour market conditions (youth unemployment). This is expected to reduce the conditional association between SES and probability of transition. However, any causal effect of SES on these variables (for example on sorting into secondary schools) could result in underestimation of the influence of SES.

The third model (M3) again estimates the 'conditional' influence of SES, this time also controlling for prior academic attainment, specifically young people's average performance in English, maths, and science at age 11. Attainment provides an imperfect proxy for ability: SES is likely to have an effect on attainment measures, suggesting that models including attainment may underestimate the influence of SES. Nevertheless, this is the preferred specification for identifying the 'conditional' effect of SES on changes in young people's expectations of applying to university.

The final two models specifically address whether changes in young people's expectations are affected by new information on academic attainment provided by examinations at age 16. The first (M4) adds to M3 a standardised variable for academic attainment at this time, interacted with the age variable indicating individuals have received their results (age 17). The coefficient on this variable is an estimate of association between a one standard deviation increase in young people's performance at age 16 and the probability of transition, conditional on family background and attainment at age 11. However, in interpreting this finding, it is important to note that individuals' performance in examinations at 16 is likely to be endogenous: young people's expectations of applying to university are likely to affect their effort at school and hence performance in these examinations. As such, the results should only be used as indicative for the question of responsiveness to new information on academic attainment; results from M3 are likely to be a more reliable guide to the association between SES and changes in expectations.

The final model (M5) builds on M4, but relaxes the implicit assumption that this new information on academic performance affects all young people in the same way. An interaction between KS4 performance and SES is added, allowing exploration of whether individuals are more or less likely to adjust their expectations in response to their results depending on their SES background. The same caveats apply in terms of the endogeneity of performance at age 16.

Given the complexity of interpreting interactions, and in the interests of parsimony, variants of M4 and M5 in which the dummy variables for each quintile group of SES have been replaced by a standardised continuous SES index variable are also estimated. This comes at the cost of assuming a linear relationship between SES and the risk of transition. However, robustness checks suggest that this does not seem to affect the overall narrative. As such, in discussing the results, these variants, referred to as M4C and M5C, are the focus.

6. Results

In models focusing on the influence of SES on transitions (M1–M3), the hazard ratios (exponentiated coefficients from the underlying regression models) are reported for each quintile group of SES, relative to a baseline category of the middle (third) quintile group. These may be interpreted as the probability that an individual in the relevant SES quintile group makes

a transition, divided by the probability that an individual in the middle SES quintile group makes a transition. As they have been exponentiated, comparisons between other groups are calculated by multiplying or dividing (rather than adding or subtracting) the hazard ratios. For example, to compare the hazards in the highest and lowest quintile groups we divide the hazard ratio for the lowest quintile group by that for the highest. In order to examine the overall patterns of young people's transitions as they age, hazard ratios from each model associated with each age are also reported, relative to a baseline of the period between the interview at age 14 and age 15.

In models focusing on the responsiveness of young people to new information about their academic attainment (M4C and M5C), the hazard ratio associated with change in SES and the hazard ratio associated with change in both SES and KS4 performance are both reported. Direct interpretation of interactions in non-linear regression models is complex. Instead, we estimate predicted probabilities of transition for an individual with 'low SES' (one standard deviation below the mean) and an individual with 'high SES' (one standard deviation above the mean) and how these change when age 16 scores increase by one standard deviation, holding all other characteristics constant at their means.

The transition from 'likely to unlikely' is reported first, in Table 4. The results from M0 suggest that individuals are most likely to make a transition between age 14 and 15 and the rate of transitions slows after this point. This reflects the findings in Figure 3. However, it could be that individuals who are most likely to make a transition have already done so before later time points, hence the remaining sample 'at risk' are less likely to make a transition (Jenkins, 2004). Controlling for factors associated with this compositional change may, therefore, reduce the apparent effect of age.

Table 4. Estimated hazard ratios of transition from reporting being likely to apply to reporting being unlikely to apply by quintiles of socioeconomic status.

	M0	M1	M2	M3	M4
Age 16	0.89 (-2.41)**	0.90 (-2.33)**	0.91 (-2.07)**	0.95 (-1.08)	0.94 (-1.22)
Age 17	0.74 (-6.60)***	0.77 (-5.82)***	0.85 (-3.48)***	0.92 (-1.83)*	0.92 (-1.76)*
SES Q1 (Low)		1.46 (6.33)***	1.54 (6.59)***	1.13 (1.80)*	1.10 (1.42)
SES Q2		1.40 (5.61)***	1.31 (4.49)***	1.17 (2.53)**	1.16 (2.42)**
SES Q4		0.75 (-4.76)***	0.80 (-3.69)***	0.80 (-3.67)***	0.80 (-3.71)***
SES Q5 (High)		0.33 (-13.45)***	0.39 (-11.89)***	0.47 (-9.66)***	0.47 (-9.59)***
Significance of SES ($P > F $)		0.00	0.00	0.00	0.00
N	9247	9247	9247	9247	9247
<i>Variables</i>					
Age	√	√	√	√	√
SES Quintile dummies		√	√	√	√
Demographics & school			√	√	√
Prior attainment				√	√
Age 16 attainment					√

Notes: Reporting hazard ratios. $P > |F|$ shows p -value from joint significance test of the hypothesis that exponentiated coefficients on all SES group dummies in the underlying conditional log-log regression model are equal to 1. Adjusted using LSYPE-provided Wave 2 survey design and non-response weights. Student's t -test statistics, based on standard errors clustered by individual's school, reported in parentheses. Estimated risks are relative to base categories of Age 15 and SES quintile group 3.

In the first model including SES (M1), the estimated hazard ratios for each of the quintile groups of SES are all statistically significant. Young people from less advantaged backgrounds are more likely to switch from 'likely' to 'unlikely', with the least advantaged group more than four times as likely than the most advantaged one. The size of the change between each quintile group increases higher up the SES distribution: the difference between Q1 and Q2 is only equivalent to a 5% reduction in the probability of transition, while the difference between Q4 and Q5 is equivalent to more than a 50% reduction.

Given previous evidence on young people's expectations of applying to university by SES, the strong relationship is unsurprising. However, the aim in the following models is to assess what, if anything, explains these gaps and whether the SES gradient persists once we have controlled for other factors.

Adding demographic and school characteristics, in M2, there is some reduction in the total socioeconomic inequality: the probability of an individual from the least advantaged quintile group switching from 'likely to unlikely' is now estimated to be just under four times greater than the probability of an individual in the most advantaged group doing so. Several of the added covariates (notably including gender, ethnicity, and school characteristics) have large hazard ratios.

Inclusion of prior attainment, in M3, also reduces the estimated influence of SES. There is no longer a significant difference between the lowest two quintile groups; conditional on other characteristics, young people in the bottom 40% of the SES distribution have approximately 40% higher probability of making this transition than individuals in the middle. Being in a higher SES group continues to be associated with large reductions in the risk of transition: young people in the highest group still have approximately 60% of the probability of making a transition as individuals in the middle.

Furthermore, introducing prior attainment reduces estimated differences in the probability of transition by age to statistical insignificance. Seemingly, the apparent influence of age on likelihood of transition was driven by the reduced presence in the sample of individuals with lower prior attainment by later time points.

In summary, there continues to be a strong relationship between young people's socioeconomic background and their probability of continuing to report being 'likely to apply' to university. Individuals from the least advantaged fifth of the SES distribution still have almost 2.5 times the probability of making a transition as individuals in the most advantaged quintile group.

What explains the reduction in the size of the SES gap once prior attainment has been included? Perhaps young people from less advantaged backgrounds are less likely to have achieved strong results at age 16, for whatever reason. It could also be that their expectations are more sensitive to the results that they receive. The final models aim to shed light on this question.

Given the likely endogeneity of performance at age 16, estimates from M3 are a better guide to the 'conditional' association between SES and the probability of transition than those from M4, although there are only slight changes in practice. At this point, for parsimony and ease of interpretation, models are reported in which SES enters using the linear index variable defined in Section 3. Comparing the results of M4 (final column of Table 4) and M4C (first column of Table 5) suggests that this simplification does not have much of an effect on other coefficients. The coefficient of interest here is on the KS4 performance variable, which

Table 5. Estimated hazard ratios of transition from reporting being likely to applying to reporting being unlikely to apply by interaction of socio-economic status and new information on attainment at age 16.

	M4C	M5C
Age 16	0.95 (-1.19)	0.95 (-1.16)
Age 17	0.92 (-1.66)*	0.97 (-0.48)
SES Z-score	0.71 (-11.36)***	0.72 (-11.06)***
KS4 Z-score (after results)	0.67 (-7.88)***	0.60 (-9.19)***
SES * KS4		0.79 (-4.03)***
<i>N</i>	9247	9247
<i>Variables</i>		
Age	√	√
SES index Z-score	√	√
Demographics & school	√	√
Prior attainment	√	√
Age 16 attainment	√	√
Age 16 attainment and SES interaction		√

Notes: Reporting hazard ratios. Adjusted using LSYPE-provided Wave 2 survey design and non-response weights. Student's *t*-test statistics, based on standard errors clustered by individual's school, reported in parentheses. Estimated risks are relative to base category of Age 15.

unsurprisingly shows that better results at age 16 are associated with a reduced probability of moving from reporting 'likely to apply' to reporting 'unlikely to apply'.

Results from M5C then provide evidence on differential responsiveness of young people to age 16 exam results. A significant estimate reported in the interaction row of Table 5 suggests that young people's SES background affects how likely they are to adjust their expectations downwards when faced with a similar set of KS4 results. The left pane of Figure 5 plots the change in the predicted probability of transition from 'likely to unlikely' associated with an increase in age 16 performance of one standard deviation for a 'high SES' individual and a 'low SES' individual. While the levels are different, both probabilities decrease by a fairly similar amount (21% to 16% for the 'low SES' individual; 14% to 7% for the 'high SES' individual) despite the large change in scores.

We turn now to consider the transition from 'unlikely to likely' (Table 6). It may well be the case that this relationship is quite different from that explaining transitions from 'likely to unlikely'. We should recall that initial selection on unobservable factors into either 'likely' or 'unlikely' could also drive some of the differences.

The unconditional relationship between young people's age and the probability that they make a transition from 'unlikely to likely' (in M0) does not seem to exhibit any of these differences: as with the opposite transition, as individuals get older they appear to become less likely to switch, albeit more dramatically by age 17.

There is a large unconditional SES (M1) gradient in young people's chances of making a transition from 'unlikely' to 'likely'. In this case, young people from more advantaged backgrounds have a higher probability: individuals from the most advantaged quintile group have more than 2.5 times the probability of making a transition as their counterparts in the least advantaged fifth. This is a large difference, although not as large as the difference between these groups in the probability of moving from 'likely to unlikely', where the

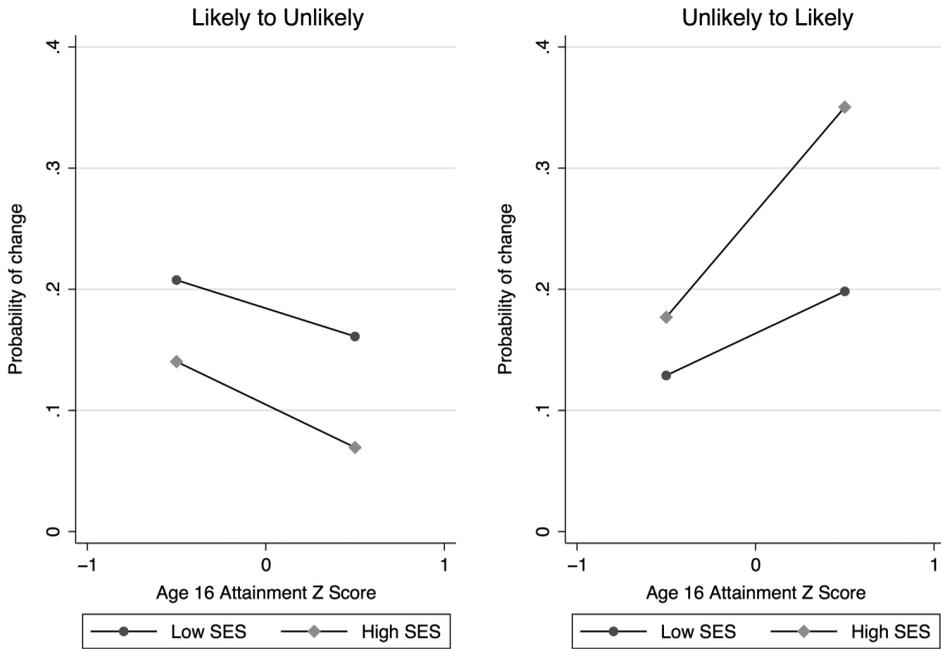


Figure 5. Predicted probability of change in expectations in response to age 16 attainment by SES. Notes: Predicted probabilities calculated from regression models MSC for the transition from likely to unlikely (left pane) and from unlikely to likely (right pane). Adjusted using LSYPE-provided Wave 2 survey design and non-response weights. All other covariates in the model are held constant at their mean values.

Table 6. Estimated hazard ratios of transition from reporting being unlikely to apply to reporting being likely to apply by quintiles of socioeconomic status.

	M0	M1	M2	M3	M4
Age 16	0.88 (-2.28)**	0.88 (-2.30)**	0.90 (-1.80)*	0.91 (-1.72)*	0.90 (-1.86)*
Age 17	0.63 (-7.90)***	0.63 (-8.13)***	0.63 (-7.90)***	0.64 (-7.61)***	0.76 (-4.48)***
SES Q1 (Low)		0.76 (-3.80)***	0.70 (-4.28)***	0.79 (-2.78)***	0.81 (-2.57)**
SES Q2		0.89 (-1.75)*	0.88 (-1.83)*	0.91 (-1.43)	0.91 (-1.38)
SES Q4		1.29 (3.42)***	1.25 (3.05)***	1.16 (2.00)**	1.15 (1.87)*
SES Q5 (High)		1.94 (7.76)***	1.92 (7.68)***	1.71 (6.25)***	1.67 (5.99)***
Significance of SES ($P > F $)		0.00	0.00	0.00	0.00
N	5330	5330	5330	5330	5330
Variables					
Age	√	√	√	√	√
SES quintile dummies		√	√	√	√
Demographics & school			√	√	√
Prior attainment				√	√
Age 16 attainment					√

Notes: Reporting hazard ratios. $P > |F|$ shows p -value from joint significance test of the hypothesis that exponentiated coefficients on all SES group dummies in the underlying conditional log-log regression model are equal to 1. Adjusted using LSYPE-provided Wave 2 survey design and non-response weights. Student's t -test statistics, based on standard errors clustered by individual's school, reported in parentheses. Estimated risks are relative to base categories of Age 15 and SES quintile group 3.

unconditional hazard ratio was greater than four. However, as with the inverse transition, will this apparent influence of SES be reduced when further covariates are added?

The additional covariates in M2 do nothing to reduce the association between SES and probability of this transition. Nevertheless, coefficients on some of the variables added at this point suggest strong relationships with the probability of transition: in particular young people from ethnic minorities and females are much more likely to switch to being 'likely to apply'.

Controlling for prior attainment does much more to explain the SES influence on young people's probability of transition from 'unlikely to likely', particularly at the more advantaged end of the distribution. Nevertheless, a large SES gradient remains, with individuals in the top SES quintile group having more than twice the hazard of moving from 'unlikely' to 'likely' as peers in the bottom group. The most advantaged fifth of the sample remain outliers from the rest of the distribution: their probability of transition is almost 50% higher than in the group below.

In contrast to the 'likely to unlikely' transition, even controlling for prior attainment does not fully explain the role of age here. While the coefficient on age 16 becomes only significant at the 10% level, the coefficient on age 17 remains highly significant. Perhaps, while it's never too late to decide against making an application to university, it can get too late for individuals to start thinking that they will. If they have not been planning to apply to university, young people will not have taken actions necessary to make a strong application. It could be that is closer to a duration effect than an age effect but is picked up by the age coefficients due to the absence of duration parameters: it may be less likely to be present for young people who only spend a single period reporting being 'unlikely to apply', for example.

In summary, as with the transition from 'likely to unlikely', there remains a large, statistically significant relationship between young people's socioeconomic advantage and the probability that they move into thinking they are 'likely to apply'.

Again, the question arises of whether young people from less advantaged backgrounds are responding differently to new information about their academic attainment. Specifically, in this case, the hypothesis that may partially explain the growing inequality in expectations is that individuals from lower SES backgrounds are less responsive to just as promising new information at age 16 as peers with similar prior academic attainment from more advantaged homes. As with the transition from 'likely to unlikely', at this point models using a continuous measure of SES are reported.

Indeed, the results (Table 7) do suggest differential sensitivity to new information on academic performance may be important in explaining the observed changes in expectations as there is a statistically significant hazard ratio on the interaction term. We again illustrate this using predicted probabilities from the model, reported in the right pane of Figure 5. As in the main models, the probability of making a transition increases with higher attainment at age 16. However, this increase is much larger for a notional high SES individual (from 18% to 35%) than a low SES individual (from 13% to 20%). This suggests that a more advantaged individual with the same prior academic performance as a less advantaged peer is more likely to respond to better academic results at age 16 by raising their expectations of applying to university from 'unlikely' to 'likely'.

Table 7. Estimated hazard ratios of transition from reporting being unlikely to apply to reporting being likely to apply by interaction of socio-economic status and new information on attainment at age 16.

	M4C	M5C
Age 16	0.90 (-1.83)*	0.90 (-1.84)*
Age 17	0.76 (-4.48)***	0.74 (-4.77)***
SES Z-score	1.28 (7.38)***	1.29 (7.40)***
KS4 Z-score (after results)	1.74 (8.37)***	1.88 (8.26)***
SES * KS4		1.18 (2.13)**
<i>N</i>	5,330	5,330
<i>Variables</i>		
Age	√	√
SES index Z-score	√	√
Demographics & school	√	√
Prior attainment	√	√
Age 16 attainment	√	√
Age 16 attainment and SES interaction		√

Notes: Reporting hazard ratios. Adjusted using LSYPE-provided Wave 2 survey design and non-response weights. Student's *t*-test statistics, based on standard errors clustered by individual's school, reported in parentheses. Estimated risks are relative to base category of Age 15.

7. Conclusions

This paper has investigated how young people's expectations of applying to university change between age 14 and age 17, just before individuals start making applications. This is a period when a substantial proportion of young people alter their expectations of applying to university. This work highlights that this change is not just from being 'likely to apply' to being 'unlikely to apply', but rather runs in both directions.

While young people across the socioeconomic status distribution start their adolescence with high expectations of applying to university, those from less advantaged backgrounds are much more likely to revise their expectations downwards and much less likely to raise their expectations during this period. This relationship persists even once many other factors correlated with SES and, perhaps most notably, young people's prior academic attainment have been controlled for. The least advantaged fifth of young people have more than twice the probability of switching from reporting being 'likely to apply' to reporting being 'unlikely to apply' as the most advantaged fifth, conditional on prior attainment. Conversely, the most advantaged fifth of young people have more than twice the probability of changing from reporting being 'unlikely to apply' to reporting being 'likely to apply' as the most advantaged fifth, again conditional on prior attainment.

The findings suggest that part of the socioeconomic status gap in university applications has roots during this period. When combined with work suggesting it is possible to affect young people's expectations of applying to university (Sanders et al., 2013), this suggests that it is not too late to target policies, both to maintain and to raise educational expectations, at bright individuals from less advantaged backgrounds during this period of their lives. However, of the two, raising expectations of applying to university may be less effective than maintaining positive expectations (Cummings et al., 2012), especially as it becomes more difficult as individuals get older.

There is also evidence that young people from differing SES backgrounds react differently to new information on their academic attainment at age 16. This differential is also asymmetric, helping to explain the growth in inequality of expectations: more advantaged young people are significantly more responsive to improved results in raising their expectations. After these exam results is a difficult point in time to reach young people, as many move between educational institutions or leave full-time education altogether. However, it may be the case that providing fresh guidance in the light of the results is very important in ensuring young people's educational expectations are appropriate.

Acknowledgements

This forms part of the Nuffield Foundation project 'Higher Education Funding and Access: exploring common beliefs' (grant number EDU/39084). The project was based at UCL Institute of Education and the Institute for Fiscal Studies. It was directed by Claire Crawford, Lorraine Dearden, John Micklewright, and Anna Vignoles. Thanks to my supervisors Lorraine Dearden and John Micklewright for their support, also to Alex Bryson and Richard Dorsett for helpful comments. I gratefully acknowledge my PhD studentship funded by UCL IoE as part of the ESRC NCRM ADMIN node.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This work was supported by the Nuffield Foundation [grant number EDU/39084].

Notes on contributor

Jake Anders is a Senior Research Associate in Quantitative Educational Research in the Department of Learning and Leadership, UCL Institute of Education, University College London. His research focuses on the causes and consequences of educational inequality and the evaluation of policies and programmes aiming to reducing it.

ORCID

Jake Anders  <http://orcid.org/0000-0003-0930-2884>

References

- Alcott, B. (2013). Predicting departure from British education: Identifying those most at risk through discrete time hazard modelling. *Widening Participation and Lifelong Learning*, 15(4), 46–64.
- Allison, P. D. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology*, 13, 61–98.
- Anders, J. (2012a). The link between household income, university applications and university attendance. *Fiscal Studies*, 33, 185–210.
- Anders, J. (2012b). *Using the Longitudinal Study of Young People in England (LSYPE)/'Next Steps' for research into higher education access* (DoQSS Working Paper 12-13). London: Institute of Education, University of London.

- Anders, J. & Micklewright, J. (2015). Teenagers' expectations of applying to university: How do they change? *Education Sciences*, 5, 281–305.
- Baker, W., Sammons, P., Siraj-Blatchford, I., Sylva, K., Melhuish, E. C., & Taggart, B. (2014). Aspirations, education and inequality in England: Insights from the effective provision of pre-school, primary and secondary education project. *Oxford Review of Education*, 40, 525–542.
- Bane, M. J. & Ellwood, D. T. (1986). Slipping into and out of poverty: The dynamics of spells. *The Journal of Human Resources*, 21(4), 1–23.
- Breen, R., & Yaish, M. (2006). Testing the Breen-Goldthorpe model of educational decision-making. In S. L. Morgan, D. B. Grusky, & G. S. Fields (Eds.), *Mobility and inequality: Frontiers of research from sociology and economics* (pp. 232–258). Redwood City, Stanford University Press.
- Cameron, S. V., & Heckman, J. J. (1999). Can tuition policy combat rising wage inequality? In M. H. Koster (Ed.), *Financing college tuition: Government policies and educational priorities* (pp. 76–124). Washington, DC: The AEI Press.
- Chowdry, H., Crawford, C., Dearden, L., Goodman, A., & Vignoles, A. (2013). Widening participation in higher education: Analysis using linked administrative data. *Journal of the Royal Statistical Society: Series A*, 176, 431–457.
- Chowdry, H., Crawford, C., & Goodman, A. (2010). Outcomes in the secondary school years: Evidence from the Longitudinal Study of Young People in England. In A. Goodman & P. Gregg (Eds.), *Poorer children's educational attainment: How important are attitudes and behaviour?* (pp. 34–43). London: Joseph Rowntree Foundation.
- Chowdry, H., Crawford, C., & Goodman, A. (2011). The role of attitudes and behaviours in explaining socio-economic differences in attainment at age 16. *Longitudinal and Life Course Studies*, 2, 59–76.
- Cummings, C., Laing, K., Law, J., McLaughlin, J., Papps, I., Todd, L., & Woolner, P. (2012). *Can changing aspirations and attitudes impact on educational attainment? A review of interventions*. JRF report, Joseph Rowntree Foundation.
- Fumagalli, L. (2012). *Great expectations. Channels and barriers to university education* (ISER Working Paper). Institute of Social and Economic Research, University of Essex.
- Ganzach, Y. (2000). Parents' education, cognitive ability, educational expectations and educational attainment: Interactive effects. *British Journal of Educational Psychology*, 70, 419–441.
- Goldthorpe, J. H., & McKnight, A. (2004). *The economic basis of social class*. (CASE paper 80). London: Centre for Analysis of Social Exclusion, London School of Economics.
- Gorard, S. (2012). Querying the causal role of attitudes in educational attainment. *ISRN Education*, 2012, Article ID 501589.
- Gottfredson, L. S. (2002). Gottfredson's theory of circumscription, compromise, and self-creation. In D. Brown (Ed.), *Career choice and development* (pp. 85–148). San Francisco, CA: Jossey-Bass.
- Gutman, L. M., & Akerman, R. (2008). *Determinants of aspirations* (Research Report 27). London: Centre for Research on the Wider Benefits of Learning, Institute of Education, University of London.
- Hauser, R. M. (1994). Measuring socioeconomic status in studies of child development. *Child Development*, 65, 1541–1545.
- Hsiao, C. (1986). *Analysis of panel data*. Cambridge: Cambridge University Press.
- Jacob, B. A. & Wilder, T. (2010). *Educational expectations and attainment* (NBER Working Paper 15683). Cambridge, MA: National Bureau of Economic Research.
- Jenkins, S. P. (1995). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics*, 57, 129–138.
- Jenkins, S. P. (2004). *Survival analysis* (Unpublished manuscript). Institute for Social and Economic Research, University of Essex.
- Jenkins, S. P., & Schluter, C. (2002). *The effect of family income during childhood on later-life attainment: Evidence from Germany* (IZA Discussion Paper 604). Institute for the Study of Labor.
- Kao, G. & Tienda, M. (1998). Educational aspirations of minority youth. *American Journal of Education*, 106, 349–384.
- Khoo, S. T., & Ainley, J. (2005). *Attitudes, intentions and participation* (Longitudinal Surveys of Australian Youth Research Report 41). Australian Council for Educational Research.
- Kolenikov, S. & Angeles, G. (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth*, 55, 128–165.

- Morgan, S. L. (1998). Adolescent educational expectations: Rationalized, fantasized, or both? *Rationality and Society*, 10, 131–162.
- Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44, 443–460.
- Rampino, T., & Taylor, M. (2013). *Gender differences in educational aspirations and attitudes* (ISER Working Paper 2013-15). Institute for Social and Economic Research, University of Essex.
- Reynolds, J. R. & Pemberton, J. (2001). Rising college expectations among youth in the United States: A comparison of the 1979 and 1997 NLSY. *The Journal of Human Resources*, 36, 703–726.
- Rose, D., & Pevalin, D. J. (2001). *The National Statistics Socio-economic Classification: Unifying official and sociological approaches to the conceptualisation and measurement of social class* (ISER Working Paper 2001-4). Institute for Social and Economic Research, University of Essex.
- Sanders, M., Kristal, A., Sabri, F., & Tupper, A. (2013). *Aspiration and inspiration—A pilot study of mentoring in schools* (CMPO Working Paper 13/314). Bristol: Centre for Market and Public Organisation, University of Bristol.
- StataCorp. (2013). *Stata survival analysis and epidemiological tables reference manual: Release 13*. College Station, TX: Stata Press.
- Strand, S. (2011). The limits of social class in explaining ethnic gaps in educational attainment. *British Educational Research Journal*, 37, 197–229.