UNIVERSITY OF Southampton
Southampton Statistical Sciences Research Institute

Working Paper A08/08

Applications and Policy

Wage Expectations of UK Students: How Do They Vary and Are They Realistic?

John Jerrim

Abstract

The wage expectations of university students have relevance for human capital theory, models of student enrolment, and public policy on provision of higher education. However these expectations have been the subject of relatively little research in industrialised countries. The paper investigates students' expectations with UK survey data that are superior to those used in several existing studies in various ways, including their coverage of a representative sample of universities and the rich nature of the information on socio-economic background. The analysis first shows how wage expectations vary with time remaining at university, ability and quality of the university. The paper then compares expectations with the actual wages earned by the same cohort on graduation, using data that are drawn from an attempted census of all graduates. This provides a more reliable basis for assessing the realism in students' expectations than in existing studies. Results support the hypothesis that full-time students overestimate their starting salary, bringing a fundamental assumption of human capital theory into doubt.

Wage Expectations of UK Students: How Do They Vary and Are They Realistic?

John Jerrim

School of Social Sciences

University of Southampton

October 2008

Abstract

The wage expectations of university students have relevance for human capital theory, models of student enrolment, and public policy on provision of higher education. However these expectations have been the subject of relatively little research in industrialised countries. The paper investigates students' expectations with UK survey data that are superior to those used in several existing studies in various ways, including their coverage of a representative sample of universities and the rich nature of the information on socio-economic background. The analysis first shows how wage expectations vary with time remaining at university, ability and quality of the university. The paper then compares expectations with the actual wages earned by the same cohort on graduation, using data that are drawn from an attempted census of all graduates. This provides a more reliable basis for assessing the realism in students' expectations than in existing studies. Results support the hypothesis that full-time students overestimate their starting salary, bringing a fundamental assumption of human capital theory into doubt.

Key Words: Expectations, Higher Education, Wages
JEL Codes: I2 Education, D84 Expectations

Contact Details:
John Jerrim, Division of Social Statistics, School of Social Sciences, Southampton University, United Kingdom. E-mail: jpj106@soton.ac.uk

"The contempt of risk and the presumptuous hope of success are in no period of life more active than at the age at which young people choose their professions."

Adam Smith, The Wealth of Nations (1776), Page 109

1. Introduction

The returns to education have been widely explored in England and Wales. Some recent studies suggest students in particular subjects receive, over their lifetime, poor financial returns to their investment in university education (Vignoles 2007). The fact students still decide to enter university and take these courses is often explained by the value of non financial benefits, such as the joy of learning and experiencing independence. However, an often overlooked possibility is that students may be unrealistically optimistic about their future opportunities when deciding to go to university, as noted by Adam Smith in the quote above. This paper investigates the variations in students' expectations and whether these expectations are in line with wages in the graduate labour market. Discussion follows on the provision of higher education across Europe, with topics including student debt, fees and widening access schemes. The UK is used as an example, with specific reference to the government's objective of getting 50% of school leavers to experience higher education. The importance of the research to human capital theory is also discussed, and how this may affect academic models of school enrolment.

This topic has a small literature within America and Continental Europe, although no recent academic research on students' wage expectations has been conducted within the UK. These existing European studies are severely limited by their reliance on convenience samples drawn from a small number of institutions and subjects. Therefore results are unlikely to generalise to any meaningful population. The first part of this paper addresses this issue by using a representative sample drawn from the Department for Education and Skills Student Income and Expenditure Survey to estimate a model of student wage expectations. I test the hypothesis that students near the beginning of their course expect significantly higher wages than those about to graduate, and investigate the impact of several characteristics relating to students and their institutions. The conclusion reached is that the seniority of the student and individual ability has a significant influence on wage expectations.

The second half of this paper considers whether students hold realistic wage expectations. Existing European studies use of unrepresentative samples causes particular difficulty in comparing students' wage expectations to the actual earnings of graduates. A highly selective cross sectional survey on wage expectations is usually compared with historical data on graduate wages. Almost no attention is paid to whether the surveys are comparable, with problems such as selectivity, induced by convenience sampling or non-response bias, largely ignored. This severely hinders the existing studies assessment of how realistic students are. In comparison, this study compares a representative sample of students' wage expectations with the average realisations of groups from the same cohort, drawn from an attempted census of all graduates, providing a better basis for comparison. The comparability of the two surveys is also thoroughly discussed, with the results further checked for robustness using the Labour Force Survey. Results suggest that on average full-time students overestimate their first salary after university, though this varies with the subject being studied.

The paper begins by reviewing the current literature and describing the available datasets. A model of UK students' wage expectations follows in section 4, with

discussion of results in relation to the seniority of the student and various background characteristics. The final two sections compare students' expectations with actual graduate wages, before a discussion of what the findings imply for academic models of school enrolment and higher education policy across Europe.

2. Current Literature and Research Questions

There have been a small number of studies investigating students' wage expectations across America and Continental Europe. A common theme is that students who are further through their course have lower wage expectations than those at the beginning, reflecting better knowledge of their own ability and chances in the graduate labour market. Betts (1996) finds that students do not gather information until a late stage. He therefore concludes that students near the beginning of their course have reasonably poor labour market knowledge. Brunello et al (2001) show a similar pattern in their cross European study; students further through their course tend to not only expect lower wages, but are also less optimistic about their employment prospects. This begs the question, will the difference in wage expectations, based on the seniority of the student, remain once views of employability have been controlled for? Indeed will the same pattern be observed within the UK at all, using a nationally representative sample of all students?

A topic that has received rather less coverage is the role that ability and university prestige play in students' wage expectations. Smith and Powell (1990) took samples from two universities in America that differed in quality. Students at the elite university were found to have higher wage expectations, conditional on their pre-university high school rank. Brunello (2001) looked at the expected wage gain[1] in relation to university status, and found that only tighter admission criteria had a significant impact on expectations. However neither study draws their sample from a large number of institutions within one country. The UK is a particularly interesting setting for such research, with the number of universities having grown dramatically since government expansion of the higher education sector in 1992, creating a large variation in terms of standards and prestige. Moreover previous research in the UK suggests that the quality of an institution may affect the actual wages of graduates. Chevalier and Conlon (2003) find a premium of going to a prestigious UK university above and beyond the influence of ability. They conclude that this presents an economic argument for these institutions being allowed to charge higher tuition fees. However do students expect such a premium in their wages? If not, then one may question whether students would be willing to pay the higher tuition fees suggested.

Significant investigation has also been conducted into how socio-economic differences affect students' wage expectations. Parents are assumed to be one of the most important sources of students' labour market knowledge, with expectations based around what they earn. Webbink and Hartog (2003) found students from high income families expect significantly more than those from poorer backgrounds, but that they are also more likely to overestimate their future wage. Smith and Powell (1990) also found this positive association between parents' income and students' wage expectations. One piece of work conducted in the UK by Williams and Gordon

[1] Expected salary with a degree minus expected salary without a degree.

(1981)[2] looked at the impact of socio-economic variables on the wage expectations of students at the end of compulsory education. However they found that socio-economic status had little direct influence on students' expected lifetime gain from going to university. Other variables typically investigated include gender, age and the education and occupation of parents. However less attention has been paid to differences based on characteristics such as ethnicity. For instance, do ethnic minorities anticipate some form of discrimination in the labour market and therefore lower their wage expectations?

Some of these existing studies go onto make a rough comparison between students' expectations with wages in the graduate labour market. For instance, Wolter (2002) shows that students tend to overestimate their wage with a degree. Smith and Powell (1990) suggest students are well informed about average wages, but tend to overestimate their own returns. The one known European study that uses longitudinal data, by Webbink and Hartog (2003), comes to a different conclusion; students can accurately predict their starting salary. However a difficulty encountered in the existing studies is that they tend to compare a highly selective convenience sample of students' expectations with historical data on graduate wages. The characteristics of the two samples covered in each survey are usually not even discussed, even though the sampling designs mean that they relate to very different, and incomparable, populations. Even Webbink and Hartog (2000) advise caution generalising results in their longitudinal survey, due to the highly selective nature of follow-up[3]. As such, no paper thus far convincingly illustrates whether students' expectations are realistic. Indeed there is some disagreement in the literature. Whereas Webbink and Hartog (2003) boldly label their paper,

"Can Student's Predict their Starting Salary? YES!",

Betts (1996) concludes that on average students overestimate their starting wage by 10%, while Brunello (2001) suggest the figure could be even higher than this. This paper hopes to resolve this conflict by comparing two surveys that cover largely comparable populations, to assess how realistic UK students are in their expectations. Issues of non-response and comparability are addressed directly, with further robustness checks using an additional data source.

The results have substantial policy implications for Europe's expanding higher education sector. In particular, the UK government is aiming to get 50% of school leavers to enter university, with several schemes, such as "Aim Higher", in place to encourage young adults to continue their education. However are these schemes simply reinforcing students' unrealistic expectations? Young people may be entering university based on unrealistically high expectations of their prospects on graduation. This may be further exacerbated by government policy that encourages young people to go to university, highlighting the magnitude of possibilities the student will have on graduation, that never actually materialise. A further possibility is that students are willing to take on high levels of debt because they believe their future wages will enable repayment of their student loans. Gustman and Stafford (1972) show that the

[2] Two known UK studies have been conducted, one by Williams and Gordon (1981) and the second by Bosworth and Ford (1985). Also Brunello (2001) contained some information from the University of Sterling and University of Essex, though the sample sizes were small.

[3] Out of an initial sample of 3,845, only 676 (17%) of cases have complete information. This is discussed further in section 7.

higher the income expectations of students, the more they consume. However if their expectations are unrealistic, students may over consume during university, leading to difficulties and debt in later life. Therefore this research also adds to the debate on how to finance higher education across Europe, paying particular attention to the provision of student loans and debt.

On the basis of the international literature and current policy interest across Europe, the research questions to be explored in this paper are as follows:

1. Do students who are further from graduation have greater wage and employment expectations?

2. Do students at elite universities have significantly higher wage expectations than those at less prestigious institutions? What is the relationship between ability, university prestige and wage expectations?

3. Are parental income and ethnicity associated with students' wage expectations?

4. Do students have realistic expectations? Do students who are studying a subject directly leading to a career have more realistic wage expectations?

To my knowledge, this paper provides the first study in Europe using a representative national sample of all university students. The first question follows much of the existing research, but extends the analysis to show how students' views of their employment prospects influence their wage expectations. On the other hand, the second question has received little attention in the existing literature, due to the reliance on convenience samples taken from a small number of institutions. The definition of 'elite' in this work is whether the institution belongs to the 'Russell Group'; a self-selected alliance of the top 20 research-intensive universities[4]. Question three attempts to look at some previously neglected variables such as differences in wage expectations between ethnic groups. Finally, I investigate whether students are realistic, the first such study conducted in the UK.

[4] For further details see http://www.russellgroup.ac.uk/

3. Data on Wage Expectations

One reason why more research has not been done in this area is the lack of available data. One possible source is the Association of Graduate Recruiters Graduate Career Survey. However this study only targets the "top 30" UK universities[5], and therefore does not cover the whole UK student population, leading to an unrepresentative sample. Several methodological problems also exist with the sampling strategy used and with the reliability of responses. The 2004/5 Student Income and Expenditure Survey (SIES) is an alternative source. This survey was carried out using face-to-face interviews between January and March 2005 by the Institute of Employment Studies and the National Centre for Social Research on behalf of the then Department for Education and Skills.

The purpose of the study was to generate a representative sample of all higher education students in England and Wales, in order to investigate income and expenditure patterns. One strength of using this dataset is that it contains detailed information on a number of potential explanatory variables. This allows analysis of potential sources of variation in wage expectations that have been neglected in previous studies. Information is provided on the students' current year of study and the length of their course, providing valuable information regarding the first research question. The number of universities included in the survey provides a large sample of students from a range of institutions. This allows a detailed investigation across both universities and subjects within one country; a further topic with little coverage in the existing literature. There is also information on students' background, including ethnicity, social class and previous schooling. Other controls such as gender and whether the student is classed as 'dependant'[6], meaning they are in full-time education and had their parents' income taken into consideration when applying for student support, are included. For 'dependant' students, there is also an approximate measure of family income, though it can only be taken as a proxy due to the way this data has been collected and recorded[7]. Unfortunately some other important information is missing; in particular there is no indicator of student ability.

A complex sampling design was used to ensure a representative cross-section of students was selected. Universities were sampled using a probability design based on the size of the institution. There was also stratification by region and whether it was a "pre 1992" or "post 1992" university.[8] A sample of 80 universities, from a population of 132, was drawn, with probability proportional to size. In total, 69 universities agreed to take part. All these universities were included in the final sample, with the intention of contacting 240 randomly selected students from each institution. Separate samples of full-time and part-time students were drawn, with special previsions made for those institutions with medical schools.[9] 25 Further Education Colleges (other degree awarding institutions) were also approached, with 19 taking part. From each of

[5] A "top 30" university in this case is defined by the Association of Graduate Recruiters. The majority of universities included in the survey are Russell Group institutions, known for their excellence in research.
[6] Full details are given in appendix 1 about the survey definition of this variable.
[7] Further details are given in appendix 2
[8] A "post 1992" university is an institution that achieved university status in 1992 or later. This date marks a major change in the UK higher education sector, when several polytechnic institutions were given degree awarding powers. This increased the number of students at universities dramatically.
[9] Further details can be found in the 2004/2005 SIES technical report.

these institutions, 60 students were randomly selected. Across all institutions, a total of 16,524 students were selected to take part. These students were each mailed an initial "opt-in" questionnaire, where they were asked to provide some basic information and whether they consented to be contacted to take part in the research. 7,548 (45%) opt-in questionnaires were returned, with 5,810 (35%) giving their consent to take part. In total 4,570 names were issued with 3,548 interviews achieved.[10] For the purpose of this study, students who did not report their expected starting salary, along with those studying at further education colleges or for qualifications other than at degree level, were dropped. A further 50 (2%) observations were dropped from the dataset, where the expected starting salary was below £3,000[11]. The final dataset contains 2,744 observations, with the sample selection rules presented in Table 1.

Table 1

The level of non-response is not negligible, and obviously has implications for the generalisability of results. Those that take part in the survey could be systematically different to those who opt out. To address this, observations were weighted to correct for the probability of a student being selected and responding.[12] A second stage of weighting was also conducted to ensure the sex and age profile of students matched that of Higher Education Statistics Authority records. An important implication for the research is whether the sample drawn accurately reflects the wider student population. The SIES 2004/2005 report states,

"As can be seen, this was an ambitious methodology but one which succeeded in producing the objective of a nationally representative student sample for interviews." (P 10)

It does indeed appear that every attempt has been made to investigate and correct for any bias in the sample, though it should be noted that the use of sample weights can only correct estimates in terms of observable characteristics. In comparison to most of the studies on wage expectations discussed in section 2, the SIES data has the advantage that it is designed to draw a representative sample from the population of students, rather than relying on a simple convenience sample. Moreover, data is drawn from around 70 institutions across the whole spectrum of subjects, whereas most of the existing research can only boast a handful of subjects from a couple of universities. Although non-response does cause some limitations, it is reasonable to say the SIES is much more likely to be representative of the wider student population than any previous study and therefore provides a better source for analysis. Furthermore, this is the first study to consider the impact of non-response on estimates,

[10] Another institution that mainly involves part-time distance learning, The Open University, was in the original dataset but was dropped as these students did not give details on their wage expectations.
[11] Almost half (23) of these values were at £1, and thus largely reflect illogical answers. Results were also checked for robustness using £8,000 as the minimum allowed expected salary. When this is done, all the substantial conclusions in the following sections remain intact.
[12] Weights were calculated as the *inverse* of the probability of being both selected and responding to the survey, and were the product of five conditional probabilities. The loss of effective sample size due to weighting was only moderate for full-time students (where the effective sample size was 88 per cent of the actual sample size) but relatively high for part-time students (where the corresponding proportion was 62 per cent). Further details are provided on non-response and weighting in the 2004-2005 Student Income and Expenditure technical report.

with the robustness of results considered to various specifications and use of response weights.

Another critical part of the survey is how students report their wage expectations. They were each asked the following question:

"What sort of salary do you expect to be earning in the first job you take once you have graduated?"
Interviewer comments: If not sure of the exact amount, please give your best estimate.

Students are clearly asked for their expected salary, to be recorded in an open text field, allowing precise estimates. This is interpreted as students giving the mean of all possible outcomes they face. In other words, students are providing the arithmetic mean for the entire distribution of all possible outcomes.[13] A further issue is that the question asks students about the first job they take after university. Students are not asked explicitly whether they expect this to be full-time or part-time work, or if this will be temporary while they look for a job directly related to their career aspirations. Nethertheless, Manski (1996) suggests that students interpret questions regarding future salary expectations on the assumption that they will be in full-time employment. Thus it seems reasonable to assume expected salary corresponds to students' first full-time job after university.

When interpreting the data it is also assumed that students are providing a gross, yearly figure. Although ideally this would be made explicit in the question, it seems reasonable to assume students would report figures in this way, as it is the standard method of advertising salaries in the UK. Assumptions must also be made about how students deal with inflation when forming their wage expectations. The most common assumption is that students do not consider inflation, and are thus reporting in 2005 prices, as discussed in Wolter (2002), Manski (1993) and Dominitz & Manski (1996)[14]. This is the approach also taken in this paper.

Brunello (2001) and Dominitz & Manski (1996) also note that respondents tend to round their estimates to questions surrounding expectations to the nearest 5 or 0. The histogram of expected salaries below, and Table 2, shows that the distribution of log expected wages is broadly symmetric, though there is bunching of estimates to the nearest £1,000. Large spikes are especially prelevant at multiples of £5,000 (for instance £15,000, £20,000 and £25,000), with other instances at £12,000 and £18,000 (equivalent to a salary of £1,000 and £1,500 per month). Brunello et al (2001) go on to say that there is no evidence to suggest students do not take care when completing their questionnaire.[15] However no existing study tries to take this explicitly into account. This paper goes a stage further and checks the robustness of results to this heaping in the data. A description of the techniques used can be found in appendix 3.

[13] Ideally, a precise definition would be provided to the students as Manski (2004) suggests when eliciting students' median expectation of their future wage distribution. However I feel this is a reasonable assumption of how students would interpret the question posed.
[14] One exception is Webbink and Hartog (2003), where they use longitudinal data, but their assumptions about inflation are never made clear. Here they compare expectations in 1991 with realisations in 1995, with it seemingly implicitly assumed that respondents give their expectation in 1995 prices (i.e. they account for future inflation and productivity growth in their forecasts).
[15] Manski finds a similar phenomenon when eliciting individuals subjective probabilities, where individuals round to the nearest 5% or 10%.

Figure 1. Distribution of log expected starting wage

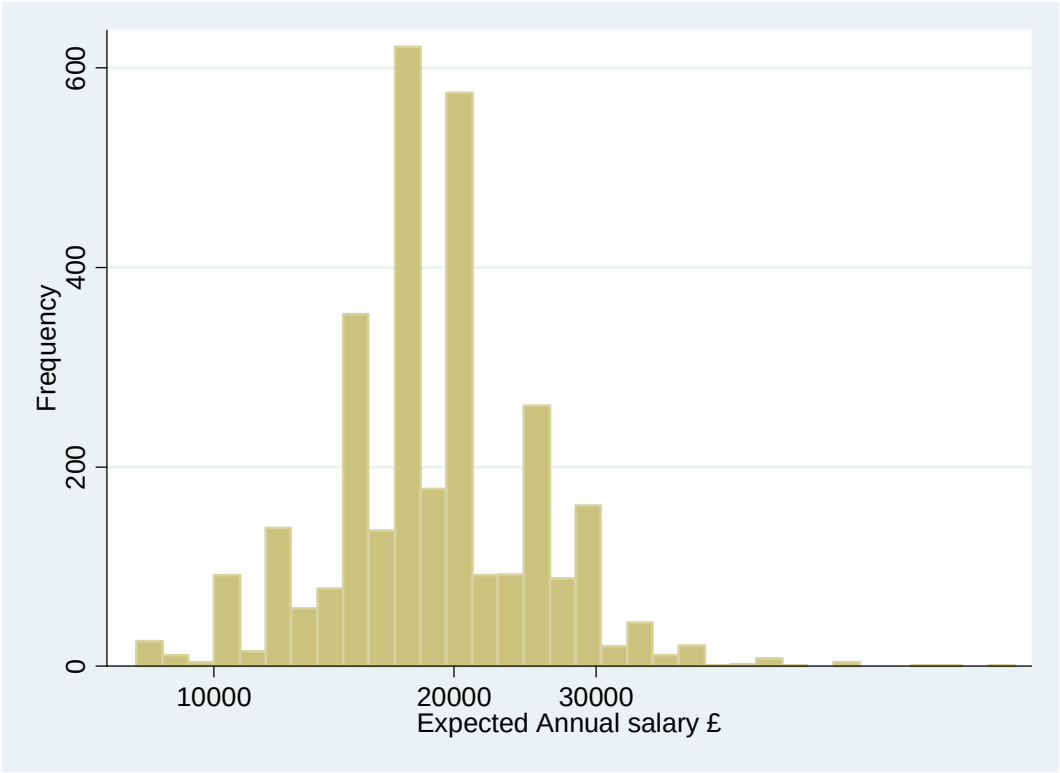


Table 2

A final point that often concerns economists when using subjective data is that respondents have little incentive to report their expectations accurately. A recent study by Botelho and Pinto (2004) tested this issue in an experimental setting, and found financial incentives have little impact on the accuracy of students reported expectations. Moreover Manski (2004) puts over compelling evidence on the measurement and use of expectations in economic research. These studies conclude that economists' scepticism of subjective data is largely unwarranted.

To investigate research questions one to three, a statistical model of wage expectations using the following specification[16] has been developed:

$$\text{Log}(W_{i,j}) = \alpha + \boldsymbol{\beta}\mathbf{X}_i + \psi P_i + \text{ϒ}U_j + \boldsymbol{\lambda}\mathbf{V}_i + \Gamma A_I + \xi_{i,j}$$

With W = Student's expected wage
X = Matrix of background characteristics (e.g. gender)
P = Proximity to graduation
U = University type
V = Student's view on their employability
A = Proxy for individual ability
ξ = error
i = for individual i
j = for university j

[16] Peer group affects were also considered in the initial model selection process. However no evidence of peer group affects was found, with all results small and statistically insignificant.

The first specification will exclude the matrix $\mathbf{V}_i$, that refers to students views of their employability[17]. It is then included in all following specifications. Each of the first two specifications will also *exclude* the proxy for individual ability (A_i), which shall then be included in specifications 3 and 4. The use of this proxy brings about some difficulties. It is calculated by taking the average university entry score[18] for each combination of subject (broken down into 20 subjects) and university based on High Education Statistics Authority (HESA) administrative data[19]. For example, the average entry score for Social Science students at Bristol University, calculated using the HESA data, is 400. Anyone in the SIES, who was at Bristol University doing a Social Science course, then had this imputed as their ability score. Institution and subject are highly correlated with students' UCAS score, thus this technique seems to provide a reasonable proxy for the missing ability variable[20]. However this does limit the assumptions one can make about the error term, and how the complex sampling design is taken into account. For instance, when including this proxy it is no longer sensible to use a fixed affects model including a dummy variable for each institution, as they will suffer a problem of collinearity[21]. Moreover, a random affects model was considered inappropriate, as it is highly likely the university ("level two") affect would be correlated with terms at the individual ("level one") level. For instance, the university affect would necessarily be correlated with the ability proxy itself, thus violating a standard assumption for using a random affects model. Hence, in these instances, the effects of clustering and stratification shall be accounted for by adjusting the standard errors. Results for specifications 1 and 2, which do not contain the ability proxy, will also be presented in this way for consistency. However a fixed affects model will also be presented for these initial specifications as a check for robustness. In all specifications, the error term is assumed to be normally distributed and will contain, amongst other things, information relating to missing variables. Summary statistics for the explanatory variables, and their relationship with wage expectations, is presented in Table 3.[22]

Table 3

[17] One may suggest the variables in matrix $\mathbf{V}_i$ are potentially endogenous. In this model, I assume that if individuals are less optimistic about the labour market, they will lower their wage expectations. However, another possibility is that students who expect higher wages, and who possibly also have a higher reservation wage, limit their labour market opportunities and are therefore less optimistic about their post university prospects. Leaving $\mathbf{V}_i$ out of the first specification gives an indication of results if one considers this possible endogeneity issue to be a serve problem.

[18] Known as UCAS score in the UK, which reflects performance in public exams that are typically taken at age 18

[19] This is the same data source used in sections 6 and 7 of this paper. It relates to the UCAS score of students who finished university in 2005. For further details see :http://www.hesa.ac.uk/index.php/content/view/97/136/#alevel

[20] Moreover, although Chevalier and Conlon (2003) argue that there are some students with high UCAS scores at lower quality institutions, their difficulty in finding common support in their propensity score matching estimates illustrates that institution, in particular, helps to serve as a good proxy for ability.

[21] The results including the ability proxy in a fixed affects model is actually presented later in table 7 to illustrate this point.

[22] All the following analysis use the sample weight provided in the SIES dataset. Table 3 contains the unweighted sample size, while the summary statistics are provided using the weights. The weights have a substantial influence on the gender composition, increasing the proportion of men from 33% to 46%.

4. Proximity to Graduation

A number of existing studies have found junior students to be more optimistic than those nearing graduation. Betts (1996) concluded that students lowered their expectations due to "learning effects", where individuals discover more about their ability and the labour market as they move through tertiary education. Brunello et al (2001) found similar results in their cross European study, identifying senior students to be less optimistic about wage levels and employment prospects. The initial research question follows these studies and asks if students further from the labour market are more positive about their employment prospects and have higher wage expectations. This is then extended by investigating whether wage expectations still differ after controlling for students' views on their post graduation prospects.

Full results are provided in Table 4, containing two specifications of the model set out in section 3. The first specification follows the traditional approach in the literature and does not contain students' views of their employability after university (**V**) as explanatory variables. These are then included in the second specification.[23]

Table 4a and 4b[24]

Initial results, using the first specification, support the existing studies. Students one year away from graduation expect around 3% more on average than final year students, while those 2 or more years away expect around 7% more.

An interesting extension is whether students who are further from the labour market also hold more positive views on their employability after university. Two variables within the SIES shed some light on this issue. Firstly, respondents were asked whether it will be hard for them to get a graduate job. In particular, students were asked to what extent they agree or disagree with the statement:

'The growing number of graduates will make it hard to get a graduate job'

Secondly, students were asked about their post university plans, including whether they have considered taking a temporary job[25]. These two questions provide a reasonably good indication of student views on their employability. Tables 5 and 6 contain cross-tabulations between these variables and how close the student is to graduation.[26]

Table 5

Table 6

[23] The results and interpretation presented in this section are robust to each of the subsequent specifications used, including the introduction of a fixed affects model.

[24] Results have been split into two tables. Table 4a contains variables directly relating to the research questions posed, while 4b contains the other control variables.

[25] Students were asked if they planned to get a job related to their future career, a temporary job, continue studying or go travelling. They could identify more than one option; therefore this gives a rough indication of students' future plans.

[26] Logistic and ordinal regression was also used to investigate how various variables affect the responses to these two questions. However, cross-tabulations have been presented rather than the results from these models for ease of interpretation, with little evidence lost in relation to the research question posed.

Both sets of results indicate final year students generally hold more negative views. Only 2% of students who are two or more years away from graduation considered taking a temporary job after university, compared to nearly 10% of final years. This could be interpreted in several ways. It is suggested that final year students are less positive about their labour market prospects. However an alternative explanation could be that these students are looking to delay the coming of the "real world". Investigation of Table 5 however shows that a high proportion of final year students, compared to those one or two years away, agree or strongly agree that the growing number of students will make it hard to get a graduate job. The combined evidence does indeed suggest final year students are less optimistic about the graduate labour market. This may be due to the fact that final year students probably know more about their expected grade, and link this to their employability.[27] Alternatively, given the survey is conducted between January and March, they may well have already started their hunt for a graduate job, and have thus far been unsuccessful.

A question that is ignored in the current literature is whether, after controlling for views on employability, students at the beginning of their course still expect a higher starting wage than those in their final year? The two additional variables, analysed in the cross-tabulations, are included into the regression model in the second specification[28]. The impact of being further from graduation on wage expectations has been significantly reduced. The coefficient for students a year away from graduation stands at 1.6%, and is not statistically significant. Previously students two years or more away from graduation expected a 7% premium compared to final years. This almost halves to 3.6% when their future plans and opinions about the graduate labour market are taken into consideration.

This pattern could represent either a cohort or age effect. Given that other research offers similar results, it seems reasonable to suggest this represents a changing of students' views as they progress through university, rather than a difference between particular cohorts. At first, it seems there is a large difference in wage expectations between year groups. However students appear to differ in their views of the labour market in at least two aspects, namely their employability and the wages on offer. Once views of employability have been controlled for, the difference in wage expectations appears to be reasonably small.

[27] Indeed a recent paper by Chevalier (2008) indicates that first year undergraduate students over-estimate their ability. It could be that as students move through university they learn more about their ability and alter their expectations on the receipt of this new information.

[28] Further specifications of the model follow, with little influence on the substantial result and subsequent analysis.

5. University Prestige, Ability and Family Background

One fundamental drawback in the existing literature is the lack of available data drawn from a range of institutions within one country. Convenience sampling, generally of small sizes, also makes research on background characteristics rather limited. Although socio-economic class is often well covered, differences in other minority groups, such as Black and Asian students, has received less attention. The UK is a particularly well placed to address these issues.

It is also an interesting setting to investigate the effect of institution on wage expectations, due to its large and expanding higher education sector and the drive to widen participation. Moreover, existing research in the UK highlights how attending a prestigious UK institution can increase future wages over and above individual ability. For instance, Chevalier and Conlon (2003) suggest that the wage premium of going to a Russell Group is up to 6% after ability has been controlled. However, do UK students expect to receive such a premium? Our attention now turns to these issues.

I. University Prestige, Ability and Expectations

Students at older, more prestigious universities are hypothesised to have higher wage expectations. This is thought to reflect, in part, these students being of higher ability. Chevalier and Conlon illustrated a 'Russell Group premium', where students at these universities receive higher wages. This research investigates whether students at Russell Group universities internalise this premium in their expectations. In particular, do students at Russell Group universities hold wage expectations above and beyond what can be explained by their ability[29]?

In the initial specifications (Table 4), universities have been split into three groups, Russell Group, Other Pre-1992 and post 1992[30]. Students' university entry score (UCAS score) reflects performance in public exams that are typically taken at age 18. This varies greatly between the three university groups. Students at Russell Group universities enter with an average UCAS score of 400, compared to 329 for other pre 1992, and 237 for post 1992, institutions. [31] The results from the first two specifications indicate that students at other pre-1992 universities expect almost identical wages to those at post 1992 institutions. This is despite the fact that the former tend to admit higher ability students and generally have a better reputation within the UK higher education sector. However, as expected, students at Russell Group universities expect significantly more than the other groups[32].

---

[29] I am assuming university type to be exogenous in the model of wage expectations; going to a prestigious university raises students' wage expectations. However, one may consider university type to be potentially endogenous. It could be argued that students at Russell Group universities face higher costs, such as having to move further to attend a prestigious institution. Students with higher wage expectations may be more willing to suffer these costs, and thus influence their choice of university. Though one can not rule this possibility out, I believe treating university type as exogenous is not an unrealistic assumption to make.

[30] "Post 1992" universities gained their university status in or after 1992. "Other Pre 1992" institutions had university status before 1992 but are not member of the Russell Group. These universities generally take high standard students but do not have the research standing of the Russell Group.

[31] Entry scores were calculated from 2005 HESA student record data.

[32] This becomes insignificant in the fixed affects specification, though the coefficient is still large. This is due to standard error becoming significantly larger due to the correlation between the university dummies and "university type" (Russell Group, Pre 1992, Post 1992) variables.

Is this just a reflection of the Russell Group universities having students of higher ability? In the model specifications thus far a measure of individual ability has been omitted. To address this problem, the proxy for ability, discussed further in section 3, is imputed for each individual in the dataset. This proxy is based on the subject the student is studying and the university they are attending. Specification 3 now contains this control for ability, with results in Table 7.

Table 7

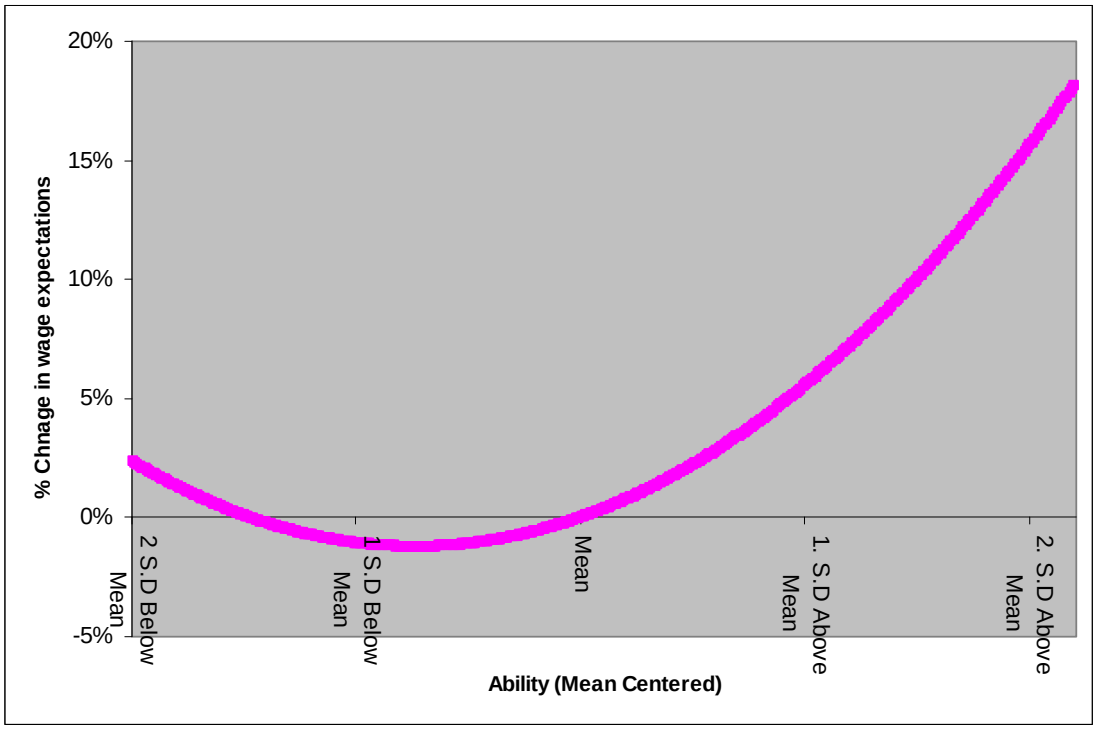
The near 7% premium initially identified in specification 2 is almost entirely explained by differences in ability. Differences in wage expectations between students at post 1992 and Russell Group universities are now statistically insignificant. However, students at Other Pre 1992 universities now expect less than both those at Russell Group and Post 1992 institutions. Moreover the model selection process suggests that the ability variable should not enter as a linear term to the model, but rather as a quadratic. Specification 4 presents this version of the model, with results also presented in Table 7. Once included, the variable reflecting differences between the three institutions has now reached statistical insignificance. It appears that students at the prestigious Russell Group do not expect a wage premium above and beyond their ability, despite Chevalier showing this to be the case in the graduate labour market. Chevalier and Conlon suggest that the "Russell Group premium" provides institutions an economic argument for being allowed to charge higher tuition fees. However it seems this premium is not present in students' expectations. Therefore it must be brought into question whether students will actually be willing to pay these higher fees, when it is unclear whether they anticipate any economic benefit.

An additional point of interest is that the role of ability in wage expectations may well be more complex than the current literature suspects. There is much discussion in the schooling behaviour literature on what variables students' condition their beliefs upon, and especially the role of ability. Freeman (1971) assumes youth do not condition their expectations on ability, while Manski (1993) shows how schooling choices critically depend on how youth use ability in forming their expectations. Several other studies make varying assumptions, though very little research exists focusing on the relationship between ability and wage expectations.

Figure 2 illustrates the consequences of the quadratic ability term in specification 4. Those with imputed UCAS score one (91 points) and two (182 points) standard deviations below the mean (290 points) have very similar wage expectations to those of average ability. However those with ability one standard deviation above the mean expect a salary 5% higher than someone of average ability. Similarly those two standard deviations[33] above the mean expect around 16% more than someone of average ability.

[33] Around 3.5% of observations are two standard deviations above the mean, while 1.5% are two standard deviations below. The distribution of UCAS score has moderate positive skew.

Figure 2. The Relationship between Ability (UCAS score) and Wage Expectations



Notes: Diagram produced using specification 4 in table 7, holding all other variables constant at their reference value

This seems to suggest that low ability students expect the same wages as those of average ability[34], while high quality students expect a premium. It appears that low ability students do not expect to be punished in the labour market for their lower grades, while those of high ability expect to be rewarded for theirs. Government schemes aimed at widening access could also be something to do with this result, highlighting the average graduate wage to groups on the margin of going to university, typically of lower quality, who then internalise these figures in their expectations. Hence low ability students may anchor their expectations to the wages of average graduates, while high ability students, realising they are of superior quality, use the average graduate wage as a lower bound. Of course caution must be exercised when interpreting this result, due to the ability variable being an imputed proxy[35], even though it is statistically significant at the 1% level. These results suggest that the relationship between ability and expectations may be quite complex. Simplistic assumptions of a linear relationship may not necessarily hold true, indicating the need for future research.

[34] Further analysis, not presented, looked at ability in quintiles. The results suggested there was no statistically significant difference in wage expectations between individuals in the bottom and middle quintiles, though there was between those in the middle and the top.

[35] Indeed, a fixed affects regression model is also ran containing a dummy variable for each university. When this is done, the ability term becomes insignificant. However this is likely to reflect that we have used institution in part to form the proxy for ability. Thus the imputed UCAS score is likely to be highly collinear with the institution dummies, explaining its insignificance in this specification.

ii. Parental Income and Ethnicity [36]

Various background characteristics have been investigated in previous research, with particular interest surrounding the socio-economic status of the student. It is hypothesised that students' parents are a critical source of labour market information, and that students' wage expectations will increase with family income.

On the other hand, in the current literature there is little consideration of differences in wage expectations between ethnic groups. They may, for example, expect to suffer some form of discrimination in the labour market and thus, ceteris paribus, have lower expectations. Alternatively, much American research, such as Rouse (2004), notes that ethnic minorities tend to be overly ambitious. Is this the same in Europe and the UK?

The results are consistent across all specifications and support the hypothesis that students from richer backgrounds have higher wage expectations. All groups presented expect significantly more than those with parents earning below £20,000. There is also a monotonic trend; the higher the income group, the greater the wage expectation. [37] It could be that students from a rich background expect this high salary in order to maintain a high standard of living. Alternatively as university participation rates have increased over time, students may use their parents' income as a lower bound, and expect a higher salary due to their better education. Another possibility is that students may think their parents have connections in the labour market that will help secure them a lucrative job. However this variable could also be reflecting unobserved factors, such as parents' influence on intelligence and work ethic, which are also correlated with wage expectations.

The second hypothesis, that students from a minority background may expect to suffer some form of discrimination in the labour market, does not seem to hold. Results suggest that Black and Asian students expect a significantly higher starting wage than white students.[38] This conclusion is robust to all the specifications used. Hence it seems that ethnic minorities do not expect to suffer discrimination in the graduate labour market. Indeed quite the opposite appears to be true, these groups tend to be more optimistic about their future earning potential than their white peers.

---

[36] It should be remembered that when interpreting the results for parents' income, the measure can only be considered as a proxy, with further details in appendix 2. Also data is only available for students classed as "Dependant" (59% of the total).

[37] Out of the 1838 dependant students, 164 (9%) did not report a figure for parents income. These students had higher wage expectations than the highest category included in the regression (those with parents earning over £40,000). A logistic regression, not presented, was carried out to investigate if certain groups are more likely to not report a figure for parents' income. The results of the logistic regression suggested that students who have parents generating most of their income from pensions or investments are less likely to report a figure. To the extent that this exhibits wealth, for instance early retirement or being able to live off investments without working, the result fits quite well with the observed pattern; the better off the students' parents, the higher their wage expectations.

[38] The Black and Asian groups were combined due to small sample sizes within each. In initial regressions the two groups were entered separately, producing similar coefficient estimates and standard errors, significant at the 10% level. A test was performed of whether Black and Asian have equal regression coefficients, resulting in the null hypothesis not being rejected.

6. Data on Realised Wages

The preceding analysis illustrates that there are quite large differences in students' wage expectations. This may simply reflect the different labour market opportunities that students face. An alternative hypothesis is that students have unrealistic expectations and significantly overestimate starting wages. Manski (2004) describes various ways in which expectations and realisations can be compared, with longitudinal data the most direct method of comparison. Unfortunately, there is almost no suitable panel data available across Europe[39] and none in the UK. Thus attention is turned to Manski's second method; using repeated cross sectional surveys to evaluate average expectations and realisations. Therefore a second data source is needed that contains information on wages in the graduate labour market for the same cohort of students. Data on graduate wages, corresponding to the same year, is drawn from the Higher Education Statistics Authority (HESA) Destination of Leavers Survey (DLHE). This section describes the additional data source and methods for comparing the two surveys.

The DLHE is an attempted census of all 2004-2005 graduates' employment circumstances, including their current salary, six months after completing university. Students are contacted directly by the institution they studied at by postal questionnaire, with non-respondents followed up in a telephone interview. This results in a survey response rate of around 80%. The results are then linked with administrative data about the student collected by HESA, providing a rich source for analysis. Variables within the dataset include socio-economic status, university entrance (UCAS) score, degree classification, subject of study and where the student lived while at university.

The target of this survey is obviously a great deal wider than that of the SIES. Several sample selection procedures, available in Table 8, were applied to the data to ensure the two sources were comparable. The table also illustrates the sensitivity of average wages to each of the sample selection procedures.

Table 8

As in the SIES, only students who attended a university in England and Wales were considered. The data was also restricted to only those students who had finished their first undergraduate degree; with those completing postgraduate study, or university courses below the level of bachelors degree, excluded.[40] It is assumed that respondents to the SIES were reporting their first wage expected after their undergraduate degree. However it is impossible to rule out the possibility that some students reported their wage expectation under the assumption that they were going to continue in full-time education and gain a post graduate qualification. Again further analysis suggested the impact of this assumption on results is small, and does not

[39] Even the one longitudinal study by Webbink and Hartog suffers quite substantial methodological problems. Their study suffers both high non-response to the survey and missing data on the wage variables. Moreover they show that the missing wage data is not at random, it comes from groups likely to have lower wages in the graduate labour market. Later in this section, the difficulty of non-response in graduate surveys is shown, and illustrates that missing data is likely to come from groups who are likely to be earning a lower wage.

[40] In the UK this includes HND, HNC, foundation and access courses, amongst others. These are generally thought of as a qualification below degree level, and have therefore been excluded.

affect the substantial result. Only students who reported salaries of £8,000 or more and were working full-time are included[41]. Since students were asked for their full-time annual equivalent wage, £8,000 was identified as the lower bound for logical responses due to minimum wage laws in the UK.[42] The same rule is also applied to the SIES data, so that only expected and actual wages above £8,000 are considered. Moreover, to limit the potential influence of previous labour market experience, and to target the particular group of interest, in both surveys the sample was restricted to those below the age of 25. This has also been done so that an additional data source, the labour force survey, can be used as a check for robustness of results[43].

The DLHE has many features that make it a strong candidate to compare with the SIES data. The information was collected in January 2006 and specifically refers to graduates who were final year students in 2004-2005 when the SIES was conducted. However in this analysis, to ensure a reasonable sample size, wage expectations from students in all year groups are used.[44] It is again assumed that students in all years report their expectations in current (2005[45]) prices[46].

The questions posed in each survey also relate closely to one another.

SIES
"What sort of salary do you expect to be earning in the first job you take once you have graduated?"
Interviewer comments: If not sure of the exact amount, please give your best estimate.

DLHE
"What was your annual pay to the nearest thousand £, before tax?"
If you were employed less than a year or were part-time, please estimate your pay to the full-time annual equivalent.

The SIES asks about salary expectations in students' first job after graduation and the DLHE records information on salary sixth months after finishing university. In the vast majority of cases, the difference between these two definitions is likely to be

[41] Dominitz &Manski (1996) shows that students tend to report their expectations conditional on working full-time.
[42] The adult minimum wage at this time was £5.05. Assuming the minimum amount of time required in a full-time job is 30 hours per week, for 52 weeks a year, this generates a full-time annual income of around £7,900. Only around 1% of observations were dropped from the DLHE using this sample selection rule, with 2% dropped from the SIES.
[43] The wording used in the labour force survey means only those under 25 have been considered. This is discussed later in this section.
[44] Indeed it may be argued that we are not using repeated cross sectional surveys in the strictest definition. In part, we are assuming that all students in the SIES face the same distribution of actual wages, represented by the salary recorded for the graduating 2004/05 cohort. However it seems highly likely the distribution of graduate wages will remain stable considering the short space of time. Manski (2004) discusses this assumption in more detail.
[45] Since the SIES was conducted in early 2005, while the DLHE recorded actual wages in 2006, an adjustment has been made for inflation. The wages in the DLHE were scaled back to 2005 prices using the Retail Price Index (2.8% for the year in question).
[46] Ideally, students would have been formally instructed not to consider inflation in the wording of questions, as in Dominitz & Manski (1996). They report that students generally adhere to this, and do not consider inflation in their wage expectations. Moreover Brunello (2001) use similar wording to the SIES, in that students are not directly informed how to deal with inflation. They also assume students report their expectations in current prices, and find inconsistencies in their data with the idea that respondents may inflate their expectations to try and account for inflation.

minimal. Previously it was stated that students in the SIES are thought to provide estimates of a gross, annual salary. The DLHE survey asks students to provide an estimate for their full-time equivalent annual wage before tax, providing a closely matched definition. A final issue is that the DLHE survey asks students for their wage to the nearest thousand[47], while expectations in the SIES were recorded in an open text cell. However section 3 described how students' expectations tend to bunch around the nearest thousand, meaning this is unlikely to induce any substantial bias.

However, for all the benefits of the DLHE, there are some difficulties with response rates and the selectivity of respondents. Although the DLHE is an attempted census of graduates, there is quite a large degree of non-response to the question about salary[48]. Moreover, the salaries for those students who go into postgraduate courses, or those that are unemployed, also go unobserved. To illustrate the potential difficulty this may cause, after the sample selection procedures have been applied, 145,517 observations remain. Of these only 45,906 (32%) individuals are in the labour market and report their salary. Missing data is either due to self-selection out of the labour market, not responding to the survey or missing salary information. As a result, there may be a selectivity problem when comparing the two surveys. Differences recorded could be a result of who is responding to each of the surveys, rather than actual differences between students' expectations and realisations. A further issue maybe that certain groups have higher drop out rates than others, leading to different proportions recorded in each survey. For instance students from lower socioeconomic backgrounds may be more likely to drop out of university. Therefore one would observe a higher proportion of this group in the SIES, with data recorded during university, than the DLHE, with data recorded for graduates.

Perhaps the greatest worry is that the missing salary data in the DLHE comes disproportionately from groups who are likely to be earning a particularly high wage. This would mean the recorded wages are lower than the "true" graduate starting salary and thus artificially support the hypothesis that students have unrealistically high wage expectations. To investigate this further, a probit regression for the item non-response to the salary variable has been conducted[49]. Results can be found in appendix 4, while Table 9 illustrates the probability of certain hypothetical groups responding to the question regarding salary.

Table 9

[47] The mid-point of this band has been used for all subsequent analysis. The top band in the DLHE was £50,000 and above. However, after the sample selection rules in table 8 have been applied, only 0.1% of observations were in this category, with negligible impact on results.

[48] For the sample selection criteria described above, 87,327 individuals are in full-time work and have responded to the survey. Of these, 45,906 (53%) report a salary.

[49] Here only item non-response is considered and not unit non-response to the entire questionnaire, though it should be remembered the overall response rate to the questionnaire is around 75%. Moreover there is likely to be selectivity into the labour market present, for example those who self-select into unemployment or further full-time study will have their salary automatically unobserved.

Person A has the characteristics of someone with excellent labour market prospects and thus expected to be a high wage earner. They have a good degree from a top university and are now in a graduate job. This individual has a 76% chance of responding to the salary question in the DLHE. Individuals B and C illustrate the impact of certain characteristics on salary response rates. Person B only differs from A in that he is in a non graduate job, yet there is only a 62% chance of him responding. If we add further characteristics that are likely to mean lower wages in the labour market, such as gaining a 2.2 in Art and now being in a admin job, the probability of responding drops to 45%. Person D is someone who is likely to have a particularly difficult time in the graduate labour market. Her probability of responding is just 31%, compared to 76% for someone who is likely to be a particularly high earner. This clearly illustrates that, if anything, the DLHE is likely to provide an **upwardly** bias estimate of the true graduate wage. To investigate this further, response weights were created by calculating the inverse of the predicted probability of students responding to the salary question. These were applied to the individuals who reported their salary[50]. The weights lowered the estimated average graduate wage, for full-time students, from £16,455, to £15,996, a drop of 3%[51].

The selected SIES and DLHE samples, for those who reported wages, were then compared in terms of characteristics that could be observed within both populations. The results are shown in Table 10, both with and without the response weights for the two surveys applied[52].

Table 10[53]

There are some differences in the social class and gender composition of the two surveys, though the use of the SIES weights (see section 3) alters the proportions of the latter quite significantly. For social class, there is a problem with missing data in the DLHE, which is likely to be causing some mismatch between the two surveys. However, even with this difficulty, the differences in observable characteristics are not particularly big, although statistically significant due to the large sample size, and are unlikely to be causing any substantial bias. To check robustness, estimates are presented both with and without the sampling weights, to analyse the sensitivity of results.

One further issue is that although the populations appear broadly similar in terms of characteristics that are observable in **both** surveys, there may still be differences in characteristics that are unobservable. These characteristics could go unmeasured in either, or both, of the surveys. As a further check for robustness, the UK Labour Force Survey will be used as an alternative source of information on graduate wages.

[50] Response weights have also been calculated for part-time students.
[51] It should be noted this is significantly lower than the official figure HESA publishes. Appendix 5 fully documents why this is the case, and how the figures presented here are calculated.
[52] The SIES response weight applied refers to those described earlier in section 3, that were contained in the dataset. The DLHE weight refers to those I have created via the probit model just described.
[53] Initial analysis, not presented, suggested a difference in the proportion of medical students contained in the two surveys (8% of the SIES compared to 2% of the DLHE). Thus medical students will be excluded in many parts of the analysis, due to the difference between the two surveys and the quite different labour market these individuals face.

7. Comparison of Expected and Actual Wages

Two methods are used to compare students' average wage expectations with the average actual wages observed. Firstly the ratio of the expected mean (median) wage is compared to the mean (median) of the actual wage. The second method is to graphically represent the distributions, via kernel density estimates, of the actual and expected wage to identify differences[54].

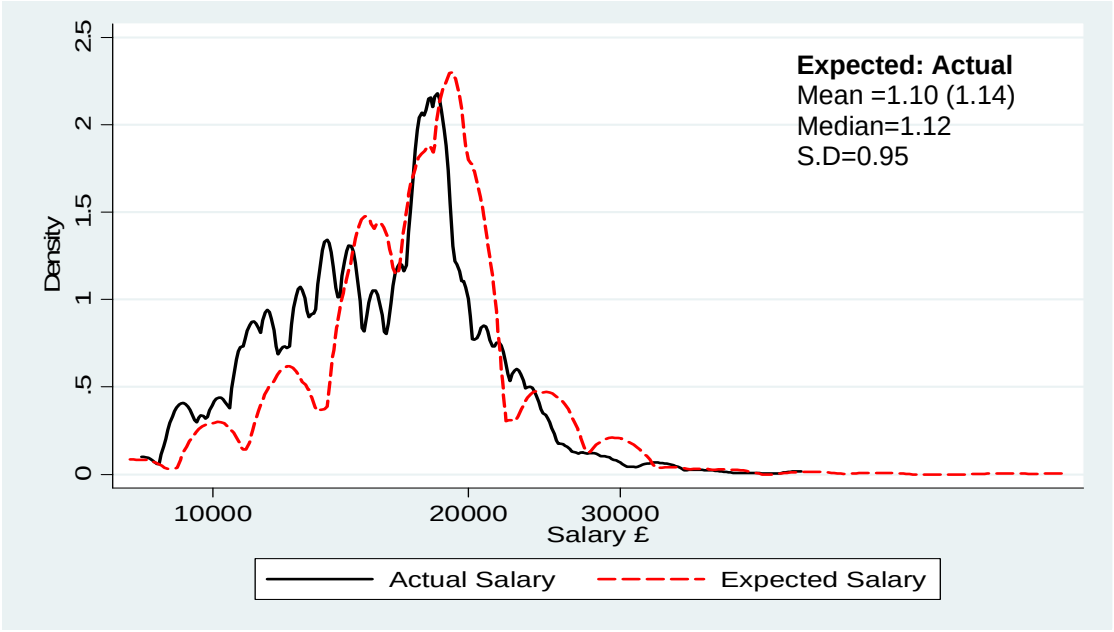
Do students have realistic expectations?

The initial hypothesis is whether students overestimate wages in the graduate labour market. The population was divided into full and part time students, with summary statistics and kernel density estimates reported below.

[54] The kernel density estimates have used the default smoothing applied by Stata.

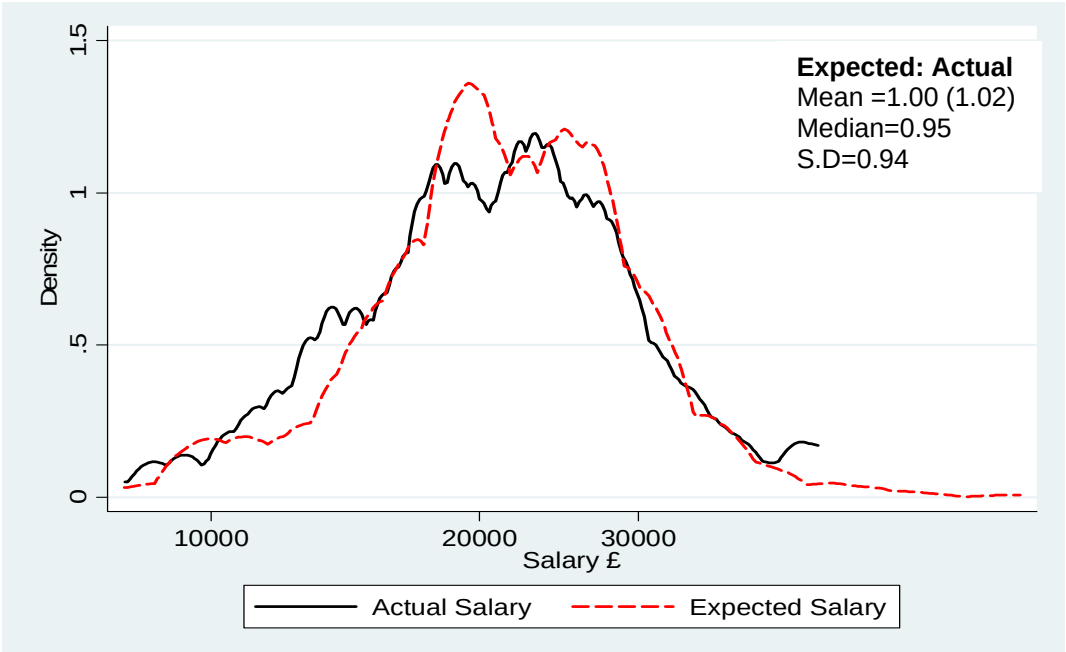
Figure 2. Kernel density estimates of actual versus expected log wages[55]

Full-time Students



Notes: 1 Kernels and ratios are estimated with unweighted data
2 Figure in brackets denotes when both SIES and DLHE weights are applied to data

Part-time Students[56]



[55] Medical students have been excluded, due to the differing proportion of these groups in the two surveys and the large influence they have on both expected and actual wages. Analysis was also conducted containing medical students, with very similar results.
[56] For part-time students, the sample selection rules differ slightly, to ensure adequate sample sizes. In particular, students of all ages and those doing "other" types of university courses, though not higher degrees, are included. Table 8b gives further details.

The ratio of the mean (median) expected salary to actual salary for part-time students is 1 (0.95), suggesting that, on average, part-time students have realistic expectations. Furthermore, the kernel density estimate illustrates how closely the distributions of part-time students' expectations and actual wages match. However a different conclusion is reached when looking at the results for full-time students. The kernel density estimate of students' expected wage appears to be to the right of the estimate for actual wages, especially near the lower tail, suggesting overestimation by students. This overestimation is also illustrated by the ratio of average expected to average actual salary in Table 11. Expectations are around 10% (£1,600) higher than average wages in the graduate labour market, using the means, and 12% (£1,900) higher using the medians. However when using the sampling weights to correct for non-response, the overestimation appears even higher at around 14% (£2,200). This means that students on average, after tax, overestimate starting wages by over half the yearly fee now charged in the UK for university tuition.[57] Moreover Table 8 suggests this result is robust to the sample selection rules applied. Thus there seems sufficient evidence to support the hypothesis that, on average, full-time students, as a group, have unrealistic expectations of future wages.

As a check for robustness, the Labour Force Survey (LFS) has been used as an alternative data source for wages of recently qualified graduates. Data comes from 10 quarterly surveys, running from September 2005 to March 2008. These dates were chosen as they relate to when the students covered in the SIES would have graduated and entered the job market. Moreover, from the September to December 2005 survey onwards, respondents were asked the question:

"Which, if any of these qualifications did you gain in the last 12 months?"

"Degree level qualification including foundation degrees, graduate **membership** of a professional institute, PGCE, or higher"

Notice that this includes higher degrees and membership of a professional institute, for instance gaining a chartered accountancy qualification. I assume those under 25 are likely to only have a degree as their highest qualification level. Restricting the sample to this group also limits the role of previous labour market experience in unduly influencing expected or actual wages. These restrictions mean the reported wage will almost certainly relate to respondents' first jobs after gaining a degree, providing a good match for the wage expectation question in the SIES. Other restrictions in the LFS include only looking at those working full-time, earning over £8,000 a year and hold a first degree as their highest qualification. Gross weekly wages are used from respondents in their fifth, and final, wave of the survey, and scaled up to the annual equivalent.[58]For those surveys conducted between 2006 and 2008, wages have been deflated to 2005 prices using the Retail Price Index, under the previously stated assumption that students report expectations in today's (2005)

[57] These particular students would have actually paid an upfront tuition fee of around £1,200 per year. Tuition fees changed for students starting after 2005 to a maximum of £3,000 per year, payable after graduation.

[58] Respondents to the LFS are asked for their wages in the first and fifth wave. However since we are using 10 consecutive waves, there would be a problem of double counting people if wage data was taken from both wave 1 and 5. For instance, someone who was in wave 1 during September-December 2005 would be in wave 5 during September-December 2006, and hence have their wages recorded twice. Hence only wages in wave 5 are used.

prices[59]. The final sample size is 194 observations.[60] Results are presented in Table 11.

Table 11

The estimate for mean starting wages from the LFS is £16,073 (standard error £370), very close to that recorded in the DLHE (£16,455), particularly after the DLHE has been weighted for non-response (£15,996). While the ratio of the means, at 1.13, is only slightly larger than when using the DLHE, the ratio of the medians is dramatically different at 1.19[61]. Moreover the difference between the average SIES expected salary and the average LFS wage is statistically significant, even with the limited sample size. Hence it appears that the preceding results are indeed robust to the data source used.

Does the realism of students' wage expectations vary by background characteristics?

In the preceding section, wage expectations were found to vary between groups based on several different background characteristics. An interesting question is whether students who expect higher wages actually secure this premium in the labour market, or are they, on average, more unrealistic?[62] This proves challenging methodologically without longitudinal data. The samples selected in the SIES and DLHE can be restricted further, for example to look at men and women separately, though this can obviously only be done for characteristics observable in both surveys[63]. Moreover there is likely to be further compositional issues, similar to those discussed in Table 10, particularly with the reduced sample sizes. This, coupled with the non-response in the DLHE, limits our ability for a more in-depth analysis. Nethertheless Table 12 provides results, though these should largely be treated as indicative rather than definitive.

Table 12

Analysis in section 5 showed that junior students have higher wage expectations than those about to graduate. Assuming that all students report their expected salary in

[59] Even without this assumption, the ratio of expected to actual median wage is 1.15 using the labour force survey.

[60] Two outlying observations have been dropped due to their large influence on the mean wage in this small dataset. These individuals had wages over 2 times bigger (over £80,000) than the next largest observation (£40,000), and were over 6 standard deviations higher than the mean. Robustness was checked by including these two observations, with all the substantial results remaining intact (see footnote 60)

[61] Note with the outliers the mean salary is £16,749, and the median £15,163. With the outliers trimmed to £50,000 the salary is £16,420 and mean ratio 1.10, while the median remains at 1.19. Even with including these trimmed outliers, there is still a statistically significant difference between the average expected and average LFS salary.

[62] One possible reason why the average expectation and realisation may differ is the sampling variation in the two data sources. To investigate this, the 95% confidence interval for students' expected wages was compared to the DLHE estimate of the actual wage. Since the DLHE is an attempted census, with a very large number of responses, the sampling variability is so small that this is assumed to be the true population wage. In all cases, the actual wage was outside the confidence interval for the expected wage, suggesting the difference observed can not be attributed to sampling variability.

[63] Hence factors such as parents' income can not be explored.

2005 wages[64], there is also evidence that junior students are more unrealistic. Whereas final year students tend to overestimate their starting salary by only 6%, on average, those who have just entered university overestimate by around 14%. An important implication is that students who have just made the decision to invest in university education have especially inflated expectations. Human capital models of school enrolment rely on the assumption that students are able to accurately assess their future income streams under alternative investment decisions. However this analysis indicates that students are not even particularly good at making the initial assessment of their starting salary. It was also shown that students at Russell Group universities expect higher wages than those at other pre 1992, and post 1992, institutions. There is little evidence that students at Russell Group universities are any less realistic than those at post 1992 institutions, who overestimate on average by 14% and 11% respectively. However students at other pre 1992 universities appear to be much more accurate, overestimating by only 6%, on average. Other points to note is that there is no evidence to suggest differences based on gender or ethnicity, with overestimation by each group, on average, of around 11%.

Does the accuracy of students' wage expectations depend on the subject they study?

Though full-time students seem to be too optimistic on average in their wage expectations, students in some subjects may have access to better information on wages, and thus make better predictions. It is hypothesised that students who are studying a subject leading to a particular career will be more realistic, as they are likely to research specific jobs and have better knowledge of the labour market they face. Alternatively students who take language and art based courses are likely to enter a far more general labour market, with less certainty about their future career prospects. Results are provided in Table 13[65], with kernel density estimates for various subjects shown in Figure 3.

Table 13

[64] Of course this assumption plays an especially prominent role here. One could argue that if students do take into account inflation and economic growth into their estimates, then this would explain much of the difference between year groups.

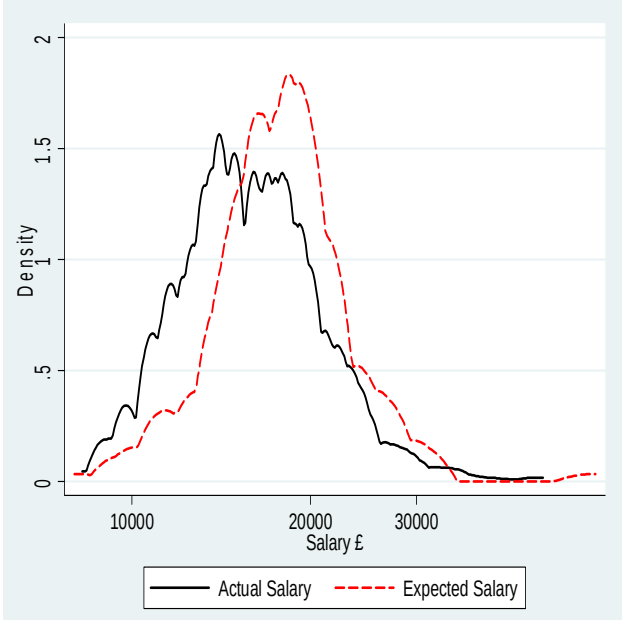
[65] Again, there maybe some compositional issues surrounding the two surveys that may hamper comparison. Caution is thus advised when interpreting the results, though I believe the general pattern to hold.

Figure 3. Unsmoothed Kernel Density Estimates

Engineering, Computer Science, Maths

Education

Business and Admin



History, English and languages

Notes: 1 Kernels and ratios are estimated with unweighted data

The results show that on average there is overestimation in all subjects, except Medicine, where there is quite a large degree of underestimation. Law also appears as an outlier, with especially large overestimation.[66] One possibility may be that students wishing to enter a legal profession have to complete additional study at a Law school after university. Therefore the vast majority will not be in the labour market yet and their wages unobserved; there is the problem of selectivity. Those that have entered work straight from their course are likely to be in a much less lucrative position than Law graduates who continue their legal training and later enter the profession[67].

As hypothesised, students who are studying a subject that directly leads to a career have more realistic wage expectations. Medicine and Education both appear at the top of the table, and suggest these groups on average are very close, or even underestimate, their future wage. In contrast to the hypothesis, students in subjects Allied to Medicine appear no better than the average. However during the period studied, the number of positions available for newly qualified physiotherapists and nurses was particularly low. Thus these students, expecting to enter these occupations, were likely to have struggled to find jobs in their desired fields and thus had to compete in the more general graduate labour market.

The kernel density estimates in the upper panel of figure 3 illustrates how closely the distributions for actual and expected salaries match for Education and Engineering, Maths and Computer Science. A comparison to the distributions for Business, and History and Languages, clearly illustrates the superior estimates made by those entering careers in Education. Interesting patterns also occur between the subjects that lead to the more general labour market. Those subjects where the errors seem to be largest generally involve language, rather than technical, skills. For example, the Social Sciences, Languages and Business courses mainly involve writing essays, while subjects teaching more mathematical skills, such as Computer Science and Engineering, appear to contain more realistic students. Science can be identified as an exception to the discussion above; it is a course that develops technical skills, but whose students appear unrealistic. This could reflect students believing that they will receive a high wage because they have chosen to take on a traditionally challenging subject. In reality however, the technical skills they have built up may be required by relatively few employers, or they end up in a field unrelated to their degree. This may mean they suffer an unanticipated need to retrain in an unrelated discipline and have to accept a lower starting salary, or perhaps move into a non-graduate job.

One may be tempted to suggest that the very different gender composition of subjects is leading to these differences in overestimation. Table 14 provides the ratio of average expected to actual wages broken down by subject and gender.

Table 14

Again, due to the small sample sizes in particular subjects, these results are indicative rather than definitive. Nethertheless the results do seem to suggest similar patterns for men and women. Both, on average, underestimate in Medicine, while overestimating

[66] This could simply be reflecting sampling error, due to the reasonably small sample size. Expected salaries are significantly different from actual wages in all cases, except for education. This suggests sampling variability is not responsible for the observed differences.
[67] Excluding this group from the analysis for "All full-time students" does not change the results substantially.

in the Sciences, Business and Language based subjects. In many instances, when the ratios do differ for men and women, this can be explained by the small sample size. This is the case for Education and subjects Allied to Medicine, particularly for men. On the other hand, there is a tentative suggestion that men and women, on average, have less accurate expectations when they are studying a subject traditionally associated with the other gender. For instance, men appear to overestimate, on average, more than women in the Arts and Languages. Meanwhile women make worse estimates in the sciences. However, our sample has only limited capacity to explore this possibility gender minority bias, requiring further evidence to explore substantiate whether this is truly the case.

Discussion

These findings have significant implications for both academic models of human capital and higher education policy across Europe. Under a human capital model, students enrol in education where the perceived benefits, both financial and non-financial, are greater than the costs. However this research has been limited to one estimate made by students relating to one specific time in the future. How much we can say about the realism of students about their future earnings over the course of their whole lifetime is quite restricted. Nor do we have students' views of the counterfactual; what they would expect to earn had they not gone to university. This study has also only considered the financial aspects of studying for a degree, and not the non-monetary benefits of university to both the individual and society. Nethertheless, the findings suggest that students tend to overestimate their first wage after university. In a human capital model, as described above, starting wages carry more weight than those at later time points, due to what may be fairly heavy discounting of future income[68]. The fact starting wages carry quite substantial importance means we must at least question whether students are realistic about the value of holding a degree. Moreover, economic models that assume students have perfect foresight, and therefore know with certainty what they will earn on graduation, also appear to based on a rocky foundation.

Of course, it is still possible for university to be a good investment, even under such conditions. Many individuals will still find university both a financially and culturally profitable experience, even if it is not to the level they once expected. Nethertheless it is equally plausible that by overestimating future wages, some students may mistakenly choose to go to university, who will not receive the benefit they expect on enrolment. The UK Class of 99' report by Purcell, Elias, Davies and Wilton (2005) illustrates such feelings in qualitative research, as shown below:

> 'I would have still ended up in the position I'm in now if I would have carried on working full-time…. I applied for over two hundred jobs, I felt this degree was a total waste of time; I was a self-funding student, which was a waste of money. I'm still paying for it now, I'm a single parent and to be honest it was the biggest waste of time and money that I've ever spent'.

[68] The survey that Wolter (2002) uses contains a question that attempts to elicit individuals time preferences. Using this method, he finds students discount future wages by between 7% and 8% per annum. He notes that this is well below some other studies, such as Osterbeek and Ophem (2000) who suggests the average discount rate is 19%. Both of these cases illustrates how students may have high discount rates, and therefore put a particularly high value on their starting salaries.

> ……..everyone tells you if you do a degree the world will be your oyster, you'll earn loads of money. No'. Page 194

Other aspects of the research may also have importance for higher education policy. Students build up debt while at university, when income is low, and expect to pay this back when they have a job after graduation. Gustman and Stafford (1972) also show that students with higher wage expectations tend to consume more at university. From an economic point of view, students are using credit markets, in part, to smooth their consumption over time. However, if students overestimate their future wage, they may also be overestimating their ability to pay back the money they borrowed. This may lead to students taking on too higher levels of debt that they later struggle to repay, due to the fact they are not in as well a paying job as expected. It may also mean they are willing to take on debt to pay for high tuition fees when entering university, but regret this decision in hindsight when paying back the money is harder than they once expected. This obviously has direct implications for the provision of credit for students, with loans for fees and maintenance provided by governments and banks. An oversupply of easy, cheap money could well lead some individuals into difficulties later in life; greater caution is needed on the level of debt incurred. Any higher education system in Europe looking to expand by offering students cheap loans should consider this before pursuing such a policy.

Another important issue is how this relates to widening access schemes proposed by European governments, and in particular the UK target of getting half of all school leavers to experience higher education. The benefits of university are widely promoted by governments, and in particular career prospects, to encourage individuals to continue their education. However this practise may just enhance students' unrealistic expectations, which certainly seems to be the case in the quote above. Another important point is whether students are being given accurate information about salaries and employment prospects from the various available sources. For instance the Association of Graduate Recruiters conducts a bi-annual survey, which states that the median starting salary of graduates in 2005 was £22,000[69]. This is well above the population average figure for graduates, because it only looks at certain jobs with large UK employers. Moreover this research even points to non-response bias that is likely to be present in the DLHE, often quoted by the media and government ministers as the "true" graduate wage. This information may well have an impact on students' views of the graduate wage and inflate their expectations to unrealistic levels. More information for students on the distribution of starting salaries, and their likely place on the scale, is required to make sure that individuals are making well informed decisions when continuing their university education.

[69] See http://www.agr.org.uk/news/agr_in_the_news/id.31.html

8. Conclusions

This paper set out to explore heterogeneity in UK students' wage expectations and to identify whether they held "realistic" views of the graduate labour market. In doing so, this provides the first study in Europe to explore wage expectations using a nationally representative sample of students. The results highlight:

- Wage expectations vary significantly based on how far the student has progressed through university, though this is largely explained by their differing views on their employability.

- Students' idiosyncratic characteristics play an important role in determining their wage expectations.

- While the quality of the university does not influence wage expectations, previous academic performance does.

- While part-time students seem realistic in their wage expectations, those studying full-time tend, on average, to overestimate by around 10%.

In particular, this paper shows how, as students' progress through university, their views on life as a graduate change. Final year students are less optimistic about their ability to land a "career job" and their starting salary. This shows how students learn about their own ability and the labour market through their time in higher education, and that prospects may not be as bright as they once expected.

There appears to be a significant difference based on wealth, with students from high income backgrounds expecting a greater salary than their low income peers. However the initial hypothesis that students from ethnic minority backgrounds may expect some form of discrimination in the labour market, and therefore estimate a lower wage, is rejected. Ethnic minorities actually expect a higher salary than their white peers.

The second half of the paper investigated whether students' wage expectations are realistic. This is the first European study to compare wage expectations from a nationally representative sample to the actual wages earned by graduates from the same cohort. The evidence suggests that full-time students overestimate wages in the graduate labour market. However some groups are particularly prone to overestimate their starting wage, such as those at the beginning of their course and those at a post 1992 institution. Moreover students studying a subject leading to a particular career expect salaries reasonably close to the observed data structure. On the other hand the expectations of those in Science, Language and Humanities courses are quite a way off the reality of the graduate labour market.

Due to the lack of current investigation, and the importance to several higher education policies across Europe, there is potential for a much wider scope of research. Longitudinal data, following students for a period of several years until they are well established in the labour market, would further the current research. Recording students' perceptions of wages with and without a degree would also provide an interesting insight into student enrolments and decision making.

This paper has provided an initial investigation into how students view their financial futures after university, and if they are expecting too much from the labour market in obtaining a higher education qualification. However a much greater depth of research is needed to understand the financial nature of students' decision making and the possible consequences for European higher education policy.

Tables

Table 1. Sample selection rules

Rule	Sample
Initial sample	3,548
Salary expectations missing	3,375
Further Education colleges dropped	3,170
"Other" degrees dropped	2,791
Expectations below £3,000 dropped	2,744

Table 2-Peaks in the distribution of expected wage

Expected Salary £000	Frequency	%
8	22	1.4
10	75	2.7
12	129	4.7
13	43	1.6
14	68	2.5
15	304	11.1
16	115	4.2
17	135	4.9
18	364	13.3
19	92	3.4
20	473	17.2
21	37	1.4
22	70	2.6
23	35	1.3
24	40	1.5
25	202	7.4
26	30	1.1
27	26	1
28	49	1.8
29	10	0.4
30	126	4.6
35	28	1
40	16	0.6
50	8	0.3

Notes 1 Data does not add up to 2,744 due to some students not predicting salaries at these round intervals

Table 3. Exploratory statistics

Category	N (Unweighted)	P90 (£000)	P10 (£000)	Mean (£000)	Median (£000)	Standard Deviation (£000)
Proximity to Graduation						
Final Year	946	25.0	12.0	18.1	18.0	6.4
1 Year	758	25.0	12.5	18.6	18.0	6.2
2 or More Years	1,040	26.0	13.0	19.4	19.0	7.7
Gender						
Male	896	27.0	13.0	19.5	19.0	7.5
Female	1,847	25.0	12.0	18.1	18.0	6.5
University Type						
Russell	678	30.0	13.0	20.0	20.0	7.3
Other Pre 1992	502	25.0	12.0	18.0	18.0	7.1
Post 1992	1,564	25.0	12.0	18.4	18.0	6.5
Parents Income (£ per annum)						
Below 20,000	356	23.0	12.0	17.0	17.0	5.4
20,001-40,000	559	23.0	12.0	18.0	18.0	6.1
40,001+	635	25.0	12.0	18.8	18.0	6.3
Dependant student/ No data[1]	1,165	29.5	15.0	20.0	19.5	7.8
Ethnic Group						
White	2,343	25.0	12.0	18.5	18.0	6.9
Asian	134	28.0	12.0	19.9	20.0	6.5
Black	110	30.0	14.0	20.9	20.0	7.4
Mixed	155	30.0	14.0	20.0	19.0	6.2
All Groups	**2,744**	**28.0**	**12.5**	**19.6**	**18.5**	**6.5**

Notes: 1 See appendix 1

2 Data does not add up to 2,744 in all cases due to missing data

3 The sample size given is unweighted. The summary statistics are provided after using the SIES sample weights

Table 4a-Regression results.

Variable	Specification 1 Coefficient	Specification 1 Standard Error	Specification 2 Coefficient	Specification 2 Standard Error	Fixed Affects (Specification 2) Coefficient	Fixed Affects (Specification 2) Standard Error
Future Plans (Ref: Career Job Only)						
Temporary Job Only	-	-	-0.224*	0.026	-0.219*	0.025
Either a Career or Temporary Job	-	-	-0.126*	0.033	-0.116*	0.031
Further Study or Travel	-	-	-0.007	0.013	-0.006	0.013
Hard to get Grad Job (Ref: Agree)						
Neutral	-	-	0.013	0.018	0.015	0.017
Disagree	-	-	0.056*	0.014	0.054*	0.013
Missing	-	-	0.078*	0.031	0.069*	0.031
Proximity to Graduation (Ref: Final Year)						
1 Year	0.030*	0.014	0.013	0.014	0.009	0.014
2 or More Years	0.066*	0.017	0.034*	0.017	0.031	0.017
University Type (Ref: Post 1992)						
Other Pre 1992	-0.006	0.019	-0.003	0.017	-0.066	0.124
Russell Group	0.074*	0.02	0.065*	0.019	0.111	0.080
Parents Earnings (Ref: Below £20,000)						
£20,001-£40,000	0.053*	0.022	0.046*	0.021	0.045*	0.021
£40,001+	0.071*	0.024	0.065*	0.022	0.058*	0.022
Independent Student or Missing data	0.119*	0.03	0.094*	0.029	0.094*	0.029
How parents earns (Ref: Work)						
Benefits	-0.052	0.046	-0.054	0.042	-0.063	0.042
Investments	0.118*	0.031	0.109*	0.028	0.117*	0.029
Ethnic Group (Ref: White)						
Black/Asian	0.064*	0.025	0.063*	0.024	0.059*	0.024
Mixed/Other	0.033	0.024	0.025	0.024	0.018	0.023

Notes: 1 * Indicates significance at the 5% level

2 Results have been split into two tables. Table 4a contains variables directly relating to the research questions posed, while 4b contains the other control variables.

3 A chow test was conducted as to whether the results should be reported separately for men and women. The test illustrated that there is no evidence of any further structural differences.

4 A fixed affects model has also been developed for specification 1. All substantial results remain in place, though the coefficient for students one year away from graduation is reduced to 0.025 and is outside the range of statistical significance.

Table 4b-Regression results

Variable	Specification 1 Coefficient	Specification 1 Standard Error	Specification 2 Coefficient	Specification 2 Standard Error	Fixed Affects (Specification 2) Coefficient	Fixed Affects (Specification 2) Standard Error
Total Income						
Mean Centred (per £10,000)	0.038*	0.017	0.038*	0.018	0.030*	0.018
Study Mode (Ref: Full-Time)						
Part-Time Student	-0.047	0.026	-0.044	0.025	-0.039	0.026
Earnings From Work						
Mean Centred (per £10,000)	0.026	0.022	0.026	0.021	0.030	0.020
Part-Time Student *Earnings From Work	0.166*	0.032	0.152*	0.031	0.150*	0.030
Subject Area (Ref: Medicine)						
Allied To Medicine	-0.206*	0.033	-0.192*	0.031	-0.201*	0.029
Sciences	-0.271*	0.033	-0.215*	0.031	-0.225*	0.030
Maths & Computer Science	-0.209*	0.037	-0.171*	0.038	-0.194*	0.037
Engineering & Technology	-0.170*	0.034	-0.153*	0.033	-0.166*	0.029
Architecture & Building	-0.233*	0.034	-0.209*	0.034	-0.210*	0.034
Social Studies	-0.233*	0.032	-0.189*	0.030	-0.199*	0.027
Law	-0.172*	0.044	-0.135*	0.041	-0.154*	0.041
Business & Admin	-0.230*	0.034	-0.182*	0.034	-0.184*	0.033
English, languages, Classics	-0.302*	0.039	-0.247*	0.037	-0.260*	0.034
History & Philosophy	-0.326*	0.042	-0.264*	0.043	-0.277*	0.042
Arts	-0.362*	0.037	-0.322*	0.035	-0.321*	0.033
Education	-0.220*	0.032	-0.202*	0.030	-0.225*	0.028
Combined	-0.275*	0.039	-0.229*	0.038	-0.242*	0.037
Other	-0.309*	0.054	-0.268*	0.051	-0.265*	0.049
Entry Qualification (Ref: A-levels)						
GNVQ/AVCE	-0.060*	0.023	-0.068*	0.024	-0.064*	0.023
Other	-0.01	0.015	-0.017	0.015	-0.016	0.015
Age						
Mean Centred	0.013*	0.005	0.014*	0.005	0.015*	0.005
University Location (Ref: Other England)						
London	0.091*	0.018	0.093*	0.020	0.163	0.116
Wales	-0.057*	0.022	-0.059*	0.021	-0.008	0.092
Gender (Ref: Male)						
Female	-0.053*	0.013	-0.053*	0.013	-0.050*	0.012
University Dummies	-	-	-	-	X	X
Constant	10.032*	0.042	10.041*	0.041	10.034*	0.061

Notes: 1 * Indicates significance at the 5% level

2 X indicates the university dummy variables have been included but values not reported

Table 5. Student response to whether they believe 'the growing number of graduates will make it hard to get a graduate job'

	Final Year %	1 Year %	2+ Years %
Strongly Agree/Agree	56.5	46.4	48.0
Neutral	14.5	15.4	16.4
Strongly Disagree/ Disagree	26.5	34.6	33.2
Observations	946	758	1,040

Table 6. Response to whether students have considered getting a temporary job

	Final Year %	1 Year %	2+ Years %
No	90.6	94.9	98.1
Yes	9.4	5.2	1.9
Observations	946	758	1,040

Table 7a. Regression results for students' wage expectations: Specification 3 and 4

Variable	SPECIFICATION 3 Coefficient	SPECIFICATION 3 Standard Error	SPECIFICATION 4 Coefficient	SPECIFICATION 4 Standard Error	Fixed Affect Coefficient	Fixed Affect Standard Error
Future Plans (Ref: Career Job Only)						
Temporary Job Only	-0.221*	0.025	-0.219*	0.025	-0.218*	0.025
Either a Career or Temporary Job	-0.125*	0.033	-0.123*	0.033	-0.115*	0.031
Further Study or Travel	-0.008	0.013	-0.007	0.013	-0.006	0.013
Hard to get Grad Job (Ref: Agree)						
Neutral	0.016	0.018	0.014	0.018	0.015	0.017
Disagree	0.055*	0.013	0.052*	0.013	0.054*	0.013
Missing	0.079*	0.031	0.080*	0.031	0.071*	0.031
Proximity to Graduation (Ref: Final Year)						
1 Year	0.014	0.014	0.013	0.014	0.009	0.014
2 or More Years	0.036*	0.017	0.036*	0.017	0.030	0.017
University Type (Ref: Post 1992)						
Other Pre 1992	-0.049*	0.018	-0.030	0.018	-	-
Russell Group	-0.022	0.024	-0.024	0.026	-	-
Parents Earnings (Ref: Below £20,000)						
£20,001-£40,000	0.043*	0.021	0.041*	0.021	0.045*	0.021
£40,001+	0.062*	0.023	0.059*	0.023	0.057*	0.022
Independent Student or Missing data	0.096*	0.030	0.098*	0.030	0.096*	0.029
How parents earns (Ref: Work)						
Benefits	-0.050	0.042	-0.054	0.042	-0.063	0.042
Investments	0.114*	0.028	0.118*	0.029	0.118*	0.029
Ethnic Group (Ref: White)						
Black/Asian	0.066*	0.024	0.062*	0.024	0.059*	0.024
Mixed/Other	0.026	0.024	0.022	0.023	0.018	0.023
UCAS Score (Ref: Mean)						
Per 100 Increase	0.051*	0.012	0.036*	0.012	0.009	0.019
UCAS Score Squared	-	-	0.022*	0.006	0.015	0.011

Notes: 1 * Indicates significance at the 5% level

2 Results have been split into two tables. Table 7a contains variables directly relating to the research questions posed, while 7b contains the other control variables.

3 University Type has been excluded from the fixed affects specification due to its highly correlated relationship with the university dummies and ability variable.

Table 7b. Regression results for students' wage expectations: Specification 3 and 4

Variable	SPECIFICATION 3 Coefficient	SPECIFICATION 3 Standard Error	SPECIFICATION 4 Coefficient	SPECIFICATION 4 Standard Error	Fixed Affect Coefficient	Fixed Affect Standard Error
Total Income						
Mean Centred (per £10,000)	0.039*	0.017	0.037*	0.017	0.030	0.018
Study Mode (Ref: Full-Time)						
Part-Time Student	-0.041	0.025	-0.042	0.025	-0.039	0.025
Earnings From Work						
Mean Centred (per £10,000)	0.029	0.022	0.030	0.021	0.034	0.022
Part-Time Student *Earnings From Work	0.148*	0.032	0.146*	0.031	0.149*	0.032
Subject Area (Ref: Medicine)						
Allied To Medicine	-0.178*	0.031	-0.162*	0.031	-0.182*	0.029
Sciences	-0.191*	0.031	-0.184*	0.031	-0.207*	0.032
Maths & Computer Science	-0.150*	0.039	-0.154*	0.037	-0.184*	0.038
Engineering & Technology	-0.122*	0.033	-0.112*	0.033	-0.146*	0.032
Architecture & Building	-0.197*	0.035	-0.181*	0.035	-0.192*	0.034
Social Studies	-0.170*	0.030	-0.160*	0.028	-0.179*	0.027
Law	-0.136*	0.041	-0.128*	0.041	-0.144*	0.040
Business & Admin	-0.155*	0.035	-0.150*	0.035	-0.166*	0.033
English, languages, Classics	-0.237*	0.036	-0.227*	0.035	-0.243*	0.032
History & Philosophy	-0.247*	0.043	-0.233*	0.044	-0.257*	0.040
Arts	-0.310*	0.034	-0.292*	0.034	-0.300*	0.034
Education	-0.198*	0.027	-0.183*	0.027	-0.203*	0.028
Combined	-0.217*	0.037	-0.203*	0.037	-0.222*	0.038
Other	-0.240*	0.050	-0.234*	0.050	-0.245*	0.048
Entry Qualification (Ref: A-levels)						
GNVQ/AVCE	-0.062*	0.024	-0.065*	0.024	-0.064*	0.022
Other	-0.013	0.015	-0.016	0.015	-0.016	0.015
Age						
Mean Centred	0.014*	0.005	0.015*	0.005	0.015*	0.005
University Location (Ref: Other England)						
London	0.099*	0.020	0.096*	0.019	0.168*	0.069
Wales	-0.037	0.021	-0.039	0.022	0.002	0.094
Gender (Ref: Male)						
Female	-0.053*	0.013	-0.051*	0.012	-0.049*	0.012
University Dummies	-	-	-	-	X	X
Constant	10.049*	0.040	10.025*	0.039	10.019*	0.061

Notes: 1 X indicates the university dummy variables have been included but values not reported

Table 8a. Sample selection Rules for Full-Time Students

Sample Selection Rules	DLHE Sample Size	Average Starting Salary (2005 prices) £	SIES Sample Size	Average Expected Salary £	ratio
Start	256,507	17,145	2,659		
England & Welsh universities only	224,226	17,196	2,659	18,540	1.08
First degree only	180,911	16,650	2,393	18,605	1.12
Salary above £8,000	178,491	17,112	2,339	18,946	1.11
Employed within UK only	168,673	17,081	2,339	18,946	1.11
Over 24 excluded	145,517	16,691	1,923	18,830	1.13
Doing "something else" (e.g. further study/travel) excluded	117,660	16,691	1,923	18,830	1.13
Only those working Full-time	87,327	16,849	1,923	18,830	1.12
Missing salary data excluded	45,906	16,849	1,828	18,830	1.12
Medics excluded	44,436	16,455	1,666	18,084	1.10
SIES weights used	44,436	16,455	1,666	18,288	1.11

Notes:
1. “First degree” in the DLHE survey means excluding those in postgraduate study and foundation courses. The SIES collected expectations data only from those doing a first degree or a foundation course and not from postgraduate students. This explains why there is a large drop in DLHE salary, but not in the SIES expectations.

2. The starting figure given in the table is in 2005 prices. The starting salary in the raw survey data, in 2006 prices, is £17,625. This has been scaled down using the Retail Price Index based on the assumption students report expected wages in current (2005) prices, while the DLHE was conducted in 2006.

3. The SIES weights correct for non-response and to ensure the age and sex profile of the survey is representative of the population of students. They are described in footnote 12 on page 7.

Table 8b. Sample selection rules in the DLHE for Part-Time Students

Sample Selection Rules	DLHE sample size	Average Salary (2005 prices) £
Start 2006 Prices	59,965	
Start 2005 Prices	59,965	25,898
England & Welsh universities only	52,606	25,981
Postgraduate students excluded	28,416	22,126
Salary above £8000	28,416	22,753
Employed within UK only	27,353	22,646
Doing "something else" (e.g. further study/travel) excluded	24,219	22,646
Missing salary data excluded	9,842	22,646
Medics Excluded	9,834	22,316

Table 9. Hypothetical respondents to DLHE salary question

	Person A "Good Prospects"	Person B
Gender	Male	Male
Degree Class	1st	1st
University Type	Russell Group	Russell Group
Work Status	Full-Time	Full-Time
Subject	Engineering	Engineering
Ethnicity	White	White
Disability	None	None
Graduate Job	Graduate Level	*Non-Graduate Job*
University Location	England	England
Term Time Accommodation	Away from parents' home	Away from parents' home
Tariff (mean=300)	430 (1 S.D above mean)	430 (1 S.D above mean)
Home Location	London	London
Degree a job requirement?	Formal Requirement	*Not Required*
Type of job (SOC)	Managerial	Managerial
Probability of responding	76%	62%

	Person C	Person D "Poor Prospects"
Gender	Male	*Female*
Degree Class	*2.2*	*3rd*
University Type	Russell Group	*Post 1992*
Work Status	Full-Time	*Full-Time*
Subject	*Art*	*Art*
Ethnicity	White	*Black*
Disability	None	*Yes*
Graduate Job	*Non-Graduate Job*	*Non-Graduate Job*
University Location	England	*Wales*
Term Time Accommodation	Away from parents' home	Away from parents' home
Tariff (mean=300)	430 (1 S.D above mean)	*170 (1 S.D below mean)*
Home Location	London	*North East*
Degree a job requirement?	*Not Required*	*Not Required*
Type of job (SOC)	*Admin Job*	*Admin Job*
Probability of responding	45%	31%

Notes: 1. Calculated using the probit model in appendix 4.

Table 10. Comparison of SIES and DLHE samples (weighted proportions in brackets)

Full-Time Students	SIES %	DLHE %
Gender *		
Male	33.2 (46.5)	38.9 (41.4)
Female	66.8 (53.5)	61.1 (58.6)
Ethnicity *		
White	86.0 (86.7)	88.3 (87.8)
Asian	5.7 (5.6)	8.0 (8.4)
Black	2.9 (2.4)	1.6 (1.8)
Mixed/Other	5.5 (5.4)	2.1 (2.0)
University Group		
Russell Group	22.9 (28.4)	23.4 (22.7)
Other pre 1992	20.6 (18.9)	21.4 (20.3)
Post 1992	56.5 (52.8)	55.2 (57.0)
Social Class (Parents Occupation)[¥] *		
Managerial/Professional	58.9 (60.0)	56.7 (56.3)
Intermediate	20.9 (20.1)	28.6 (28.4)
Routine/Manual	20.3 (19.1)	14.7 (15.3)
Living Arrangement [Ψ]		
Parental Home	21.9 (19.4)	21.2 (22.4)
Living Away from Home	78.1 (80.6)	78.8 (77.6)
Subject*		
Allied To Medicine	5.5 (4.5)	7.0 (6.3)
Sciences	9.5 (10.4)	8.7 (8.1)
Maths/ Comp Science	5.8 (6.6)	9.3 (8.9)
Engineering & Technology	4.0 (5.0)	5.1 (4.7)
Architecture & Building	1.7 (2.0)	1.6 (1.7)
Social Sciences	13.7 (13.9)	10.1 (9.7)
Business	8.5 (8.8)	14.6 (14.1)
Languages	6.1 (5.8)	8.7 (8.5)
History	4.7 (5.3)	4.8 (5.0)
Art	13.9 (13.3)	8.8 (11.6)
Education	8.2 (6.9)	5.8 (5.6)
Combined	4.2 (3.9)	0.6 (0.5)
Psychology	1.9 (1.4)	4.1 (4.0)
Sports Science	1.7 (1.7)	3.2 (3.1)
Law	5.0 (5.2)	2.5 (2.6)
Mass Communication	3.0 (2.9)	3.3 (3.7)
Other	2.6 (2.4)	1.8 (1.9)
Number of Observations	1,666	44,436

Part-Time Students	SIES %	DLHE %
Gender*		
Male	33.1 (46.3)	42 (43.4)
Female	66.8 (53.6)	58 (56.6)
Ethnicity*		
White	87.7 (89.4)	90.9 (88.9)
Asian	4.1 (2.0)	4.2 (4.9)
Black	3.6 (4.1)	3.5 (4.3)
Mixed/Other	1.4 (4.1)	1.4 (1.7)
University Group*		
Russell Group	12.1 (5.6)	3.4 (4.5)
Other pre 1992	15.5 (11.1)	27.6 (19.0)
Post 1992	72.3 (83.3)	69 (76.5)
Subject*		
Medicine	3.3 (2.9)	0.1 (0)
Allied To Medicine	11.2 (8.4)	16.6 (18.4)
Sciences	3.1 (4.0)	3.3 (3.0)
Maths/ Comp Science	5.6 (6.4)	6.9 (6.9)
Engineering & Technology	8.6 (16.9)	8.9 (8.2)
Architecture & Building	5 (6.7)	5.7 (6.1)
Social Sciences	9.6 (6.8)	11.1 (8.9)
Business	8.1 (9.0)	10.9 (11.0)
English & Languages	2.8 (2.4)	1.7 (1.8)
History	4.1 (2.8)	3.0 (3.2)
Art	4.0 (4.6)	1.2 (2.0)
Education	21.8 (18.0)	14.4 (16.3)
Combined	3.9 (3.3)	9.2 (7.0)
Psychology	0.4 (0.3)	2.5 (2.1)
Law	3.7 (3.3)	2.6 (2.5)
Mass Communication	0.9 (0.6)	0.6 (0.1)
Other	3.9 (3.8)	1.6 (1.6)
Number of Observations	784	9,842

Notes: 1 Proportions when weights are used are reported in brackets. The weights used in the SIES correct for non-response and ensures the age-sex profile of the sample matches that of the student population. These are described in footnote 7 on page 12. The DLHE weights refer to non-response weights I have calculated to take into account missing data for graduates salary. These were based on the probit model in appendix 4. Further details can be found on page 20.
2 ¥ Given for the 20,156 (46%) observations with social class known in the DLHE.
3 Ψ17,435 unknown. % given for the 57,829 observations with living arrangement reported.
4 DLHE and SIES figures are for those who reported wages and after sample selection procedures have been applied
5 Medics have been excluded after initial analysis showed a greater number in the SIES than DLHE
6 * Indicates chi-squared test statistically significant at 5% level

Table 11. Comparison of average salary in the SIES, DLHE and LFS for full-time students (weighted estimates in parenthesis)

	SIES Average Expected Wage £000	DLHE Average Wage £000	LFS Average Wage £000	Ratio SIES:DLHE	Ratio SIES:LFS
Mean Wage	18.1 (18.3)	16.4 (16.0)	16.1	1.10 (1.14)	1.13
Median Wage	18.0	16.0	15.1	1.12	1.19
N	1,684	47,702	194		

Notes: 1. See Table 10, footnote 1 for more details on the weighted estimates.

2. SIES weighted by those provided in the dataset. DLHE weights are those calculated in section 7.

3. With the two outliers included, the mean (median) LFS wage is £16,749 (£15,163), suggesting a ratio of 1.08 (1.19). All these figures are statistically significant at the 5% level compared to the expected salary.

4. Without the adjustment for inflation, the mean (median) LFS wage would be £16,865 (£15,600) without the outliers, and £17,557 (£15,600) with. All these figures are statistically significant at the 5% level compared to the expected salary, except for the mean non-inflation adjusted salary when the outliers are included (£17,557). Nether-the-less, there is substantial evidence to support the robustness of results to the various assumptions being made.

Table 12. Comparison between mean expected and mean actual wages for full-time students, based on background characteristics (weighted estimates in parenthesis)

	Mean Expected wage £000	Mean Actual wage £000	Ratio
All Full-time Students			
Final Year	17.4 (17.7)	16.5 (16.0)	1.06 (1.10)
1 Year	17.9 (18.2)	16.5 (16.0)	1.09 (1.14)
2 or More Years	18.8 (18.9)	16.5 (16.0)	1.14 (1.18)
Ethnic Group			
Black/Asian	19.2 (19.6)	17.3 (17.2)	1.11 (1.14)
White	17.9 (18.1)	16.1 (15.8)	1.12 (1.13)
Mixed	18.6 (19.0)	16.8 (16.9)	1.11 (1.13)
University Type			
Russell Group	19.2 (19.4)	17.3 (17.1)	1.11 (1.13)
Other pre 1992	17.6 (17.8)	16.6 (16.5)	1.06 (1.08)
Post 1992	17.8 (17.9)	15.6 (15.4)	1.14 (1.16)
Gender			
Male	18.7 (18.9)	16.9 (16.6)	1.11 (1.14)
Female	17.8 (17.8)	15.8 (15.6)	1.12 (1.14)

Notes: 1 See Table 10, footnote 1 for more details on the weighted estimates.

2 Excludes medical and part-time students, due to different proportions found within the two sources.

3 For all variables the mean expected and actual salary are significantly different at the 5% level

4 Social Class has not been investigated due to a large amount of missing data in the DLHE for this variable.

Table 13a. Comparison between expected and actual mean wages for subject groups (weighted estimates in parenthesis)

Subject	Sample Size SIES	Mean Expected Wage £000	Mean Actual Wage £000	Ratio
Medicine	162	26.5 (25.8)	28.8 (28.5)	0.92 (0.91)
Education	137	17.9 (17.9)	17.5 (17.4)	1.02 (1.03)
Engineering, Maths & Computer Science	164	19.5 (19.2)	18.4 (17.5)	1.06 (1.10)
Allied To Medicine	92	18.9 (19.0)	17.1 (16.7)	1.10 (1.14)
All Full-Time Students (excluding medics)	1,684	18.1 (18.3)	16.5 (16.0)	1.10 (1.14)
Social Sciences	230	18.5 (18.9)	16.6 (16.4)	1.11 (1.17)
Art	238	16.3 (16.4)	14.4 (14.2)	1.13 (1.15)
Business, Admin & Management	143	18.4 (18.9)	16.3 (16.1)	1.13 (1.17)
History, English & Languages	232	17.5 (17.7)	15.3 (15.1)	1.15 (1.17)
Psychology, Sports Studies, Combined and Other	204	17.7 (17.7)	15.3 (15.3)	1.15 (1.16)
Sciences	161	18.3 (18.2)	15.8 (15.7)	1.16 (1.16)
Law*	83	20.7 (20.6)	14.8 (14.8)	1.35 (1.39)

Notes: 1 See Table 10, footnote 1 for more details on the weighted estimates.

2 For all variables the mean expected salary is significantly different from mean actual salary at the 5% level, except for Education.

3 "All full time students" excludes medicine students

.

4 Law is somewhat of an outlier, due to the large number of students continuing into postgraduate training. Excluding this group from the aggregate analysis has little influence on the overall result.

Table 13b. Comparison between expected and actual median wages for subject groups

Subject	Median Expected £000	Actual Median £000	Ratio Expected: Actual (Medians)
Medicine	26.0	28.7	0.91
Education	18.0	18.0	1.00
Allied To Medicine	18.0	17.0	1.06
Art	15.0	14.1	1.06
Engineering, Maths & Computer Science	20.0	18.0	1.11
All Full-Time Students (excluding medics)	18.0	16.0	1.12
Social Sciences	18.0	16.0	1.12
Psychology, Sports Studies, Combined and Other	17.0	15.1	1.13
History, English & Languages	18.0	15.1	1.19
Sciences	18.0	15.1	1.19
Business, Admin & Management	18.0	15.1	1.19
Law*	2000	14.1	1.42

Table 14a. Comparison between average expected and average actual wages of men in each subject group

Subject	Sample Size SIES	Mean Expected Wage £000	Mean Actual Wage £000	Ratio
Medicine	55	27.5	29.0	0.95
Education	21	18.3	17.2	1.06
Engineering, Maths & Computer Science	115	19.6	18.7	1.04
Allied To Medicine *	16	19.9	16.7	1.19
All Full-Time Students (excluding medics) *	559	18.7	16.9	1.11
Social Sciences*	74	19.7	18.1	1.09
Art*	71	16.7	14.8	1.13
Business, Admin & Management *	60	18.9	17.1	1.10
History, English & Languages *	60	18.5	15.5	1.19
Psychology, Sports Studies, Combined and Other	63	17.3	16.4	1.05
Sciences *	56	18.1	16.5	1.10
Law *	23	20.9	15.9	1.31

Table 14b. Comparison between average expected and average actual wages of women in each subject group

Subject	Sample Size SIES	Mean Expected Wage £000	Mean Actual Wage £000	Ratio
Medicine *	107	26.0	28.6	0.91
Education	115	17.8	17.8	1.00
Engineering, Maths & Computer Science	49	19.1	18.1	1.06
Allied To Medicine *	76	18.7	17.0	1.10
All Full-Time Students (excluding medics) *	1,125	17.8	15.8	1.12
Social Sciences*	156	17.8	16.1	1.10
Art*	167	15.8	14.7	1.08
Business, Admin & Management *	83	18.0	15.9	1.13
History, English & Languages *	172	17.1	15.6	1.10
Psychology, Sports Studies, Combined and Other *	141	17.7	15.1	1.18
Sciences *	105	18.3	15.8	1.16
Law *	60	19.8	14.7	1.34

Notes: 1 SIES sample size refers to unweighted data

2 Sample size for men and women differ substantially due to differential response rates. See footnote 22 and table 10 for further details

3 Results were also investigated with the use of sampling weights. The substantial results in Table 14 remain.

4 * Indicates ratio is statically different from 1 at the 5% level.

Appendix 1. Definition of Part-time and Dependent Students

The SIES used the following definition for whether a student is classed as a dependant or independent student:

Dependent students	These are full-time students: • who had applied for student support and their parent/step parent, legal guardian's income had been taken into account • or were aged under 25 years, were unmarried and had not applied for student support. **Independent students** are all part-time students and full-time students not in the above category.
Part-time student	Students (excluding OU) on an undergraduate or PGCE course lasting at least one academic year and equivalent to at least 50 per cent of a full-time course. OU students were included if they were studying for an undergraduate degree or PGCE course and eligible for financial support (which involved taking or registering for a course worth more than 60 credit points). NB This profile means that the OU students included in this survey are not typical of all OU students.

Source: Extract from SIES 2004-2005 report

The National Centre for Social Research was also contacted to clarify the definition used for a dependent student. The questions asked and answers given appear below.

Q: Is a person, who has applied for student support, unmarried, under 25, but **has not had** their parent's income taken into account classed as a dependent or independent student?

A: Independent student

Q: Is a person, who has applied for student support, unmarried, under 25, **had** their parent's income taken into account**, but is above the threshold** for any further support other than the basic level, a dependent or independent student??

A: Dependent student

This highlights that whether parents' income has been taken into consideration is important to deciding how the student is classed. Anyone who is 25 or over, married or has not had their parent's income taken into account is classed as an independent student. Those whose parent's income has been taken into account are dependent students. One problem identified from the responses may be that parents know they are above the income threshold for any further support and hence do not disclose this information on the UCAS application form. Therefore it is possible that some respondents, who are actually dependent students, are mistakenly classified as independent. The data records there to be 1,278 independent students with 691 of these part-time students and 587 full-time. It is likely that there is some measurement error in recording some of these 587 full-time students as independent, for the reasons stated above.

Appendix 2. Construction and difficulties in measuring parental income

Parental income has been recorded in band widths of £5,000 or £10,000, though the respondent could decide whether to report the information in a gross or net amount. It is necessary to try and put the data on the same scale. In this instance, it was decided to put all net data on the gross scale. The modifications to the data are shown below.

Table 15. Conversion between net and gross parents income

Net Value £	Gross Value £	Category £
0-5,000	0-5,000	20,000 and below
5001-10,000	5,000-12,000	20,000 and below
10,001-15,000	11,900-19,400	20,000 and below
15,001-20,000	19,400-27,000	20,001-40,000
20,001-25,000	27,000-34,000	20,001-40,000
25,001-30,000	34,000-41,000	20,001-40,000
30,001-40,000	41,000-58,000	40,001+
40,001-50,000	58,000-75,000	40,001+
50,001+	75,001+	40,001+

There are some limitations to this technique. The choice of groups is largely dictated by the data. For instance a net salary of £15,000 is roughly equal to a gross salary of £20,000. Being that £15,000 and £20,000 are both cut off points, it is sensible to create a category of income £20,000 or below to minimize overlap between groups. With the categorizations used, there should be little overlap, though this can not be totally avoided. For instance a student may know that their parent earns £15,100, equivalent to £19,500 gross, per year. However this student would be put into the group £20,000-£40,001 gross per annum because of the overlap problem. It is reasonable to suggest that the analysis will not be severely affected by this, as the overlap is small. Some other assumptions must be made about this variable. The question asked is about the total income of parents. This may complicate the conversion between net and gross. In particular the tax, if only one parent is earning the income, is greater than if two parents are working. For example, consider two households with £30,000 NET income. Household A has one parent working who earns the whole £30,000. The gross equivalent is £41,000 per year. Household B however has 2 parents earning £15,000, with gross equivalent being £38,800. Hence there is a difference between the gross equivalence due to the tax system that would put household A into the £40,000+ bracket and household B into the £20,001-£40,000 group. Furthermore there is an issue that some forms of income are not taxable, such as child benefit. For simplicity, it has been assumed that all parental income is taxable and has been generated by one adult in the household. With the boundaries chosen the effect is probably quite small, but it is still important to note this difficulty. It is also important to recognize that this variable may suffer a reasonably large degree of measurement error, as it relies on students reporting their parents' income. Hence the quality of the variable relies on students accurately knowing their parents' income. A further point to note is that some (36) students failed to state whether they were reporting figures in gross or net terms. In this instance it has been assumed that students are reporting gross figures.

It should also be noted that this is only one of several possible ways to classify this variable. One drawback is that the measurement of parents' income is quite coarse, as there are only three groups. An alternative is to take the midpoint of the groups, and create a quasi-continuous variable. This would have the benefit of providing a broader sense of parents' income, though the difficulty of conversion between net and gross income still exists. In both ways of handling the data there is a significant chance of measurement error. Thus the variable should be viewed as an approximate measure of parents' income.

Appendix 3. Interval regression Results

Table 16a. Interval regression results

Variable	Specification 4 Coefficient	Standard Error
Future Plans (Ref: Career Job Only)		
Temporary Job Only	-0.228	0.025
Either a Career or Temporary Job	-0.096	0.023
Further Study or Travel	-0.005	0.013
Hard to get Grad Job (Ref: Agree)		
Neutral	0.021	0.018
Disagree	0.059	0.012
Missing	0.081	0.033
Proximity to Graduation (Ref: Final Year)		
1 Year	0.014	0.014
2 or More Years	0.035	0.017
University Type (Ref: Post 1992)		
Other Pre 1992	-0.031	0.018
Russell Group	-0.024	0.026
Parents Earnings (Ref: Below £20,000)		
£20,001-£40,000	0.031	0.019
£40,001+	0.053	0.020
Independent Student or Missing data	0.087	0.030
How parents earns (Ref: Work)		
Benefits	-0.067	0.043
Investments	0.116	0.029
Ethnic Group (Ref: White)		
Black/Asian	0.055	0.025
Mixed/Other	0.019	0.023
UCAS Score (Ref: Mean)		
1 Standard deviation Change	0.036	0.011
UCAS Score Squared	0.023	0.006

Note: 1 * Significant at 5% level

Table 16b. Interval regression results continued

Variable	Specification 4 Coefficient	Standard Error
Total Income		
Mean Centred (per £0000)	0.036	0.017
Study Mode (Ref: Full-Time)		
Part-Time Student	-0.033	0.025
Earnings From Work		
Mean Centred (per £0000)	0.030	0.021
Part-Time Student *Earnings From Work	0.141	0.032
Subject Area (Ref: Medicine)		
Allied To Medicine	-0.162	0.032
Sciences	-0.183	0.032
Maths & Computer Science	-0.141	0.035
Engineering & Technology	-0.111	0.034
Architecture & Building	-0.182	0.036
Social Studies	-0.155	0.029
Law	-0.114	0.043
Business & Admin	-0.149	0.036
English, languages, Classics	-0.225	0.036
History & Philosophy	-0.235	0.046
Arts	-0.285	0.035
Education	-0.181	0.028
Combined	-0.199	0.038
Other	-0.215	0.048
Entry Qualification (Ref: A-levels)		
GNVQ/AVCE	-0.072	0.023
Other	-0.013	0.015
Age		
Mean Centred	0.013	0.005
University Location (Ref: Other England)		
London	0.105	0.021
Wales	-0.036	0.022
Gender (Ref: Male)		
Female	-0.052	0.013
University Dummies		
Constant	10.009	0.040

Students' estimates tend to bunch around round numbers, despite an open text field allowing precise estimates to be recorded. How may we take this into account when estimating regression coefficients? One possibility is to assume students do not expect to get exactly the salary they report, but give a "ball-park" figure. A student who estimates their salary does not expect to receive exactly this amount, but somewhere around the figure. As an example, a student predicting a salary of £13,000 may be presenting their midpoint estimate or reporting to the nearest thousand. In reality they expect a salary between £12,500 and £13,499.

One way to treat the data is to suggest that students' actual estimates are unknown, but lie within a censored interval. If it is assumed that students round their estimate to the nearest thousand, intervals as described in the preceding example may be appropriate. It is also necessary to assume that the unobserved response (expected starting wage) is normally distributed.

However this does not account for the extra large peaks at certain points. Examination of the data suggests that extra clustering occurs at numbers that are rounded to the nearest £5,000. Consequently it seems appropriate to assume that students who are reporting these figures are exercising a greater degree of rounding and have a wider anticipated salary range. Therefore the interval is adjusted for figures at £15,000, £20,000, £25,000, £30,000 and £35,000 under the assumption that students round their estimate to the nearest £5,000. Hence a student who predicts a wage of £30,000 is assumed to have a censored estimate within the range £27,500 to £32,499.

Results for the censored regression model appear in the table above. Compared to the original OLS regression, very little changes with the introduction of the censoring assumption and use of interval regression. Both coefficients and standard errors are close to the OLS estimates. Most coefficients alter by around 0.2 to 0.3% suggesting that, even when assuming quite extreme rounding by students, there are limited differences compared to using ordinary least squares.

Appendix 4. Probit results for Salary non-response

Table 17a. Results for probit model of non-response

Variable	Group	Coefficient	Std. Err.
Domicile (Ref: London)	North East	0.221	0.027
	North West	-0.092	0.019
	Yorkshire	0.005	0.021
	East Midlands	0.027	0.021
	West Midlands	-0.009	0.020
	East	-0.054	0.020
	South East	0.011	0.018
	South West	0.009	0.021
	Isle of Man/ Channel Islands	-0.391	0.110
	Unknown	-0.072	0.024
University Location (ref England)	Wales	-0.240	0.021
Term-Time Accommodation (Ref: Uni Maintained)	Parental Home	-0.136	0.016
	Own home	-0.043	0.014
	Other	-0.090	0.019
	Unknown	-0.025	0.021
Degree Class (Ref: 1st)	2.1	-0.082	0.017
	2.2	-0.180	0.018
	3rd	-0.276	0.027
	Unclassified	-0.166	0.046
	Not Applicable	-0.107	0.024
UCAS Score (Ref: Mean)	100 point increase	0.039	0.006
Uni Type (Ref: Post 1992)	Russell Group	-0.097	0.014
	Pre 1992	-0.133	0.013
Subject (Ref: Medicine)	Allied to Medicine	0.126	0.040
	Biology	0.174	0.043
	Physical Sciences	0.227	0.042
	Maths	0.183	0.046
	Computer Science	0.165	0.040
	Engineering	0.249	0.042
	Social Sciences	0.210	0.039
	Law	0.139	0.046
	Business Admin	0.202	0.038
	Mass Communication	0.067	0.044
	Languages	0.102	0.040
	History	0.069	0.042
	Art	-0.117	0.039
	Education	0.012	0.042
	Combined	0.050	0.071
	Psychology	0.223	0.043
	Sports Science	0.144	0.044
	Other	0.089	0.049

Table 17b. Results for probit model of non-response continued

Variable	Group	Coefficient	Std. Err.
Graduate Level Job (Ref: Yes)	Non-graduate job	-0.115	0.030
Degree Required for job (Ref: Formal requirement)	Expected	-0.066	0.018
	Advantageous	-0.075	0.015
	No	-0.271	0.015
	Don't Know	-0.702	0.034
Job Type (Ref: Managerial)	Professional	0.042	0.020
	Associate Professional	0.059	0.018
	Admin	0.100	0.033
	Skilled Labour	-0.090	0.059
	Personal Services	-0.064	0.039
	Sales/ Customer Service	-0.053	0.036
	Construction	0.042	0.074
	Elementary	-0.366	0.043
Ethnicity (Ref: Asian)	Black	0.015	0.039
	Other/ mixed	0.135	0.037
	Unknown	0.035	0.036
	White	0.148	0.018
Disabled (Ref: No)	Yes	-0.072	0.019
Gender (Ref: Male)	Female	-0.098	0.010
	Constant	0.308	0.052

Appendix 5. The difference between HESA's official average graduate wage and the figures used in this paper

HESA report the average graduate wage for the 2004/2005 year group as £19,000[70], calculated from the DLHE survey used in this paper. This is in fact the starting wage calculated from the data (£18,531), rounded to the nearest thousand. The average graduate wage presented in this paper is lower than this official figure. The table below shows how the figures presented in this paper relate to the official HESA average. As can be seen, the difference in graduate salaries is largely due to the different samples being considered.

Among the most important points is that we are calculating wages **separately** for those who studied part-time and those who studied full-time (selection rule 1). In comparison, the official HESA figure relates to when these groups are analysed **together.** An important research result, shown in section 7, is that part-time students, on average, are actually quite realistic in their wage expectations. It is full-time students who, on average, overestimate their starting wage and this should be made explicitly clear when reporting results.

A further adjustment is that we have scaled the HESA data back to 2005 prices using RPI inflation, recorded as 2.8% in 2005 (selection rule 4). As explained in section 6, the SIES was conducted in January to March 2005, whereas the DLHE for this cohort was conducted in early 2006. The existing academic literature, and our overall understanding of data of this type, suggests that individuals report their wage expectations in current prices. Since data on actual wages has been collected a year later than expectations, it is necessary to account for the inflation over this period. It should be noted however that this has only a moderate influence on the results, and the general findings of this paper would still hold if it is not made.

Other points to note are that we have excluded medical students (selection rule 5) and those over 25 from the study of full-time students (selection rule 3). Medics have been excluded due to the different proportion of these students in the SIES compared to the DLHE. Although by excluding medics the DLHE starting salary is reduced, the same selection rule has been applied to the SIES, with an even sharper fall in expectations. This is shown in table 8. Furthermore, only students under 25 have been considered in our analysis. Our intention in studying full-time and part-time students separately was to investigate how realistic "traditional" university students are, who have little pre-existing experience of the labour market. Restricting the sample to those under 25 fits in with this research aim. Again table 8 shows that this restriction has a largely negligible influence on results.

A final point is that the official HESA data does not take into account the large number of graduates not reporting their salary, who are generally working in non-graduate jobs and have 2.2 or 3rd class degrees. When this is taken into account, I estimate that HESA's official figure is upwardly biased by around 3%, or £500 (see rule 6).

[70] See http://www.hesa.ac.uk/index.php/content/view/126/161/

Table 15. Comparison between the average graduate salary reported by HESA and the average graduate salary used in this paper

	Average Salary £	Selection Rule
Raw HESA data	20,314	
1st Degree students Only	18,381	
Only those in Full-time jobs	18,540	
Only those employed in UK *	18,531	
Only those who were studying full-time when at university	17,720	1
English and Welsh Universities only	17,743	2
Under 25 years old only	17,336	3
Scaled to 2005 prices	16,788	4
Medics excluded **	16,455	5
Weighting for non-response	15,996	6

Notes: The figures in the average salary column relate to the cumulative affect of all the sample selection rules.

* Official HESA figure
** Figure used in this paper

References

Betts, Julian R. (1996): What Do Students Know About Wages? Evidence from a Survey of Undergraduates, Journal of Human Resources 31/1, 27-56.

Bosworth, D. L. and Ford, J. (1985a), 'Income Expectations and the Decision to Enter Higher Education', Studies in Higher Education, 10, 21–32.

Botelho, A Pinto, L (2004). Students' expectations of the economic returns to college education: results of a controlled experiment, Economics of Education Review 23 (2004) 645–653

Brunello, G., Lucifora, C., Winter-Ebmer, R., (2001). The wage expectations of European college students. Quaderni dell'Istituto di Economia dell'Impresa e del Lavoro, Universita' Cattolica del Sacro Cuore, Milan, No.30,

Chevalier, A and Conlon, G (2003), Does it pay to attend a prestigious university?, CEE Discussion Papers 0033, Centre for the Economics of Education, LSE

Chevalier, A (2008) Students' Academic Self-Perception, CEE Discussion Papers, Centre for the Economics of Education, LSE

Dominitz, j & Manski, C (1996) Eliciting Student Expectations of the Returns to Schooling, The Journal of Human Resources, pp.1-26

Finch et al (2006), Student Income and Expenditure Survey 2004-2005 Report, Department for Education and Skills

Finch et al (2006), Student Income and Expenditure Survey 2004-2005 Technical Report, Department for Education and Skills

Freeman, R. (1971): The Market for College-Trained Manpower. Cambridge: Harvard University Press.

Gustman & Stafford (1972), Income Expectations and the Consumption of Graduate Students, Journal of Political Economy, Vol. 80, No. 6 (November/December, 1972) pp. 1246-1258

Manski, Charles F. (1993): Adolescent Econometricians : How Do Youth Infer the Returns to Schooling?, in Studies of Supply and Demand in Higher Education, eds. C. Clotfelder and M.Rothschild, Chicago: University of Chicago Press, 43-60.

Manski, Charles (2004): Measuring Expectations, Econometrica, 72(5), 1329-1376.

Oosterbeek, H Ophem, H (2000)."Schooling Choices: preferences, discount rates and rates of return", Empirical Economics, Vol 25, pp 15-34

Purcell, K & Elias, P & Davies, R & Wilton, N (2005): The Class of '99:A study of the early labour market experience of recent graduates

Rouse, C (2004) Low-Income Students and College Attendance: An Exploration of Income Expectations, Social Science Quarterly, Volume 85, Number 5, December 2004

Smith, A (1776). An Inquiry into the Nature And Causes of the Wealth of Nations

Smith, H & Powell, B (1990) Great Expectations: Variations in Income Expectations Among College Seniors Sociology of Education, Vol. 63, No. 3 (Jul., 1990), pp. 194-207

Vignoles, A (2007) The value of higher education in the post-expansion era Presentation to the British Educational Research Association at the Institute of Education

Webbink, Dinand and Joop Hartog (1999): Can Students Predict their Starting Salary? Yes!, mimeo, University of Amsterdam.

Williams, G & Gordon (1981), A Perceived Earnings Functions and Ex Ante Rates of Return to Post Compulsory Education in England ,Higher Education, Vol. 10, No. 2 (Mar., 1981), pp. 199-227

Wolter S.C.; Zbinden A. (2002) Labour market expectations of Swiss university students, International Journal of Manpower, Volume 23, Number 5, 2002 , pp. 458-470(13)

The AGR Graduate Recruitment Survey 2006: Winter Review, Association of Graduate Recruiters (AGR)