Migration transitions to higher educational institutions: Statistical modelling of the United Kingdom student record data

Neil Bailey*

Department of Social Statistics and Demography, University of Southampton

Jakub Bijak

Department of Social Statistics and Demography, University of Southampton

Sylke Schnepf

Unit of Econometrics and Applied Statistics, European Commission DG Joint Research Centre

Paper prepared for presentation at the European Population Conference
Budapest, Hungary, June 25th – 28th 2014

Abstract:

The number of students participating at higher education institutions (HEIs) in the United Kingdom over the past 15 years have rose sharply from 1.7million in the academic year 1995/96 to 2.6million in 2010/11. The latest figures indicate that the Higher Education Initial Participation Rate (HEIPR) for the 2011/12 academic year for English domiciled students was at a record high of 49%, this indicates that just under half of all 17 year olds that lived in England at the start of the 2011/12 academic year will participate in higher education by age thirty given the current age specific participation rates. Surprisingly, given the importance of higher education very little work has been conducted on the migratory patterns of students attending institutes of higher education in the UK. With the use of the Student Record Dataset of the Higher Education Statistics Agency - which contains detailed information on every student recorded as attending an institute of higher education in the UK - this paper uses a series of statistical techniques to gain an in-depth understanding of how student migration transitions are impacted by the student's characteristics (age, ethnicity and social background), the course they study and the institute they attended. The results indicate that there was a strong statistically significant relationship between student migration transition and social background status and ethnicity. Students from less advantaged families and non-white ethnic groups were much more likely to attend a local university or commute than their more advantaged white counterparts. The results also indicate that the distance migrated by students was impacted by the individuals ethnicity, social background and age. Distance was also affected by course studied and institution attended, with more prestigious and remote universities attracting students the furthest distances.

This paper presents work in progress and is being distributed to readers to stimulate discussion and elicit comments – Please do not cite without permission; comment, criticisms and suggestions very welcome!

*Corresponding author. Email address: Neil.Bailey@soton.ac.uk

1. Introduction

The aim of this paper is to gain an in-depth understanding of how the migration transitions of people entering Higher Education Institutions (HEIs) in the United Kingdom (UK) are impacted by their personal characteristics, the course they study and the institute they attend.

The number of students attending HEIs in the UK has steadily expanded over the last half century, while over the past 15 years student numbers have risen sharply from 1.7 million in the academic year 1995/96 to 2.6 million in 2010/11 (Higher Education Statistics Agency 2012). The latest figures indicate that the Higher Education Initial Participation Rate (HEIPR)[1] for the 2011/12 academic year for English domiciled students was at a record high of 49% (BIS 2013). This indicates that almost half of all 17 year olds that lived in England at the start of the 2011/12 academic year will participate in Higher Education (HE) by the age of thirty, assuming the current age-specific participation rates.

This expansion in student numbers was driven by government policy and the major restructuring of higher education in the United Kingdom in the early 1990s, which saw the emergence of the 'post 1992' University – former polytechnics and colleges given university status through the Further and Higher Education Act (HMSO 1992; Christie 2007). These changes were implemented under the premise of building a workforce capable of sustaining the shift towards a knowledge economy and promoting economic growth. These policies saw incentives introduced to increase the number of school leavers entering higher education to 50 per cent (Munro et al. 2009), which has nearly been achieved.

Nevertheless, the expansion of HE was not primarily driven by the desire to increase the numbers of people participating in HE. It was also underpinned by the desire of government and policy makers to bring a more diverse set of non-traditional students into universities as a means to counter problems of social exclusion and poverty within the UK (Christie 2007). Non-traditional students refer to students who would not, in previous generations, have been expected to attend university. Of a particular interest was whether improvements had been made in the representation of previously under-represented groups, such as students

[1] The Higher Education Initial Participation Rate (HEIPR) is calculated individually for the four constituent countries of the United Kingdom. A description of how the measure is calculated and links to the most updated statistical releases are available at: https://www.gov.uk/government/collections/statistics-on-higher-education-initial-participation-rates

from lower socio-economic backgrounds and ethnic minority groups (Chowdry et al. 2008). The expansion of non-traditional student representation in HE remains a major policy issue within the UK today. The 'Widening Participation Policy' (HEFCE 2013) states that anyone with the ability who wants to go to university should have the chance to do so, whatever their economic or social background.

With the population in the UK progressively increasing their participating in HE, an understanding of the motivations of students who relocate for HE is becoming increasingly important. The motivations of HE students often differ widely between individual students, affecting the distance students choose to relocate and where they choose to relocate. The decision to migrate in order to attend a HEI will be influenced by many overarching factors. These factors include ethnicity, socio-economic status, parental background and educational achievement. Many previous studies have indicated that social inequality exists for many factors across the life course such as educational attainment, access to HE and future career earnings (Blanden and Gregg 2004; Blanden and Machin 2004; Blanden et al. 2010). However, despite anecdotal evidence of motivations and the relationship to the spatial patterns of mobility, there has been little attention paid to the spatial patterning of HE students attending HEIs in the UK and how these differ between certain social groupings.

Therefore, this paper aims to investigate if inequalities or dissimilarities were present in the migration patterns of student entering into HEIs in the UK by answering the following main research questions:

1. How did a student's social background, ethnicity or gender impact on the migration outcomes experienced in order to attend a HEI?
2. How did the student's characteristics, course studied or institute attended impact on the student's migration outcome?

The analysis presented in this paper was conducted by implementing three different techniques of evaluating a student's migration transition. The first technique used a logistic model to evaluate what variables impacted on probability of whether a student either migrated or stayed local to attend a HEI. The second technique used distance travelled by the student in a Tobit regression model to investigate how distance was associated with the variables within the data. The final technique used a the typology that categorised all students that attended a HEI in the UK into one of eight categories which illustrated the type of transition the student experienced in-order to attend a HEI. This typology was then used as the dependent variable in a multinomial logistic regression model to evaluate what variables impacted on a student's migration category. These three techniques were chosen and reported in this order as each technique advances the information used to understand the type of migration transition undertaken to attend a HEI. The choice of three techniques was also used to show the commonalities within the methods and the data as well as to corroborate the outcomes of three different dependent variables.

The remainder of this paper takes the following structure. An exploration of the previous research conducted in the subject area of student mobility and social inequalities will be conducted in Section 2. A description of the data used in the analysis and an argument for further in-depth investigation are put forward in Section 3. In Section 4, the results of the preliminary analysis are presented and provide the evidence that supports the need to use modelling techniques in the analysis. The methodologies of the analysis are then explained in detail in Section 5 while the results of the different methodologies will be presented in Section 6 and finally, the paper and its findings are concluded in Section 7.

2. Previous Research

The overriding themes of this pre-existing research focused on the differences and inequalities observed in HE participation and attainment, the migration and housing choices of students and the distances experienced in the HE decision process.

There has been a significant amount of previous research that focused on the inequalities of access and participation in HE in the UK, and this remains a major policy issue to this day (Department for Education and Skills 2003, 2006; HEFCE 2013). Machin and Vignoles (2004) investigated the links between HE and family background by analysing the experiences of two cohorts of individuals born in 1958 and 1970. The findings indicate that educational inequality increased between the two cohorts and that the expansion in HE during this period benefitted children from richer families rather than the most able. Blanden and Machin (2004) also investigated the links between family background and HE by studying the temporal shifts in participation and attainment across parental income groups for children going to university in the 1970s, 1980s and 1990s. Their key finding was similar to that of Machin and Vignoles (2004) in that they found that the HE expansion was not equally distributed across people from richer and poorer backgrounds.

Further research by Galindo-Rueda et al. (2004) investigated whether the socio-economic gap in HE participation had widened over time and if this gap emerged on entry to university or earlier in the education system. They did this in two ways, firstly by looking at samples of school leavers at different time periods and analysing how the likelihood of them going to university differed as a result of the socio-economic status of the student's neighbourhood. Secondly, they used more detailed individual level data, to model the determinates of participation in HE, focusing on changes in the relationship between family background and participation overt time. The main findings of the study indicated that actual growth in participation amongst poorer student had been remarkably high but the gap between the rich and poor widened during the 1990s. They did however indicate that much of the class difference in HE participation seems to reflect inequalities at earlier stages of the education system. Therefore, despite decades of policy designed to widen participation, it appears from the majority of research that social inequality within HE in the UK increased during the 1980s, 1990s and early 2000s.

The most recent and detailed research into widening participation in HE was the report conducted by the Institute for Fiscal Studies (Chowdry et al. 2008, 2010), which used a unique individual-level administrative dataset that provided information on a particular cohort of state school pupils as they progressed through the education system. The report found that students from materially deprived backgrounds were much less likely to participate in HE at age 18 or 19 than students from less deprived backgrounds.

In short, the majority of previous research has indicated there were significant differences in participation rates in HE as a result of an individual's socio-economic background. The current study aims to build on these previous studies by using three different methodological techniques and the use of population data to see if these patterns are mirrored in the most up-to-date data available.

The substantial evidence of differences in participation outcomes as a result of socio-economic status is undoubted. There is, however, a further point to be considered. Previous research has also indicated that there are substantial differences in participation rates in HE across the different ethnic groups (Modood and Shiner 1994; Dearing 1997; Tomlinson 2001; Khambhaita and Bhopal 2013). Chowdry et al. (2008) found that ethnic minority students were significantly more likely to participate in HE than their White British peers, while Ball et al. (2002) argued that the differences amongst ethnic minority students cannot be fully understood without reference to their social class background. Shiner and Modood (2002) and Chowdry et al. (2008) found that there were large institutional bias with regards to ethnicity and that there were large socio-economic and ethnic gaps in the likelihood of attending high status HEIs within the UK.

The current study does not provide analysis of participation over time like the previously mentioned research. However, unlike previous research, the analysis in this chapter does use population data, to provide an insight into the current situation of the student population with regards to their socio-economic, parental and ethnic background. The current study also builds on the previously mentioned research by analysing if the findings of inequalities that were visible in participation trends are also visible in the migration transitions and distance travelled to attend a HEI.

Despite the vast amount of prior research in the area of HE, a truly quantitative analysis of socio-economic, ethnic and gender differences on the specific migration transitions experienced by students is missing. The current study will build on the prior research by applying solely quantitative techniques on population data that has not been previously conducted and therefore the current study will support or critique the findings of the previous research that was conducted using differing techniques and data sources by comparing there results to the findings produced here.

3. Data

Our student migration and characteristics information come from the Higher Educations Statistics Agency (HESA) Student Record Data. The HESA Student Record is collected in respect of all students registered at a reporting HEI in the United Kingdom, which follow courses that lead to the award of a qualification or institutional credit, excluding those registered as studying wholly overseas. However, for the purpose of the analysis undertaken in this chapter it was decided that students that migrated from overseas to attend a HEI in the UK would be omitted from the study. As mentioned previously, one of the aims of the study in this chapter is to investigate how a student's background and ethnic group could have impacted on the migration decisions of students attending HEIs in the UK. Due to the fact that all international students have migrated in order to study in the UK and due to the poor information held on international students' background and ethnicity, it was decided best to omit this group from the analysis. It must also be noted that this dataset consists of 'population data' as every student in the UK is recorded and the data is not derived from a survey. After removing the International and Open University students, the remaining data consists of 1,797,492 students that were enrolled at a UK HEI in the 2011/12 academic year. Table 1 gives a brief description of the variables used in the empirical analysis while Tables 2-4 provide cross-tabulations of the three outcome variables against ethnicity, social background and gender.

The explanatory analysis shown in Tables 2-4 illustrate than unconditionally independent of any other explanatory variables, there appears to be significant differences between ethnic groups, social background and gender in all three of the migration outcomes. These preliminary results support findings from the literature that ethnicity, social background and gender play important roles in the migration decision process (Modood and Shiner 1994; Dearing 1997; Tomlinson 2001; Khambhaita and Bhopal 2013), and this therefore suggests evidence to support future multivariate analysis of these variable later on in the paper.

Table 1 – Variable Descriptions

Name	Description
Outcome Variables	
Internal Migration – Yes or No	Had the student internally migrated to attend a HEI. A internal migration was recorded if the student's term-time address was in a different Local Authority (LA) to their domicile address
Distance Migrated	How far the student had travelled to attend a HEI. The distance in kilometres between the LA centroid for the three locational variables: domicile, term-time and institution address.
Migration Type	Typology of Student Migration which categorises student movement into 4 categories; **Local Student, Commuter/Distance Learner, Internal Student Migrant and Internal Student Migrant Commuter/Distance Learner**
Explanatory Variables	
Ethnicity	- **White** includes White and Irish Traveller. - **Black** includes Black or Black British - Caribbean, Black or Black British - African, and other Black background. - **Asian** includes Asian or Asian British - Indian, Asian or Asian British - Pakistani, Asian or Asian British - Bangladeshi, Chinese, and other Asian background. - **Other** (including mixed) includes mixed - White and Black Caribbean, mixed - White and Black African, mixed - White and Asian, other mixed background, and other ethnic background. - **Unknown** includes not known and information refused.
Social Background	Social Background variable indicates how advantageous the student's background - with regards to their socio-economic status and whether their parent(s) attended higher education - was in encouraging a student towards the traditional process of migrating away from the parental home to attend a HEI. The students were grouped into five social background categories: **Most Advantaged, Advantaged, Less Advantaged, Least Advantaged and Unknown**.
Age	The age of the student was recorded as at 31st August in the reporting period, therefore in this study the age of the student on 31st August 2011.
Gender	**Male**, **Female** and **Indeterminate**
Number of years in HE	This field indicates the number of years that the student had been enrolled in a HE course or programme leading to the student's qualification aim
Level of Study	Level of study was taken from the course aim of the student and classifies a student as either **Undergraduate** or **Postgraduate**
Institution Attended	Three variables to define the universities status: **Pre1992** – If the institution was founded before the 1992 expansion (HMSO 1992) **Russell Group** – Was the institution part of the Russell Group which represents 24 leading UK HEIs. **Top30 -** This variable indicated whether the institution attended by the student was in the top 30 of the Complete University Guide 2014 University League Table
Course Studied	Students degree course was classified into 7 categories: **Medicine, Science or Engineering, Agriculture or Veterinary, Social or Human Sciences, Business or Law, Humanities and Combined**

Table 2 - Internal Migration Outcomes by Students Ethnicity

2011/12	Ethnicity					Total
	White	Black	Asian	Other	Unknown	
All Students (%)	79.1	6.0	9.0	4.0	1.9	100
Internal Migration - Yes (%)	47.4	30.1	32.0	45.5	27.9	44.6
Internal Migration - No (%)	52.6	69.9	68.0	54.5	72.1	55.4
Mean Total Distance (km)	96.8	59.7	61.9	85.1	79.4	90.6
Local Students (%)	10.3	13.2	16.4	11.8	15.7	11.2
Commuter/Distance Learner (%)	42.3	56.7	51.7	42.7	56.4	44.3
Internal Student Migrant (%)	37.7	19.8	23.6	34.0	19.1	34.9
Migrant Commuter/Distance Learner attended local HEI (%)	0.5	0.6	0.4	0.6	0.8	0.5
Internal Migrant Commuter/Distance Learner (%)	9.2	9.7	8	10.9	8	9.1

Table 3 - Migration Outcomes by Students Background

2011/12	Social Background					Total
	Most Advantaged	Advantaged	Less Advantaged	Least Advantaged	Unknown	
All Students (%)	17.5	17.8	22.8	20.6	21.3	100
Internal Migration - Yes (%)	71.1	57.8	45.9	32.6	21.8	44.6
Internal Migration - No (%)	28.9	42.2	54.1	67.4	78.2	55.4
Mean Total Distance (km)	126.7	102.9	89.4	68.8	73.2	90.6
Local Students (%)	5.4	8.2	11.0	13.9	15.9	11.2
Commuter/Distance Learner (%)	23.4	34.0	43.1	53.5	62.3	44.3
Internal Student Migrant (%)	60.4	47.6	35.1	24.1	13.4	34.9
Migrant Commuter/Distance Learner attended local HEI (%)	0.3	0.4	0.6	0.6	0.8	0.5
Internal Migrant Commuter/Distance Learner (%)	10.5	9.8	10.2	7.9	7.6	9.1

Table 4 - Migration Outcomes by Gender

2011/12	Gender		Total
	Male	Female	
All Students (%)	42.5	57.5	100
Internal Migration - Yes (%)	49.6	40.8	44.6
Internal Migration - No (%)	50.4	59.2	55.4
Mean Total Distance (km)	99.7	83.9	90.6
Local Students (%)	10.4	11.8	11.2
Commuter/Distance Learner (%)	40	47.4	44.3
Internal Student Migrant (%)	39.5	31.4	34.9
Migrant Commuter/Distance Learner attended local HEI (%)	0.5	0.6	0.5
Internal Migrant Commuter/Distance Learner (%)	9.6	8.8	9.1

Total Population Size – 1,797,492

Source: Higher Education Statistics Agency (2013)

4. Methodology

On the basis of the arguments in sections 2 and 3, we can hypothesise that the decision to migrate in order to attend a HEI in the UK is associated with the student's individual characteristics, the course they study and the institution they attend. Therefore it was decided to analyse these associations further by investigating how these variables would impact on the migration outcomes when the explanatory variable were analysed simultaneously.

The three outcome variables used in the analysis are quite different in their format and as a result they required different methods to analyse the student migration outcomes against the explanatory variables simultaneously. This section will go through the methods used in turn and the results of these methodologies are presented in Section 5.

The first of the migration outcome variables that was analysed was the binary outcome that depicted whether a student had migrated or not, where a value of 0 (a failure) was recorded for no migration and a value of 1 (a success) if a migration had occurred.

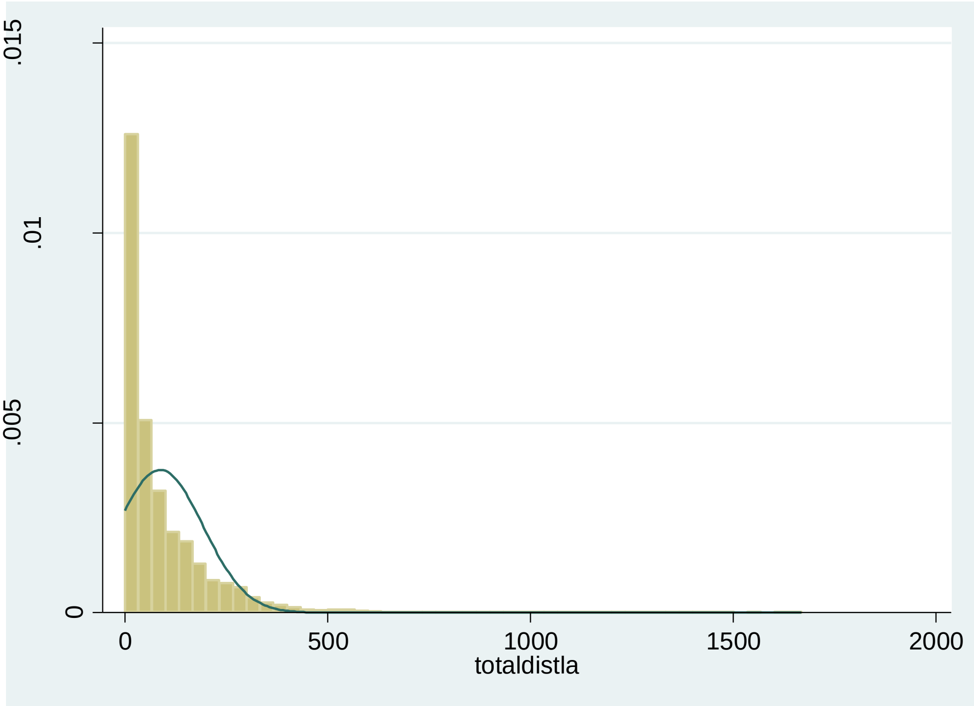
For the binary response variable of migration Y and the multiple explanatory variables x_p the rearranged logistic regression model to calculated the predicted probabilities for $\pi(x) = P(Y = 1)$, the predicted probability of a student making a migration to attend a HEI, at values $x = (x_1, \dots, x_p)$ of p predictors was (Agresti 2013, p18):

$$\pi(x) = \frac{e^{(\alpha+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p)}}{1+e^{(\alpha+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p)}} \qquad (4.1)$$

Where the parameter β_j refers to the effect of x_j on the log odds that $Y = 1$, adjusting for the other x_k . Therefore, $e^{(\beta_j)}$ was the multiplicative effect on the odds of a student migration of a 1-unit increase in x_j when the other variables levels of x_k were left constant.

The second of the migration outcome variables that was analysed was the variable of distance travelled by the student to attend a HEI where the outcome variable referred to a numerical value of distance on a continuous scale measured in kilometres. For outcome variables on a continuous scale the common method to estimate linear relationships is multiple linear regression. However, one of the criterions for linear regression is that the outcome variable is normally distributed. The distribution of the distance outcome variable is shown in Figure 1, and was clearly not normally distributed. Therefore, a different methodology needed to be used or a transformation of the outcome variable performed.

Figure 1 - Histogram to show the distribution of the outcome variable Total Distance (km)



Source: Higher Education Statistics Agency (2013)

Note: 'totaldistla' refers to the total distance measured in km for each individual student in the 2011/12 academic year. Non-UK Domiciled students were removed.

The Tobit model is a statistical model originally proposed by Tobin (1958) that was designed to estimate the linear relationship between variables when the outcome variable has a number of its values clustered at a limiting values, usually zero (McDonald and Moffitt 1980). The Tobit technique uses all observations, both those at the limit and those above it, to estimate a regression line, and as a result of the model taking into account the clustering in the outcome variable, it is to be preferred, in general, over alternative techniques that estimate the regression line only with observations above the limit (McDonald and Moffitt 1980).

As shown in Figure 1, the outcome variable of total distance was a perfect example of where the outcome variable was clustered around zero. This clustering around zero was a result of those students that did not make a migration or commute in order to attend a HEI and were therefore classified as a local student. Due to the structure of the outcome variable in this example it was decided that the Tobit model was the best methodology to analyse the impact of the explanatory variables on the distance travelled in order to attend a HEI.

In the Tobit model used here y_i referred to the total distance travelled by the student in kilometres, there were 10 explanatory variables x_p and three interaction terms. The structural equation for the Tobit model is shown in Equation 4.2 (Long 1997).

$$y_i^* = x_{1i}\beta_1 + x_{2i}\beta_2 + \cdots + x_{pi}\beta_p + \varepsilon_i \tag{4.2}$$

In Equation 4.5, $\varepsilon_i \sim N(0, \sigma^2)$. The x's are observed for all cases. y^* was a latent variable that was observed for values greater the τ and was censored for values less than or equal to τ. The observed y was defined by the measurement Equation 4.3.

$$y_i = \begin{cases} y_i^* & if \quad y_i^* > \tau \\ \tau_y & if \quad y_i^* \leq \tau \end{cases} \tag{4.3}$$

As previously mentioned, the data in this analysis were censored at zero therefore $\tau = 0$, with this in mind and combining Equations 4.5 and 4.6 the final Tobit model is shown in Equation 4.4.

$$y_i = \begin{cases} y_i^* = x_{1i}\beta_1 + x_{2i}\beta_2 + \cdots + x_{pi}\beta_p + \varepsilon_i & if \quad y_i^* > 0 \\ 0 & if \quad y_i^* \leq 0 \end{cases} \tag{4.4}$$

The third migration outcome variable that was analysed was the variable of student migration category. Multinomial regression modelling is suitable where the response variable is nominal and has three or more categories that are unordered. The basic principle of multinomial regression is the prediction of the probability of membership to each group of the outcome variable as a result of the observed explanatory variables. Therefore, for this section of the analysis predicted the probability of a student being in one of the student migration categories given their explanatory characteristics. In predicting probabilities response categories are simultaneously compared to a reference category. Multinomial regression models the log of probability ratio; the log of probability of response in one category compared to the probability of the reference category. The set-up of these models, where the outcome variable has four categories as in this analysis, are shown in Equations 4.5 (Agresti 2013):

$$\log\left(\frac{\pi_1}{\pi_4}\right) = \alpha_1 + x_1\beta_1 \quad \log\left(\frac{\pi_2}{\pi_4}\right) = \alpha_2 + x_2\beta_2 \quad \log\left(\frac{\pi_3}{\pi_4}\right) = \alpha_3 + x_3\beta_3 \qquad (4.5)$$

In Equation 4.8, π_1 is the response category 1, π_2 is the response category 2, π_3 is the response category 3 and π_4 is the response category 4 (reference category), α_i the intercept, x_i a vector of the explanatory variables and β_i the coefficients.

The reference category used in this analysis is 'Local Student' (4). Regression equations are set up for Commuter/Distance Learner (1), Internal Student Migrant (2) and Internal Migrant Commuter/Distance Learner (3). In order to ease interpretation, results from the logit equations were used to calculate the probabilities of being in a student category by transforming the equations into the predicted probabilities as shown in Equations 4.6 to 4.9 (Agresti 2013):

Probability of the reference category (local student):

$$P_4 = \frac{1}{1+e^{\alpha_1+\beta_1 x}+e^{\alpha_2+\beta_2 x}+e^{\alpha_3+\beta_3 x}} \qquad (4.6)$$

Probability of category 1 (Commuter/Distance Learner):

$$P_1 = \frac{e^{\alpha_1+\beta_1 x}}{1+e^{\alpha_1+\beta_1 x}+e^{\alpha_2+\beta_2 x}+e^{\alpha_3+\beta_3 x}} \qquad (4.7)$$

Probability of category 2 (Internal Student Migrant):

$$P_2 = \frac{e^{\alpha_2+\beta_2 x}}{1+e^{\alpha_1+\beta_1 x}+e^{\alpha_2+\beta_2 x}+e^{\alpha_3+\beta_3 x}} \qquad (4.8)$$

Probability of category 3 (Migrant Commuter/Distance Learner):

$$P_3 = \frac{e^{\alpha_3+\beta_3 x}}{1+e^{\alpha_1+\beta_1 x}+e^{\alpha_2+\beta_2 x}+e^{\alpha_3+\beta_3 x}} \qquad (4.9)$$

5. Results

5.1 Internal Migration

Out of the 1,797,492 non-international students enrolled at a HEI in the United Kingdom in the 2011/12 academic year, 44.6% migrated across a local authority boundary to do so and when considered individually, all the explanatory variables analysed appeared to have some form of association with the migration outcome.

The final logistic regression model included all the explanatory variables available to the researcher plus three interaction terms between; gender and ethnicity, gender and background and background and ethnicity. The coefficients (β), standard errors, 95% confidence intervals and the odds of a successful outcome (e^{β}) of the final logistic regression model are shown in Table 5.

It is important to note that when interpreting the figures shown in Table 5, that those variables that were involved in the significant interaction terms, their main effect terms cannot be interpreted individually since the individual main effects of interacted variables cannot be isolated. Therefore, interpretation regarding the ethnicity, background and gender variables should be in terms of their interactions in order to make appropriate conclusions.

One of the aims of this chapter was to analyse whether a student's background or ethnicity impacted on the migration outcomes experienced in order to attend a HEI. As a result, it was decided to interact these terms in the model and the outcomes were significant at the 1% level (with the exception of ethnicity unknown and Asian*High, which were insignificant at the 10% level). It was also decided to interact these two variables with gender, to see if there were any significant differences between ethnicity and background as a result of the student's gender. Again, these two new interaction terms were significant at the 1% level (with the exception of Other*Female and Unknown*Female, which were insignificant at the 10% level). It is therefore, necessary to evaluate the effects of these three variables concurrently. However it is important to state that the high significance levels again may be a product of the very large sample size and having population data and as a result the statistical inferences being made are technically referring to a super-population (Cochran 1939; Hartley and Sielken 1975; Dorfman and Valliant 2005).

Table 5 - Multiple logistic regression results of the association between student migration and the student characteristic variables

Variable (x)	Coefficient (β)	Sig.	Std. Err.	95% Confidence Interval		e^{β}
				Low	High	
Constant	0.69	***	0.01	0.66	0.71	1.99
Ethnicity						
White[a]						
Black	-0.47	***	0.02	-0.51	-0.43	0.63
Asian	-0.92	***	0.02	-0.96	-0.89	0.40
Other	-0.31	***	0.02	-0.36	-0.27	0.73
Unknown	-0.05		0.06	-0.17	0.06	0.95
Social Background						
Most Advantaged						
Advantaged	-0.26	***	0.01	-0.28	-0.24	0.77
Less Advantaged	-0.47	***	0.01	-0.49	-0.46	0.62
Least Advantaged	-0.72	***	0.01	-0.74	-0.71	0.48
Unknown	-0.95	***	0.01	-0.97	-0.93	0.39
Gender						
Male[a]						
Female	-0.14	***	0.01	-0.16	-0.13	0.87
Course Studied						
Medicine[a]						
Science/Engineering	0.23	***	0.01	0.21	0.24	1.25
Agriculture or Veterinary	0.65	***	0.02	0.61	0.68	1.91
Social or Human Science	0.03	***	0.01	0.01	0.04	1.03
Business or Law	0.05	***	0.01	0.04	0.07	1.06
Humanities	0.58	***	0.01	0.56	0.59	1.78
Combined	-0.59	***	0.02	-0.64	-0.55	0.55
Institution Variables						
Non-Russell Group[a]						
Russell Group	0.44	***	0.01	0.43	0.45	1.55
Non Top 30[a]						
Top 30	0.50	***	0.01	0.48	0.51	1.64
Post 1992[a]						
Pre 1992	0.24	***	0.00	0.23	0.25	1.27
Age						
17 years and under	-1.04	***	0.02	-1.08	-1.00	0.35
18-20 years[a]						
21-24 years	-0.72	***	0.00	-0.73	-0.71	0.49
25-29 years	-1.55	***	0.01	-1.57	-1.54	0.21
30 years and over	-2.65	***	0.01	-2.66	-2.64	0.07
Age Unknown	-2.95	***	0.18	-3.31	-2.59	0.05
Number of Years in HE						
1[a]						
2	0.17	***	0.00	0.16	0.17	1.18
3	0.38	***	0.00	0.37	0.39	1.46
4	0.67	***	0.01	0.65	0.68	1.95
5	0.74	***	0.01	0.71	0.76	2.09
6 or more	1.05	***	0.02	1.01	1.10	2.87
Unknown	0.35	***	0.10	0.16	0.55	1.42
Level of Student						
Postgraduate[a]						
Undergraduate	-0.12	***	0.01	-0.14	-0.11	0.88
Interaction Terms						
Ethnicity*S.Background						

Variable (x)	Coefficient (β)	Sig.	Std. Err.	95% Confidence Interval Low	High	e^{β}
White*Most Advantaged[a]						
Black*Advantaged	0.06	**	0.03	0.00	0.11	1.06
Black*Less Advantaged	0.14	***	0.02	0.09	0.19	1.15
Black*Least Advantaged	0.15	***	0.03	0.10	0.20	1.16
Black*Unknown	0.34	***	0.03	0.29	0.40	1.41
Asian*Advantaged	-0.02		0.02	-0.06	0.03	0.98
Asian*Less Advantaged	-0.05	***	0.02	-0.09	-0.01	0.95
Asian*Least Advantaged	-0.07	***	0.02	-0.11	-0.03	0.93
Asian*Unknown	0.42	***	0.02	0.38	0.46	1.52
Other*Advantaged	0.12	***	0.03	0.06	0.18	1.13
Other*Less Advantaged	0.11	***	0.03	0.06	0.16	1.12
Other*Least Advantaged	0.08	***	0.03	0.03	0.14	1.09
Other*Unknown	0.32	***	0.03	0.26	0.38	1.38
Unknown*Advantaged	0.00		0.07	-0.14	0.14	1.00
Unknown*Less.Adv	-0.13	**	0.06	-0.26	-0.01	0.87
Unknown*Least.Adv	-0.18	***	0.07	-0.31	-0.05	0.84
Unknown*Unknown	-0.20	***	0.06	-0.31	-0.08	0.82
S.Background*Gender						
V. Advantaged*Male[a]						
Advantaged*Female	-0.06	***	0.01	-0.08	-0.04	0.94
Less Advantaged*Female	-0.10	***	0.01	-0.12	-0.08	0.91
Least Advantaged*Female	-0.18	***	0.01	-0.20	-0.16	0.84
Unknown*Female	-0.13	***	0.01	-0.15	-0.11	0.88
Ethnicity*Gender						
White*Male[a]						
Black*Female	0.10	***	0.02	0.07	0.14	1.11
Asian*Female	0.09	***	0.01	0.06	0.11	1.09
Other*Female	0.01		0.02	-0.02	0.05	1.01
Unknown*Female	0.03		0.03	-0.03	0.09	1.03

Source: Higher Education Statistics Agency (2013)
Note: * denotes significance at the 1% level, ** 5% level, * 10% level, [a] reference category**

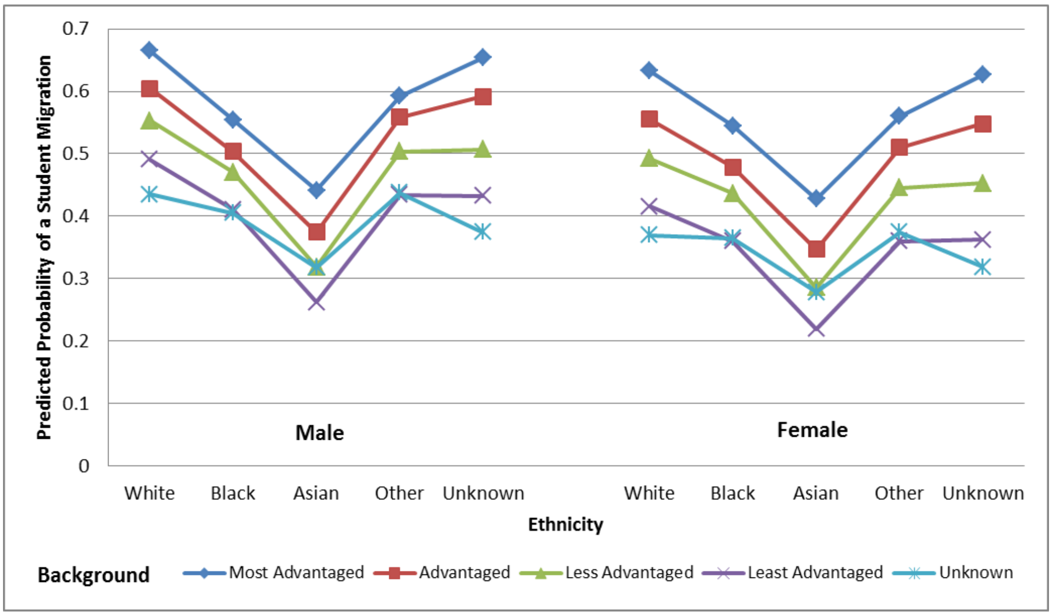
The predicted probabilities, as calculated by substituting the coefficients in Table 5 into the regression equation 4.1, of a student migration by ethnicity, background and gender are shown in Figure 2 and Figure 3. Figure 2 and Figure 3 graph the same predicted probabilities; however Figure 2 uses background as the focus variable illustrated by the lines on the graph, whereas Figure 3 switches the focus to ethnicity.

When studying the patterns in the two figures, it was important to note that the lines in the graphs are not parallel which indicated that the interaction terms in the model were significant and without these interaction terms these differences would not have been visible. In terms of ethnicity, social background and gender the predicted probabilities of student migration varied significantly. While fixing the remaining variables in the model at the reference category, it was clear to see that a White, most advantaged social background male was the type of student that had the highest predicted probability (0.633) of making a student migration. In contrast, an Asian, least advantaged social background female had the lowest predicted probability (0.219) of making a student migration.

Equally important was to consider the impact of the student's social background on the probability of student migration. Those in the most advantaged group had the highest predicted probabilities than all other social backgrounds and this was the case for all ethnicities and both genders (Figure 2). Those students in the advantaged background group had the second highest predicted probabilities and the less advantaged group the third highest and again this was the case for all ethnicities and both genders. However, there was variation in these trends with regards to the least advantaged and unknown social background groups, as the order of the predicted probabilities changed as a result of ethnicity and gender. The least advantaged social background group had the second lowest predicted probabilities for White, Black and Unknown ethnicity males but the lowest probability for Asian and Other. For females the trend was similar, however, Black, least advantaged social background females also fell below the unknown social background line unlike their male counterparts.

These patterns show that the social background of the student appeared to play a significant role in the likelihood of a student migrating in order to study. This was suggested in the preliminary analysis but has been confirmed here when all other variables were considered together. These patterns tend to support previous research regarding the tradition in UK HE for students to migrate away from their parental home in order to study at a HEI especially for those from more traditional backgrounds (Patiniotis and Holdsworth 2005). There appears to be an interaction between Asian students and the least advantaged social background group, as these students were highly unlikely to migrate in order to study.

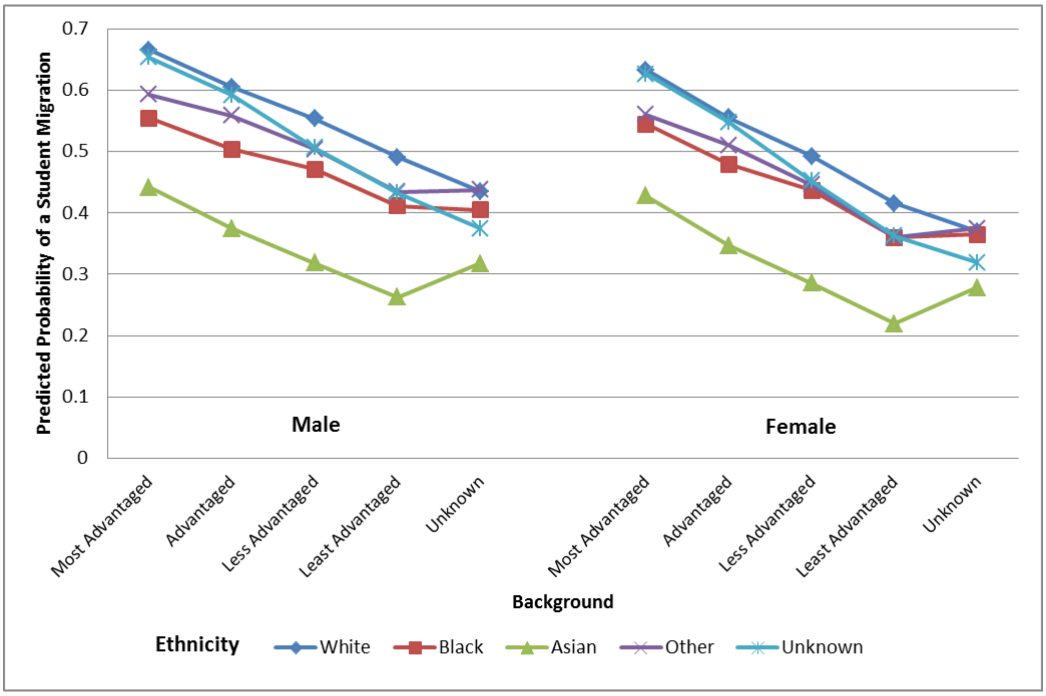
Figure 2 - Variations in the predicted probabilities of student migration by gender and ethnicity, for different social backgrounds.



Source: Higher Education Statistics Agency (2013)

Note: The predicted probabilities assume the remaining variables were set to the reference category therefore; the student studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University, was 18-20 years old and was a first year postgraduate student.

Figure 3 - Variations in the predicted probabilities of student migration by gender and social background, for different ethnic groups.



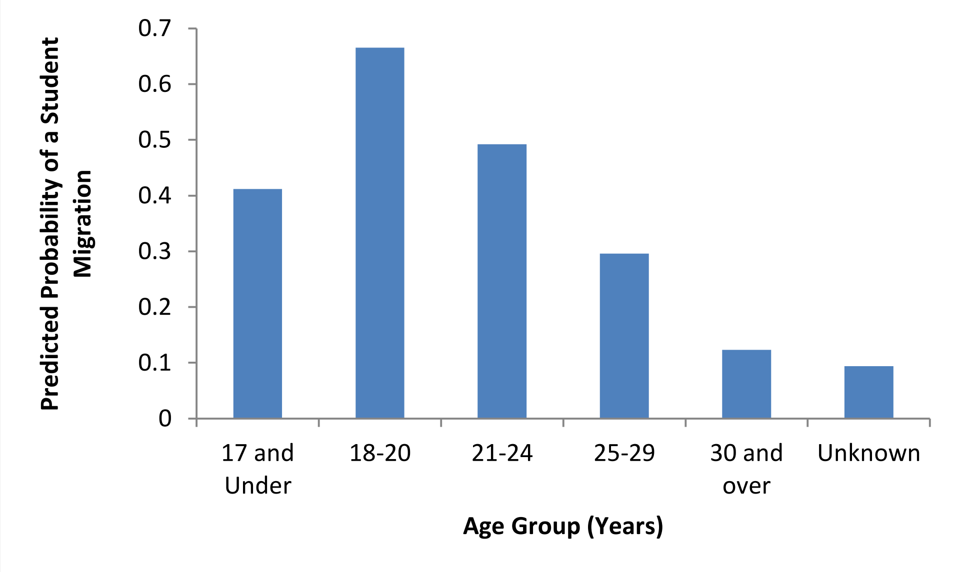
Source: Higher Education Statistics Agency (2013)
Note: The predicted probabilities assume the remaining variables were set to the reference category therefore; the student studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University, was 18-20 years old and was a first year postgraduate student.

Focusing on the differences between ethnicities (Figure 3) it was clear that those students from the Asian ethnicity group were the least likely to migrate in order to attend a HEI irrespective of social background or gender, although Asian females were even less likely to migrate than their male counterparts. In contrast, White students had the highest predicted probabilities irrespective of social background, while White males were again more likely to migrate than their White female counterparts. This trend supports the findings of Khambhaita and Bhopal (2013) who found that Asian female students were much more likely than White students to stay living in the parental/guardian home during the first year at university. In the current study a student staying in the parental/guardian home during the first year at university would be recorded as not making a migration and with the predicted probability of not migrating being highest for Asian females from the least advantaged social background it appears the findings here mirror those reported by Khambhaita and Bhopal (2013).

Another aim of this chapter was to evaluate how the student's characteristics, course studied or institute attended impacted on the student's migration outcomes. The remainder of this section analyses the probability of a student migration according to their characteristics (other than ethnicity, social background and gender), course studied and institution attended. These remaining variables were not involved in any interaction terms and therefore their main effects could be interpreted individually. All the remaining control variables in the model were significant at the 1% level.

When considering the impact on the probability of migration as a result of a student's age, clear differences were visible. In comparison to the 18-20 years (the reference category) age group, all of the remaining age groups had odds of a success less than one, suggesting that students in these age groups were less likely to migrate than those aged 18-2- years. The declines in the predicted probabilities by age are clearly illustrated in Figure 4. The highest probabilities of migration at the 18-20 years group could be a result of students in this age group following a more traditional route into HE. For example, moving straight through the education system into HE and quite often moving out of the parental home to do so. Those students in the older age groups may be less likely to make a migration to study because they have already done so for work reasons and decide to go into HE at a later age and as a result were more heavily tied to the area they resided in prior to entering HE.

Figure 4 - Predicted probabilities of student migration by age group.

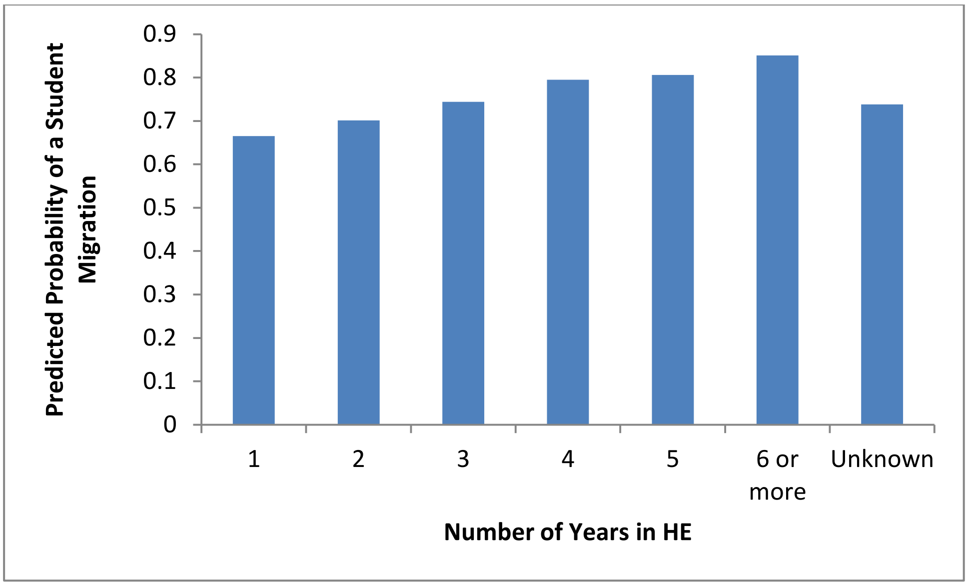


Source: Higher Education Statistics Agency (2013)
Note: The predicted probabilities assume the remaining variables were set to the reference category therefore; the student was male, white, Very High Background, studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University and was a first year postgraduate student.

Another interesting trend was the probability of migration by year of student as shown in Figure 5. The students least likely to migrate were those in their first of study, while the probability of migration increased sequentially with each year of study. This may be caused be influenced by large number of ethnic minority students that tend to remain in the parental home, especially in the first year of study (Khambhaita and Bhopal 2013). The subsequent increase in the probability of migration as people progress through university may also be a result of students deciding to migrate after the initial decision to remain in the parental home and commute.

When considering the level of study of the student the data showed that while holding all other variables constant, postgraduate students were more likely to migrate than undergraduates. However, although the variable was significant at the 1% level the difference in the odds between the two groups was quite small.

Figure 5 - Predicted probabilities of student migration by Number of Years in HE.



Source: Higher Education Statistics Agency (2013)
Note: The predicted propabilities assume the remaining variables were set to the reference category therefore; the student was male, white, Very High Background, studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University and was a first year postgraduate student.

With regards to the institution variables, all three were significant and all three variables showed that if the student attended a top30, Russell Group or Pre1992 University, then they were more likely to migrate to do so compared to if they did not.

Finally, there were some observed differences between courses studied and the predicted probability of migration. Those students that studied Agricultural or Veterinary courses and those studying Humanities had much greater probabilities of migration than those who studied Medicine (the reference category). Social or Human Science and Business or Law were very similar to the reference category, while students that studied Combined degrees were much less likely to migrate in order to attend their HEIs.

5.2 Distance

Every student in the HESA student record data used in this analysis had a total distance travelled recorded for them. This distance referred to the number of kilometres the student migrated or commuted in order to attend a HEI. These distances vary from 0km for those students who studied at their local institution to a maximum of 1723km, while the mean distance travelled across all the non-international students was 90.6km.

In the current sub-section, a Tobit Model was used to analyse the effects of ten explanatory variables and three interaction terms on the predicted total distance travelled for a student to attend a HEI, and the results of this model are shown in Table 6.

As in the logistic regression methodology the ethnicity, background and gender variables were interacted with each other in order to answer one of the main research questions of the study. The predicted total distance travelled by a student by ethnicity, background and gender are shown in Figure 6 and Figure 7. Figure 6 and Figure 7 graph the same predicted distances; however Figure 6 uses background as the focus variable illustrated by the lines on the graph, whereas Figure 7 switches the focus to ethnicity. Again it is important to note that the lines in the two graphs were not parallel which indicated the significance of including the interaction terms.

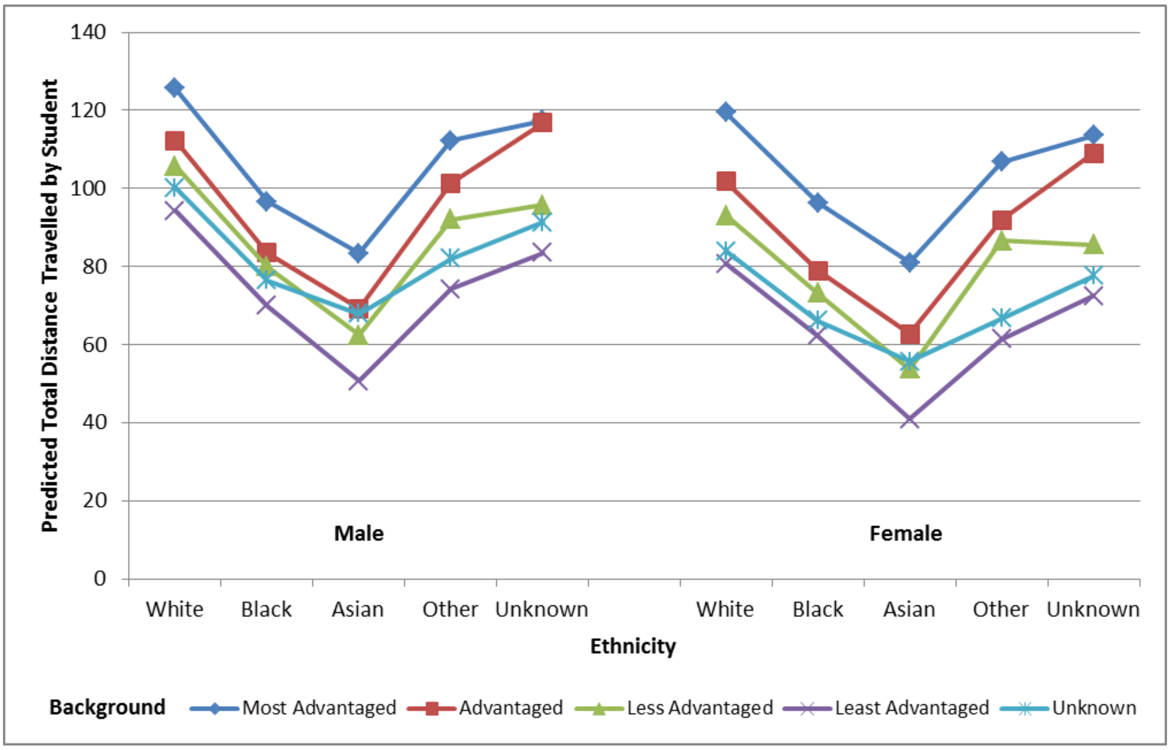
Table 6 - Tobit Model results of the association between total distance travelled to attend a HEI and the students' characteristic variables

Variable (x)	Coefficient (β)	Sig.	Std. Err.	95% Confidence Interval	
				Low	High
Constant	125.74	***	0.52	124.72	126.76
Ethnicity					
White[a]					
Black	-29.03	***	1.05	-31.08	-26.98
Asian	-42.39	***	0.89	-44.13	-40.65
Other	-13.43	***	1.08	-15.56	-11.31
Unknown	-8.30	***	2.58	-13.36	-3.24
Social Background					
Most Advantaged[a]					
Advantaged	-13.50	***	0.44	-14.36	-12.64
Less Advantaged	-19.92	***	0.42	-20.75	-19.08
Least Advantaged	-31.48	***	0.45	-32.37	-30.59
Unknown	-25.62	***	0.46	-26.53	-24.72
Gender					
Male[a]					
Female	-6.32	***	0.41	-7.12	-5.52
Course Studied					
Medicine[a]					
Science/Engineering	5.21	***	0.29	4.65	5.77
Agriculture or Veterinary	58.49	***	0.79	56.95	60.04
Social or Human Science	-2.12	***	0.28	-2.68	-1.57
Business or Law	2.34	***	0.31	1.73	2.96
Humanities	18.54	***	0.29	17.98	19.11
Combined	-20.38	***	0.88	-22.10	-18.65
Institution Variables					
Non-Russell Group[a]					
Russell Group	12.67	***	0.31	12.07	13.27
Non Top 30[a]					
Top 30	19.97	***	0.31	19.37	20.57
Post 1992[a]					

Variable (x)	Coefficient (β)	Sig.	Std. Err.	95% Confidence Interval	
				Low	High
Pre 1992	9.98	***	0.23	9.52	10.43
Age					
17 years and under	-25.38	***	1.08	-27.49	-23.26
18-20 years[a]					
21-24 years	-13.41	***	0.24	-13.89	-12.93
25-29 years	-39.70	***	0.34	-40.37	-39.04
30 years and over	-45.99	***	0.28	-46.53	-45.44
Age Unknown	-20.92	***	4.71	-30.15	-11.68
Number of Years in HE					
1[a]					
2	-0.64	***	0.21	-1.05	-0.23
3	1.34	***	0.23	0.88	1.80
4	10.78	***	0.37	10.05	11.50
5	8.23	***	0.68	6.90	9.55
6 or more	12.97	***	0.97	11.06	14.87
Unknown	-24.51	***	4.53	-33.39	-15.64
Level of Student					
Postgraduate[a]					
Undergraduate	-15.11	***	0.27	-15.64	-14.58
Interaction Terms					
Ethnicity*Social Background					
White*Most Advantaged[a]					
Black*Advantaged	0.29		1.38	-2.41	2.99
Black*Less Advantaged	3.20	***	1.19	0.87	5.52
Black*Least Advantaged	4.75	***	1.20	2.40	7.10
Black*Unknown	5.33	***	1.22	2.93	7.73
Asian*Advantaged	-0.65		1.14	-2.88	1.57
Asian*Less Advantaged	-0.96		1.03	-2.98	1.07
Asian*Least Advantaged	-1.26		1.00	-3.22	0.71
Asian*Unknown	10.25	***	1.07	8.15	12.35
Other*Advantaged	2.59	***	1.43	-0.21	5.40
Other*Less Advantaged	-0.25		1.30	-2.81	2.30
Other*Least Advantaged	-6.65	***	1.37	-9.33	-3.97
Other*Unknown	-4.56	***	1.39	-7.29	-1.84
Unknown*Advantaged	12.91	***	3.19	6.65	19.17
Unknown*Less Advantaged	-1.81		3.00	-7.69	4.07
Unknown*Least Advantaged	-2.36		3.06	-8.37	3.64
Unknown*Unknown	-0.48		2.65	-5.67	4.72
Social Background*Gender					
Most Advantaged*Male[a]					
Advantaged*Female	-4.06	***	0.57	-5.17	-2.95
Less Advantaged*Female	-6.36	***	0.54	-7.41	-5.31
Least Advantaged*Female	-7.27	***	0.56	-8.36	-6.18
Unknown*Female	-9.89	***	0.55	-10.97	-8.81
Ethnicity*Gender					
White*Male[a]					
Black*Female	5.89	***	0.73	4.45	7.32
Asian*Female	3.93	***	0.60	2.76	5.11
Other*Female	0.89		0.87	-0.81	2.59
Unknown*Female	2.48	**	1.24	0.04	4.92

Source: Higher Education Statistics Agency (2013)
Note: *** denotes significance at the 1% level, ** 5% level, * 10% level, [a] reference category

Figure 6 - Variations in the predicted total distance travelled by a student by gender and ethnicity, for different social backgrounds.



Source: Higher Education Statistics Agency (2013)

Note: The predicted distances assume the remaining variables were set to the reference category therefore; the student studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University, was 18-20 years old and was a first year postgraduate student.

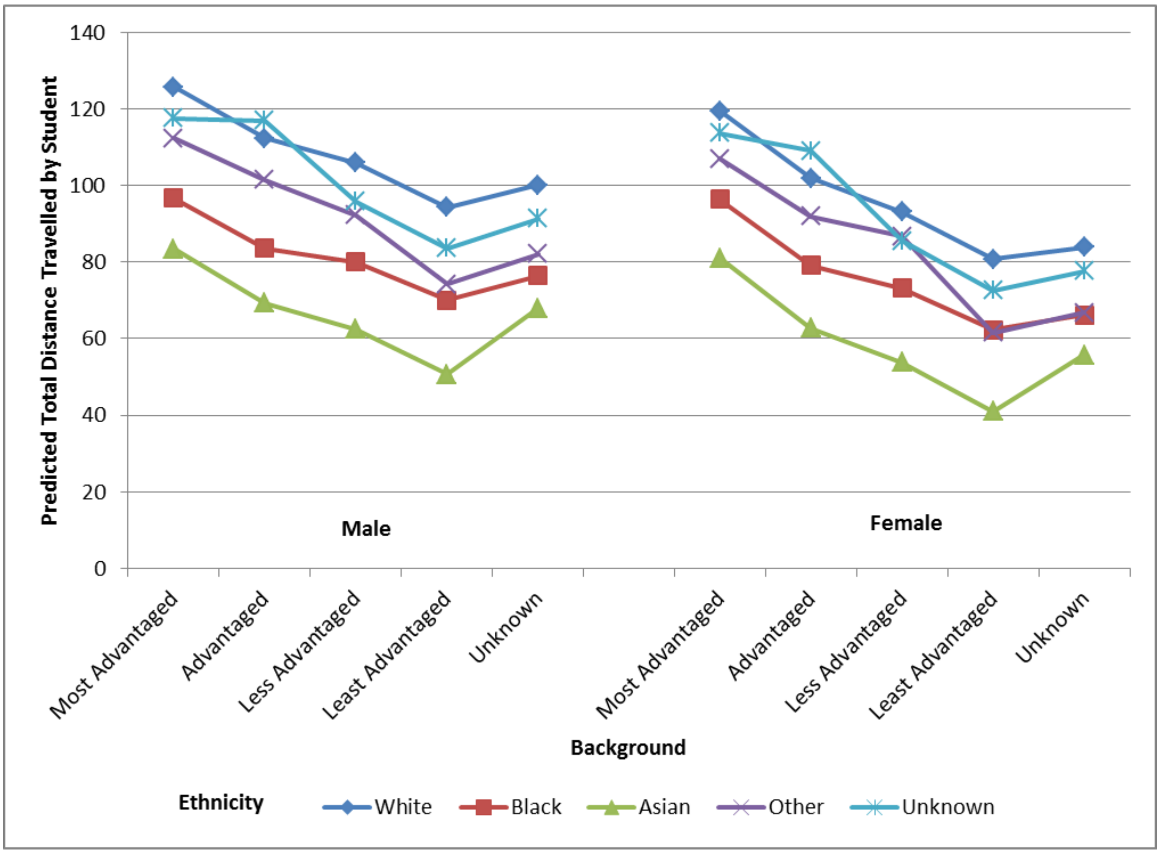
The overall results seemed very similar to those produced when investigating the probability of a student migrating. Those students predicted to travel the largest distances were the same group that had the highest probability of migration, White, most advantaged social background Males, who were predicted on average to travel 125.7 km. Similarly, the group predicted to travel the shortest distances were Asian, least advantaged social background Females, who were predicted on average to travel 41.0 km. This represented a significant difference in the predicted distances as a result of differing ethnicity and social background.

When considering the impact of a student's social background on the distance they travelled to attend a HEI (Figure 6), the patterns were again very similar to those observed in the probability of migration results Note the similarities between Figures 2 and 6. The most advantaged social background group were predicated to travel the furthest distances and those distances were greater for males than females across all ethnicities. While, the least advantaged social background group, were found to be the group predicted to travel the shortest distances across all ethnicities. The one noticeable difference in social background between the probability to migration and the predicted distance was for the social background unknown variable. The predicted distances for students with social background unknown were much higher than expected and much higher when compared to their predicted probability of migration. This may have been a result of the total distance variable measuring commuting distances as well as migration distances whereas the probability of migration did not take into account commuting in its measurement.

Shifting the focus to ethnicity (Figure 7), again the trends were similar with regards to predicted distance as they were with predicted probability of migration, again note the similarities between Figures 3 and 7. Those students of Asian ethnicity were predicated to travel the shortest distances and the difference between the Asian and Other ethnic groups were quite large. Again, female students were predicted to travel shorter distances than their male counterparts. The White group were predicted to travel the furthest distances across all social backgrounds with the exception of the advantaged social background group. This was one difference between the two outcome variables. For the advantaged social background group the unknown ethnicity group were predicted to travel the furthest distances. For all social background groups the Black ethnic group were predicted to travel the second shortest distances, however in the least advantaged social background group the Other ethnicity value was very similar to the value for the Black ethnic group.

The remaining variables in the model were not involved in any interaction terms and as a result their main effects can be interpreted individually. All the remaining control variables in the model are again significant at the 1% level (refer to previous comments with regards to significance).

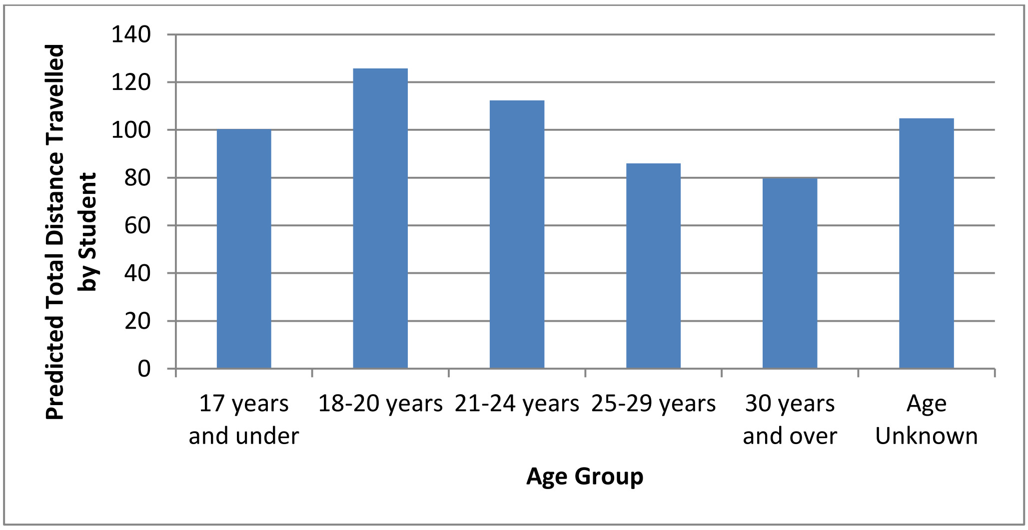
Figure 7 - Variations in the predicted total distance travelled by gender and social background, for different ethnic groups.



Source: Higher Education Statistics Agency (2013)

Note: The predicted distances assume the remaining variables were set to the reference category therefore; the student studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University, was 18-20 years old and was a first year postgraduate student.

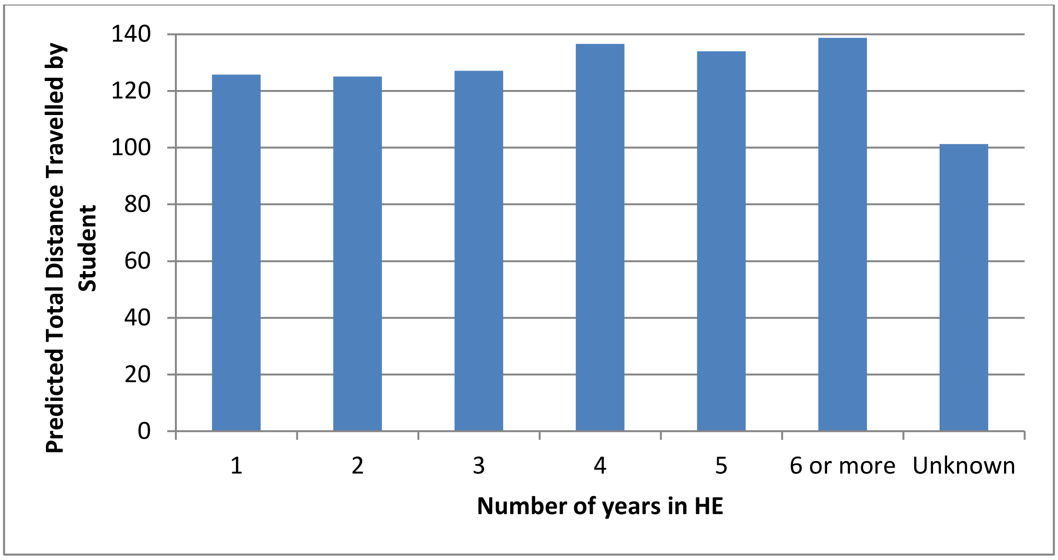
Figure 8 - Predicted total distance travelled by student by Age Group



Source: Higher Education Statistics Agency (2013)

Note: The predicted probabilities assume the remaining variables were set to the reference category therefore; the student was male, white, Very High Background, studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University and was a first year postgraduate student.

Figure 9 - Predicted total distance travelled by student by Year of Student.



Source: Higher Education Statistics Agency (2013)

Note: The predicted probabilities assume the remaining variables were set to the reference category therefore; the student was male, white, Very High Background, studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University and was a first year postgraduate student.

The direction and strength of trends between the remaining variables and the predicted distance travelled again mirror those found in the logistic regression model results. The effect of age (Figure 8) on distance was the same as observed on the probability of migration (Figure 4), the average predicted distance travelled by a student declined as age increased. While the effect of the number of years in HE (Figure 9) was also the same direction as observed on the probability of migration (Figure 5) however, the impact on the average distance travelled did not change significantly between the year groups.

With regards to the institution variables, all three were significant and all three variables showed that if the student attended a Top30, Russell Group or Pre1992 University, then they were more likely to travel further distances compared to if they did not. However, the differences in average distance only ranged from 10 to 12.7km for the three variables therefore, the differences in the average distances as of result of HEI attended were not that substantial.

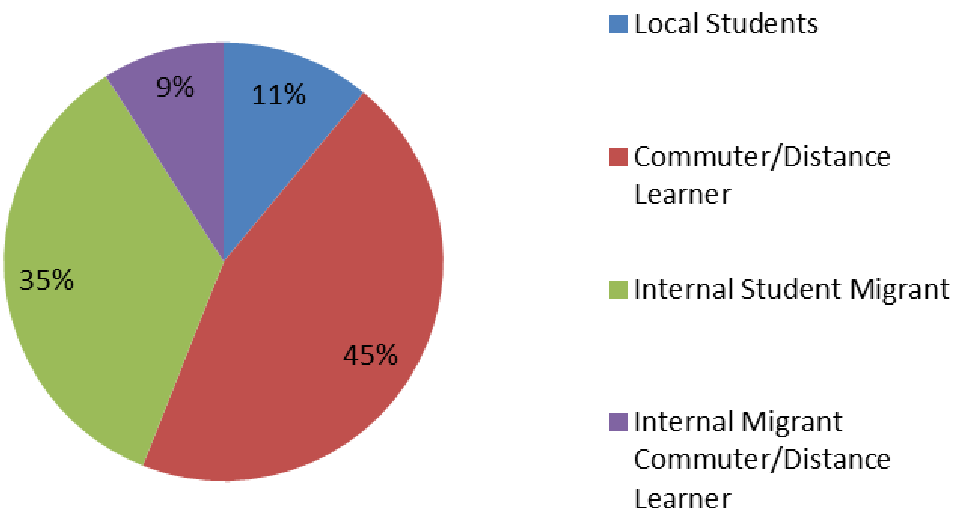
Finally, there were some observed differences between courses studied and the predicted distance travelled by the student. Those studying Agricultural or Veterinary course on were predicted on average to travel the furthest, with these students predicted to travel on average 58.5 km further than those who studied Medicine (the reference category). Those students studying Humanities had a very high probability of migration but the results in the distance outcome suggest that these migrations may actually be over relatively short distances. On average a Humanities student was predicted to travel on 18.5 km more than the reference category.

Social or Human Science and Business or Law were very similar to the reference category, while students studying Combined degrees were predicted to travel the shortest distances, but this difference of 20.4 km to the reference category does not represent a very large difference. Therefore, with the exception of Agricultural and Veterinary courses, the course studied does not affect the average predicted distance travelled by a student by very large distances.

5.3 Migrant Type

A typology of student migration categorised every student registered at a UK HEI into one of four categories that depicted the type of migration the student experienced in order to attend a HEI. The distributions of the student sub-population between these four categories are illustrated in Figure 10.

Figure 10 - Breakdown of student population by student migration category



Source: Higher Education Statistics Agency (2013)

The analysis in this sub-section develops on the binary analysis of migration as the four categories of student internal migration are directly linked to the categorisation of migration; yes or no. Local student and commuter/distance learners did not migrate and were classified as not migrating in the previous analysis; while in contrast, the internal student migrant and internal migrant commute/distance learner categories did migrate and were recorded in the migration yes category.

In the current sub-section, a multinomial logistic regression model was used to analyse the probability of a student being in one of the four student migration categories and the coefficients of the multinomial logistic model are shown in Table 7. The final chosen multinomial logistic regression model included no interaction terms and as a result the main effects of all the variables in the model can be interpreted individually.

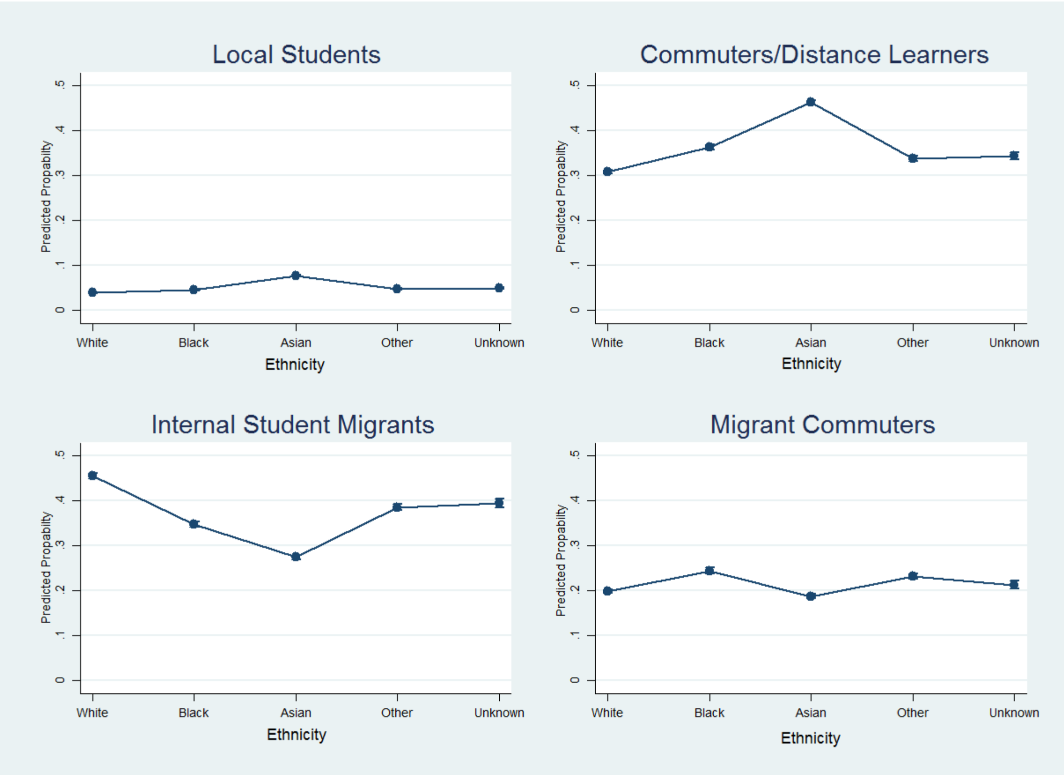
Table 7 - Multinomial logistic regression results of the association between student migration category and student characteristic variables

Variable (x)	Commuter Coefficient (β)	Commuter Sig.	Migrant Coefficient (β)	Migrant Sig.	Migrant Commuter Coefficient (β)	Migrant Commuter Sig.
Constant	2.09	***	2.33	***	1.63	***
Ethnicity						
White[a]						
Black	0.03	***	-0.41	***	0.07	***
Asian	-0.27	***	-1.18	***	-0.73	***
Other	-0.10	***	-0.36	***	-0.04	**
Unknown	-0.11	***	-0.36	***	-0.15	***
Social Background						
Most Advantaged[a]						
Advantaged	-0.04	***	-0.31	***	-0.31	***
Less Advantaged	-0.09	***	-0.62	***	-0.45	***
Least Advantaged	-0.12	***	-0.95	***	-0.78	***
Unknown	-0.06	***	-1.06	***	-0.81	***
Gender						
Male[a]						
Female	-0.02	***	-0.25	***	-0.21	***
Course Studied						
Medicine[a]						
Science/Engineering	-0.28	***	0.16	***	-0.34	***
Agr/Vet	0.54	***	0.90	***	1.48	***
Social/Human	-0.25	***	0.00		-0.56	***
Business/Law	-0.29	***	-0.05	***	-0.43	***
Humanities	-0.31	***	0.44	***	0.18	***
Combined	-0.66	***	-1.00	***	-1.22	***
Institution Variables						
Non-Russell Group[a]						
Russell Group	-0.43	***	0.21	***	-0.17	***
Non Top 30[a]						
Top 30	0.14	***	0.69	***	0.42	***
Post 1992[a]						
Pre 1992	-0.06	***	0.27	***	-0.01	
Age						
17 years and under	-0.22	***	-1.23	***	-1.40	***
18-20 years[a]						
21-24 years	-0.18	***	-0.93	***	-0.51	***

Variable (x)	Commuter Coefficient (β)	Sig.	Migrant Coefficient (β)	Sig.	Migrant Commuter Coefficient (β)	Sig.
25-29 years	-0.39	***	-2.32	***	-0.98	***
30 years and over	-0.13	***	-3.65	***	-1.65	***
Age Unknown	0.38	***	-3.69	***	-1.55	***
Number of years in HE						
1[a]						
2	-0.04	***	0.09	***	0.22	***
3	-0.10	***	0.25	***	0.36	***
4	-0.12	***	0.54	***	0.52	***
5	-0.14	***	0.51	***	0.73	***
6 or more	-0.17	***	0.63	***	1.10	***
Unknown	-0.34	***	-0.40	***	0.88	***
Level of Student						
Postgraduate[a]						
Undergraduate	-0.15	***	0.01		-0.61	***

Source: Higher Education Statistics Agency (2013)
Note: [a] denoted reference category of the independent variables, the reference category for the response variable was Local Student; *** denotes significance at the 1% level, ** 5% level, * 10% level

Figure 11 - Predicted probabilities of student migration category by ethnic group



Source: Higher Education Statistics Agency (2013)
Note: The predicted probabilities assume the remaining variables were set to the reference category therefore; the student was from the most advantged social background group, Male, studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University, was 18-20 years old and was a first year postgraduate student. The error bars indicate the 95% Confidence Interval for each predicted probability.

The impact of the student's ethnic group on the predicted probabilities of being in each of the student migration categories when holding the other variables constant at the reference category is illustrated in Figure 11. There were some clear differences in the probabilities as a result of the students' ethnicity and these differences again illustrate similar findings to the previous two outcome variables.

The predicted probability of being a local student was significantly higher for Asian students that any other of the ethnic groups. This finding echo's the results found using the previously mentioned methodologies and supports the findings of previous research that suggests that Asian students were more likely to remain in the parental home during HE (Khambhaita and Bhopal 2013) than the other ethnic groups. Asian students were also the most likely to be a commuter/distance learner, which again supports the idea that Asian students tend to remain in the parental home while studying at a HEI. Out of the four student migration categories the predicted probability of being a local student was clearly the lowest. In this chapter the whole focus has been on the LA level of geography. As discussed previously, the geographical area that constitutes a LA is relatively small and as a result very few students remain within the boundaries of a LA to study at a HEI. Therefore as a result of the MAUP (Bell et al. 2002), the results could change in there magnitude if the level of geographical analysis was altered.

The group most similar to the local student group are the commuter/distance learner group. These groups are similar in the fact that both groups do not undertake a migration to study at a HEI but they differ in the fact that the commuter/distance learner students attend a HEI in a different LA. When comparing the predicted probability plots for these two groups the trends seem similar in there patterns, but the overall probability of being a commuter/distance learner was much higher than it was to be a local student. This was not surprising when the distribution of student numbers across the four categories is considered. With regards to ethnic differences in the commuter/distance learner group it was the Asian students who again had the highest predicted probability of being a commuter/distance learner. In contrast, the White ethnic group had the lowest predicted probabilities of being a commuter/distance learner as well as being a local student. Again, these results support the findings found using the previous two outcome variables.

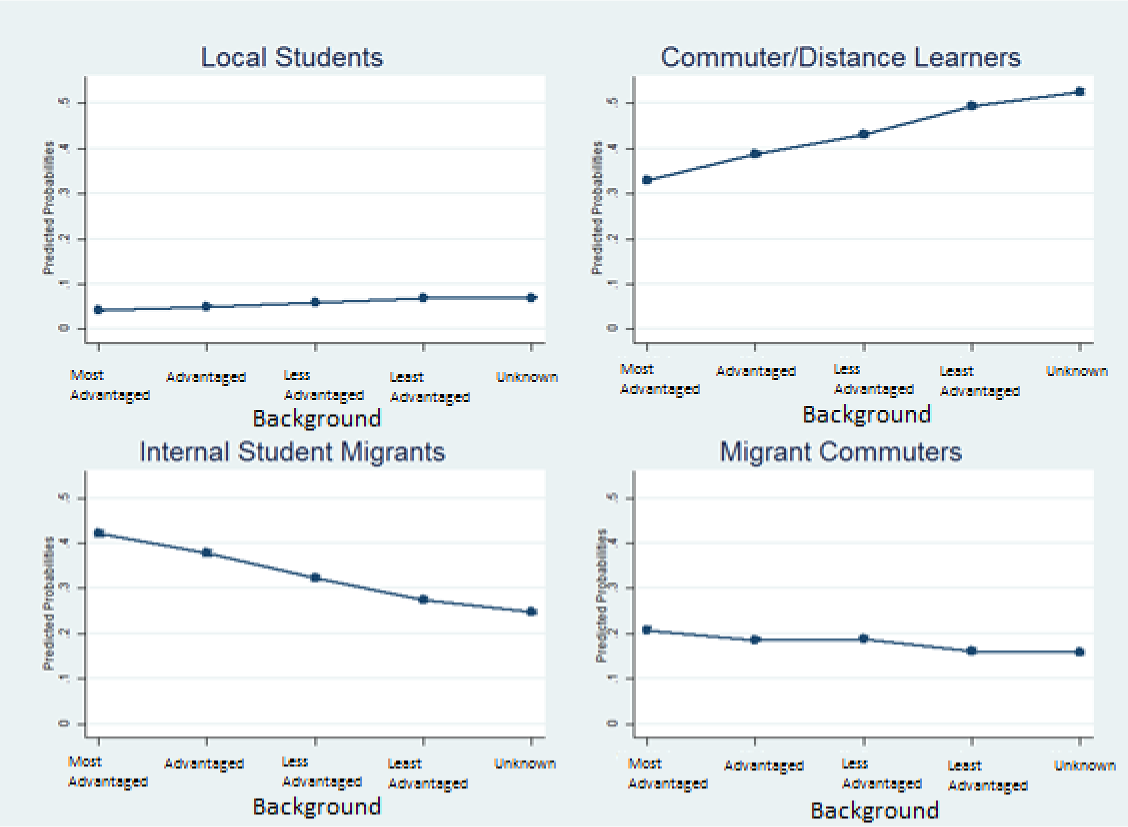
In contrast to the local and commuter/distance learner students are those students who made an internal migration. The White ethnic group had the highest predicted probabilities of being an internal student migrant while the Asian group had the lowest predicted probabilities. Again, these findings support those presented earlier where White students had the highest probability of making a migration in order to attend a HEI.

However, when you consider the migrant commuter a group a new pattern arises that has not been highlighted with the previous outcome variables. With regards to migrant commuters, the Black ethnic group had the highest predicted probability of all the ethnic groups, closely followed by the Other ethnic group. These findings were slightly surprising; however, some of this may be explain by certain factors. A large percentage of these migrant commuter students were studying in London HEIs and they are recorded as crossing LA boundaries but still remain with the area of Greater London itself. This can also be linked back to the large Black ethnic group population in London that is not found in other LAs across the UK (Office for National Statistics 2013).

While comparing differences in ethnicity is important, another part of the research aims of this chapter were to analyse the impact of a student's social background. Again, there were clear differences in the predicted probabilities of being in each of the student migration categories as a result of the student's social background. The predicted probabilities of the different migration categories by social background are shown in Figure 4-12.

The probability of being a commuter/distance learner had a linear trend with social background, the probability of being a commuter/distance learner increased as social background advantageousness decreased, while the opposite trend was apparent for internal student migrants. The differences between the social background groups for local students and migrant commuters were smaller in comparison although the direction of the trend was as expected. For local students, those with less advantageous social backgrounds had the highest predicted probabilities, while for migrant commuters, those from the more advantageous social backgrounds had the highest probabilities. The only finding that stood out in the analysis that did not in the other methodologies was the very high predicted probability of being a commuter/distance learner if your social background was unknown, however there is no logical explanation from the literature or any previous research to explain why this may be the case.

Figure 12 - Predicted probabilities of student migration category by social background

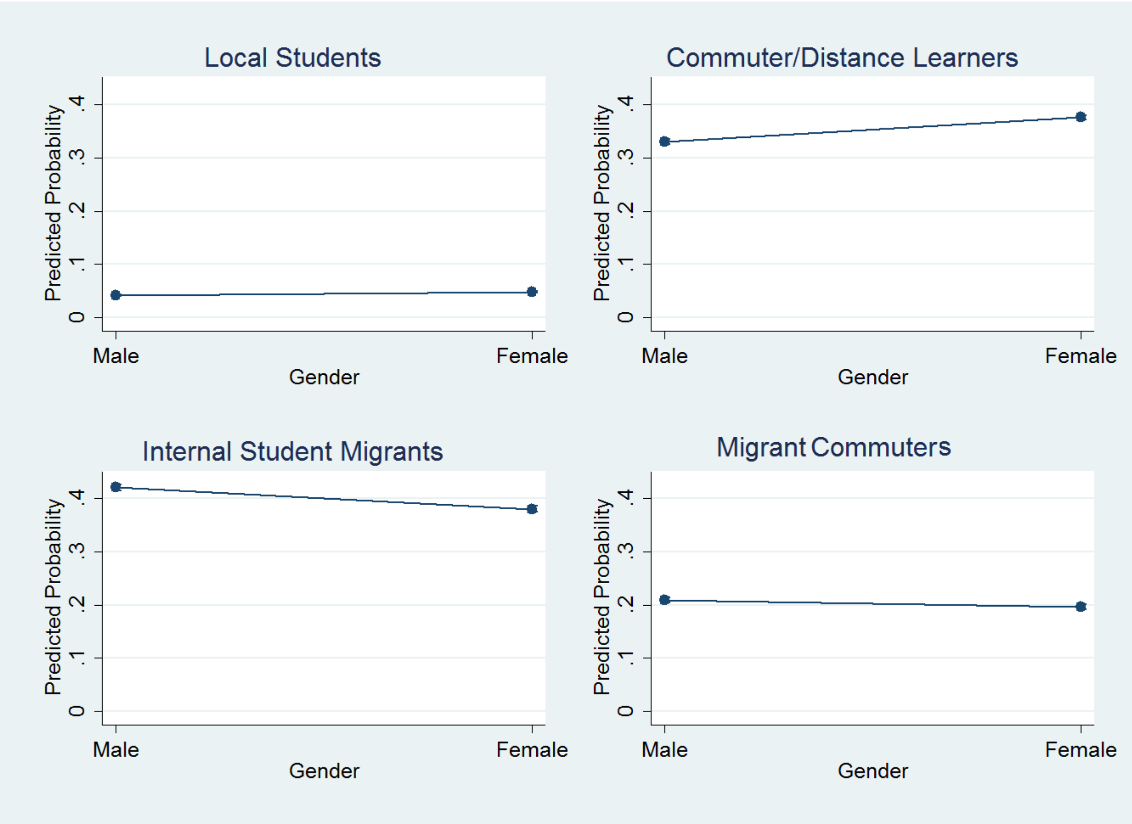


Source: Higher Education Statistics Agency (2013)
Note: The predicted probabilities assume the remaining variables were set to the reference category therefore; the student was White, Male, studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University, was 18-20 years old and was a first year postgraduate student. The error bars indicate the 95% Confidence Interval for each predicted probability.

In the previous two methodologies there were also observed differences in the migration outcomes as a result of the student's gender. The predicted probabilities of the four migration categories by gender are shown in Figure 13.

The differences in the predicted probabilities as an impact of gender are minimal, especially when compared to the differences by ethnicity and social background. Females had a higher predicted probability of being a local student than their male counterparts but the size of the difference was minimal. A similar but reversed pattern was seen for migrant commuters but again the difference between the probabilities was very small and insignificant. The differences between the genders for internal student migrants and commuter/distance learners were larger in size but again the differences were not substantial. Male were more likely than females to be internal student migrants and for commuter/distance learners the trend was reversed.

Figure 13 - Predicted probabilities of student migration category by gender

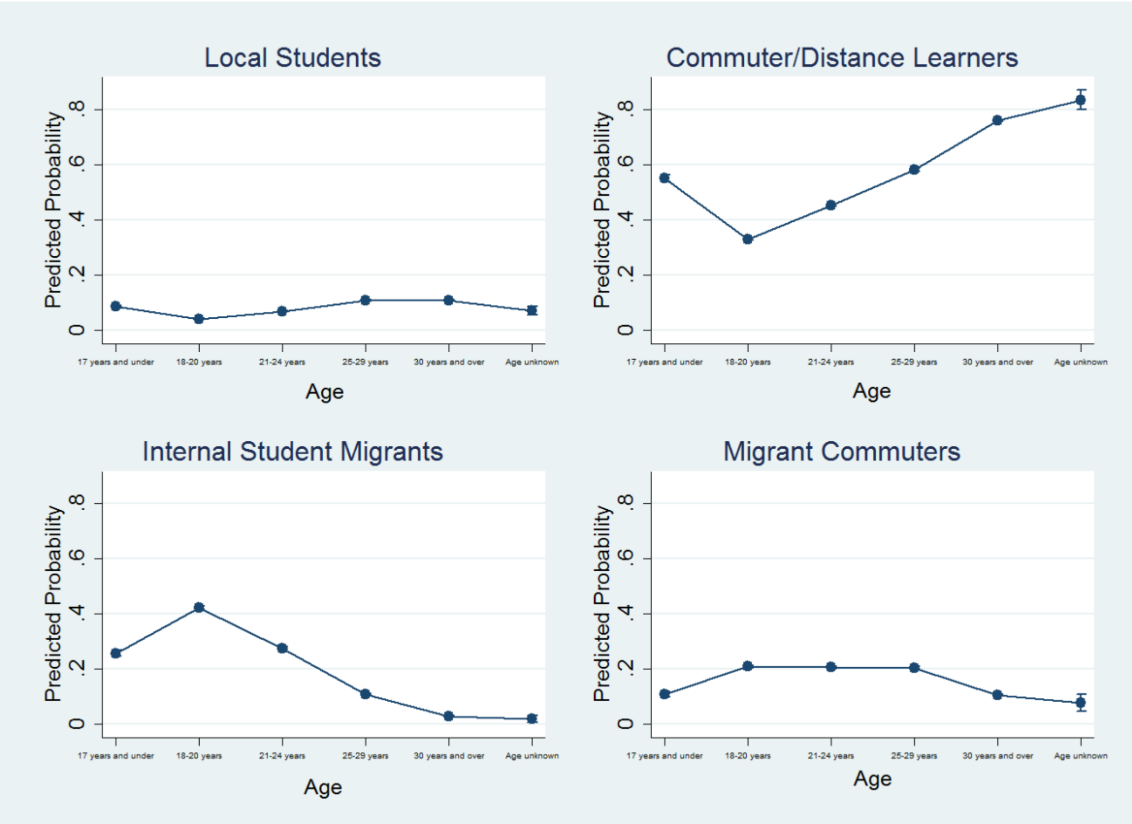


Source: Higher Education Statistics Agency (2013)
Note: The predicted probabilities assume the remaining variables were set to the reference category therefore; the student was White, from most advantaged social background, studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University, was 18-20 years old and was a first year postgraduate student. The error bars indicate the 95% Confidence Interval for each predicted probability.

There was also a clear impact of age in the previous two methodologies with the probability of migration and the predicted distance travelled declining as age increased. The predicted probabilities of being categorised in one of the four student categories by age of the student are illustrated in Figure 14.

The only major differences by age were in the internal student migrant and commuter/distance learner groups. Those aged 18-20 years were the most likely to be internal student migrants and least likely to be local students. The predicted probability of being a commuter/distance learner increased with age. These findings were not surprising. Previous empirical evidence has shown how migration intensity is interlinked with age (Wilson 2010) and as age increases so does the likelihood that the individual will have stronger ties to an area such as owning a house or having children and therefore the probability of that individual migrating to study would decrease.

Figure 14 - Predicted probabilities of student migration category by age group



Source: Higher Education Statistics Agency (2013)
Note: The predicted probabilities assume the remaining variables were set to the reference category therefore; the student was White, Male, from most advantaged social background, studied Medicine, attended a Non-Russell Group, Non Top 30 and Post 1992 University and was a first year postgraduate student. The error bars indicate the 95% Confidence Interval for each predicted probability.

Overall, the results from the multinomial analysis provided the same conclusive findings as the previous methodologies. The overall relationship between ethnicity and social background were very similar irrespective of the methodology or outcome variable used, with those from White and most advantageous social backgrounds most likely to be an internal student migrant, while those from the least advantageous social backgrounds and non-White ethnic groups were much more likely to be local students or commuters/distance learners.

6. Conclusions

This paper has provided an in-depth analysis of how student migration transitions of people entering higher educational institutions (HEIs) in the United Kingdom were impacted by the student's characteristics, the course they studied and the institute they attended. It is recognised that migration transitions of people into higher education is of great policy interest to higher educational institutions as well as government and non-government organisations as a result of the impact students have on the locations that they reside. These migration transitions of individual student's vary as a result of many contributing factors. However, this analysis aimed to investigate if any over-arching themes or trends were apparent in the data and if any patterns in the migration transition experienced by a student have been impacted by their social background, ethnic group or gender.

The aim of this paper was to answer two main research questions; how did a student's social background, ethnicity or gender impact on the migration outcomes experienced in order to attend a HEI and how did the student's characteristics, course studied or institute attended impact on the student's migration outcome?

The preliminary analysis on the explanatory variables found evidence that supports the view that these explanatory variables explained some of the differences in the migration outcomes of students. Further analysis was then conducted to answer the aforementioned research questions. The three outcome variables used in the analysis were quite different in their format and as a result they required different methodologies to model the outcomes against the different explanatory variables simultaneously. However, the findings from this paper indicate that despite the complexities and different techniques available to measure and quantify student migration, the three outcome variables used in this analysis illustrate very similar results.

The main findings indicate that ethnicity, social background and gender all have a significant impact on the student migration experience in order to attend a HEI in the UK. The most concurrent finding across the three techniques was the group most likely to migrate, travel the furthest distances and be internal student migrant were students from the White ethnic group, most advantageous background and were male. In contrast, the group of students least likely to migrate, travel the shortest distances and be local students were from the Asian ethnic group, least advantageous social background and were female.

All three techniques showed similar findings when focusing on ethnicity. Students from the White ethic group were considerably more likely to migrate in order to attend a HEI than their Black and Asian counterparts. The White ethnic group were predicted to travel the furthest distances, followed by the Black group and then the Asian group. While, White students had the highest probability of being an internal student migrant and the Asian students had the highest probability of being a local student.

Similarly to the results for ethnicity, when analysing the impact of the students social background the three techniques all produced reciprocal findings. Students with the most advantageous social backgrounds were most likely to migrate, travel the furthest distances and be an internal student migrant, and as the level of advantage in social background declined, so did the probabilities of being a migrant.

The variables of gender and age also impact on the findings in all three techniques. Males were more likely to migrate and migrate further distances than their female peers and these differences were present regardless of social background or ethnicity. There was however an apparent interaction effect particularly between the Asian ethnic group and females, with this group being considerably less likely to migrate than any other. The gender differences were least apparent in the multinomial technique for the student migrant categories, however this may be a product of this technique not including the interaction term that were present in the first two methodologies. With regards to age the concurrent theme across all the results was that as aged increased the probability that the person would migrate for HE purposes and the distance travelled significantly decreased.

The impact of the institution attended and course studied were also similar regardless of the outcome variable analysed. The probability of migration and the predicted distance travelled were all significantly impact by the type of institution attend. Students attending Pre1992, Top30 and Russell Group universities were more likely to migrate and were predicted to travel further to do so than those students that did not. With regards to course studied there was not a large amount of difference between the probability of migration and predicted distance travelled for many of the courses. However, those students studying agricultural or veterinary courses were significantly more likely to migrate and were predicted to travel further distances than those students studying other courses. This may be a result of fewer institutions offering these courses and those that do being in rural and isolated areas. As a result students have fewer choices and therefore students wanting to study these courses are forced to travel larger distances.

The results presented throughout this paper have analysed the trends between the available variables within the dataset. However, the underlying factors that influence the student migration decision process are plentiful. It must be noted here that one of the limitations of this study is that it was not possible to quantify or to take into account several of the factors identified as influencing student migration in our analysis. One of the key influencing factors that was not taken into account in this analysis was the impact of a student's achievement level on the student migration outcome. This was not included as the variables were not available in the dataset used. Student achievement is directly linked to the HE admissions process and will often influence the student in their choice of course and HEI. There have been previous studies that have linked the achievement level of students to their socio-economic status, ethnicity and the level of schooling at earlier stages of the education system. In our findings we concluded that the student migration outcome was influenced by the student's ethnicity, social background and gender. However, due to the inability of this study to disseminate the results by factors not included in the models, these findings might be a by-product of other influencing factors that cannot be identified within this study, such as student achievement and levels of deprivation.

Further extensions of this work could include changing the focus of the outcome variables. Within the dataset it is possible to see how social background, ethnicity and gender impact on the institution attended or course studies instead of the focus here on the migration transition experienced. This work could also be extended by conducting a qualitative study to find more in-depth reasoning behind the observed differences between the sub-groups in the student population as the current quantitative study can only illustrate that such differences exist but do not indicate explicitly why these differentials a so apparent.

The findings from this paper indicate that despite the complexities and different techniques available to measure and quantify student migration, the outcome variables used in this analysis all illustrate very similar results. All techniques undertaken here have suggested there to be substantial differences between the ethnic and social background groups as well as significant gender differences in the patterns of migration into HE in the UK. The use of three different techniques and the cofounding results provide statistical evidence that support the findings of ethic, social background and gender differences in the migration decision process of students and that access to HE is still not equal across the social spectrum in the United Kingdom.

References

Agresti, A. 2013. Categorical Data Analysis. 3rd Edition. New Jersey: John Wiley & Sons.

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

Bell, M., M. Blake, P. Boyle, O. Duke-Williams, P. Rees, J. Stillwell and G. Hugo. 2002. "Cross-national comparison of internal migration: issues and measures." Journal of the Royal Statistical Society Series a-Statistics in Society 165:435-464.

BIS. 2013. "Participation rates in higher education: academic years 2006/2007 – 2011/2012 (provisional)." ed. Matthew Bollington. London: Department for Business, Innovation and Skills.

Blanden, J. and P. Gregg. 2004. "Family income and educational attainment: A review of approaches and evidence for Britain." Oxford Review of Economic Policy 20(2):245-263.

Blanden, J., K. Hansen and S. Machin. 2010. "The Economic Cost of Growing Up Poor: Estimating the GDP Loss Associated with Child Poverty." Fiscal Studies 31(3):38-73.

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

Chowdry, H., C. Crawford, L. Dearden, A. Goodman and A. Vignoles. 2008. "Widening participation in higher education: analysis using linked administrative data." Institute for Fiscal Studies.

Chowdry, H., C. Crawford, L. Dearden, A. Goodman and A. Vignoles. 2010. "Widening Participation in Higher Education: Analysis using Linked Adminstrative Data." In IFS Working Paper W10/04: Institute for Fiscal Studies.

Christie, H. 2007. "Higher education and spatial (im)mobility: nontraditional students and living at home." Environment and Planning A 39(10):2445-2463.

Cochran, W.G. 1939. "The Use of the Analysis of Variance in Enumeration by Sampling." Journal of the American Statistical Association 34(207):492-510.

Dearing, R. 1997. "Higher Education in the Learning Society." In Report of the National Committee of Inquiry into Higher Education. Norwich: HMSO.

Department for Education and Skills. 2003. "Widening Participation in Higher Education." London: DfES.

Department for Education and Skills. 2006. "Widening Participation in Higher Education." London: DfES.

Dorfman, A.H. and R. Valliant. 2005. Superpopulation Models in Survey Sampling.

Galindo-Rueda, F., O. Marcenaro-Gutierrez and A. Vignoles. 2004. "The widening socio-economic gap in UK Higher Education." National Institute Economic Review 190.

Hartley, H.O. and R.L. Jr. Sielken. 1975. "A "Super-Populaiton Viewpoint" for Finite Population Sampling." Biometrics 31(2):411-422.

HEFCE. 2013. "Widening Participation Policy." ed. Higher Education Funding Council for England.

Higher Education Statistics Agency. 2012. "HESA Student Record Data - 2007/08 to 2010/11."

Higher Education Statistics Agency. 2013. "HESA Student Record Data - 2011/12."

HMSO. 1992. "Further and Higher Education Act." London.

Khambhaita, P; and K; Bhopal. 2013. "Home or away? The significance of ethnicity, class and attainment in the housing choices of female univeristy students." Race, Ethnicity and Education:1-32.

Long, J. 1997. Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage Publications.

Machin, S. and A. Vignoles. 2004. "Educational inequality: The widening socio-economic gap." Fiscal Studies 25(2):107-128.

McDonald, J.F. and R.A. Moffitt. 1980. "The Uses of Tobit Analysis." The Review of Economics and Statistics 62(2):318-321.

Modood, T. and M. Shiner. 1994. Ethnic Minorities and Higher Education: Why are There Differential Rates of Entry? London: Policy Studies Institute.

Munro, M., I. Turok and M. Livingston. 2009. "Students in cities: a preliminary analysis of their patterns and effects." Environment and Planning A 41(8):1805-1825.

Office for National Statistics. 2013. "Neighbourhood Statistics." ed. ONS. Titchfield.

Patiniotis, J and C Holdsworth. 2005. "'Seize That Chance!' Leaving Home and Transitions to Higher Education." Journal of Youth Studies 8(1):81-95.

Shiner, M. and T. Modood. 2002. "Help or Hinderance? Higher Education and the Route to Ethnic Equality." British Journal of Sociology of Education 23(2):209-232.

Tobin, J. 1958. "Estimation of Relationships for Limited Dependent Variables." Econometrica 26:24-36.

Tomlinson, S. 2001. "Some success, could do better: education and race 1976-2000." In Education, Reform and the State: Twenty-Five Years of Politics, Policy and Practice, eds. R. Phillips and J. Furlong. London: Routledge.

Wilson, T. 2010. "Model migration schedules incorporating student migration peaks." Demographic Research 23:191-222.