What is the Impact of University Expenditure on Graduates' Incomes?

George White

Economics Dissertation 2020/21

Student Number: 14328698

Supervisor: Cecilia Testa

Word Count: 7113

This dissertation is submitted in partial fulfilment of the requirements for the degree. I hereby grant permission for my dissertation to be used for teaching in future years.

Contents

INTRODUCTION 3

LITERATURE REVIEW 4

METHODOLOGY 7

DATA SOURCES 7
RELATIONSHIPS 9
MODEL 13
SUMMARY STATISTICS 14

RESULTS 15

PRIMARY MODEL 15
EXPENDITURE VARIABLE ROBUSTNESS 15
LEO MODEL 16
REF MODEL 18
FIXED EFFECTS MODEL 19

CONCLUSION 20

REFERENCES 22

APPENDICES 24

APPENDIX A: CONVERTING HESA RAW DATA TO AVERAGE INCOME DATA 24
APPENDIX B: FULL RESULTS TABLES 26

Abstract

Amongst economists there is disagreement as to whether quality inputs significantly improve students' outcomes. I look to contribute to the existing literature by examining whether a university's expenditure affects its graduates' incomes. I make use of the first release of the Higher Education and Statistics Authority's Graduate Outcomes data, and universities' financial reports, to compare 120 British universities. Despite initial findings suggesting a significant positive effect of expenditure on incomes, this result does not stand up to robustness tests and alternate models. I conclude that there is no evidence of a statistically significant relationship between universities' expenditures and their graduates' incomes.

Introduction

The link between quantity of education and incomes is well established but the relationship between quality of education and incomes, particularly in higher education, is less well understood. Many researchers have failed to find a significant relationship between quality factors, such as student-staff ratios and teachers' qualifications, and outcomes at various levels of education. In this study I examine this quality-outcome relationship using cross-sectional data on British universities. Specifically, I look at whether universities' expenditures are associated with their graduates' earnings. I make use of the Higher Education Statistics Authority's (HESA) Graduate Outcomes data for the 2017/18 graduating cohort, released in June 2020[1], along with data from universities' financial statements and The Guardian's 2018 university league table.

Despite six decades of research into the issue, there is no consensus as to whether educational quality improves outcomes. I aim to add not only a new and updated account to the literature, but also to approach the question from a different angle. Whilst wide-ranging, the existing literature generally focuses on American high school data from the mid-twentieth century with individual students as the unit of observation. My research focuses on British university data from 2018 with individual universities as the unit of observation. Such an approach will allow me to make an original contribution to the literature, rather than simply replicating existing studies.

My research will be of interest to both universities and prospective students, as well as those associated with them such as schoolteachers, parents, and compilers of the university league tables. For students, the results will offer an insight into the weight they should put on quality factors, relative to student satisfaction for example, with regards to their university destination. For universities, this research will inform what spending most benefits their undergraduates: is hiring the best possible academics in their interests? Finally, there are potential policy implications. Numerous reforms have taken place in the UK's higher education sector in the past decade including higher tuition fees, changes to loan repayments, and a removal of the cap on student numbers at a given university (Hamilton, 2016). One of the most significant outcomes of these reforms is increased marketisation, with strong incentives for universities to attract more students, given the removal of the student numbers cap and a higher proportion of their incomes coming from tuition fees. A knock-on effect could be a lower quality of education, for example if student-staff ratios increase. Such questions are not the focus of my research, but my results will offer evidence as to whether such an effect might be worth examining further.

There are many outputs of university education: tangible and intangible skills, broadened horizons, cultural exchange, and friendship, to name but a few. No single measure could claim

[1] Contains HESA Data: Copyright Higher Education Statistics Agency Limited 2020. The Higher Education Statistics Agency Limited cannot accept responsibility for any inferences or conclusions derived by third parties from its data.

to capture the entire outcome of an individual's university experience; but as an economist I measure outcome by a graduate's income as it is the main quantifiable return from studying, broadly comparable across all graduates. It is an outcome that is of elevated importance in the UK. Tuition fees trebled in 2006 and then again in 2012, leaving them amongst the highest in the world. Furthermore, student loan repayments are linked to earnings, making the financial return to higher education an important consideration for prospective students.

Similarly, there are many potential measures of a university's quality. The literature generally uses student to staff ratios (SSR), with a lower ratio assumed to be better quality due to students receiving more individual attention. I use this measure in my models, alongside university expenditure per student. I believe this is a more relevant measures as it encapsulates the number of staff as well as their perceived quality, measured by their wage. It also covers spending on materials to support learning, such as libraries and facilities, that student-staff ratios do not capture. It could also be noted that broad measures of expenditure will also capture spending unrelated to learning, such as corporate facilities and marketing, and so I will consider several different measures of expenditure per student.

From my primary model using the HESA data, I find initial evidence in favour of expenditure per student being significantly positive, robust to different measures of expenditure. However, due to data constraints, I am unable to fully control for subjects studied at each university, and so that model alone can only provide weak evidence. Further analysis that makes use of a similar dataset, which allows subjects to be controlled for using subject fixed effects, fails to support the findings of my primary model. I conclude that, within the data available to me, there are no consistently significant effects of university expenditure on graduates' incomes.

My study has four main sections: a literature review, methodology, results, and conclusions. The literature review considers the quality-outcome relationship in both school and university settings. The methodology section discusses the sources of the data, the characteristics and relationships of these data, and a specification of the primary model I estimate. The results section details the findings from this primary model and adds discussion regarding the strengths and weaknesses of this model, before providing robustness checks using different data sources and methods. I conclude with a discussion of my overall findings and their implications, whilst considering the limitations of this study and ways future research could improve upon it.

Literature Review

I will consider literature examining the effect of expenditure on student outcomes in both school and university[2] contexts. Whilst the latter is evidently more relevant to my dissertation, I will still examine the school-related literature for three reasons. Firstly, the

[2] The literature generally refers to higher education institutions as "colleges". I will refer to them as "universities" for consistency.

theory behind the two topics is very similar, so lessons from the literature can be applied to my research, and secondly, research in the school-context is often able to draw on better data, as well as scenarios that better lend themselves to econometric analysis. Finally, research on universities in this context is relatively sparce compared with that on primary and secondary education (James, et al., 1989).

The natural starting point for examining the effect of expenditure on schooling outcomes is the Coleman Report, a US government commissioned report into racial inequality in its education system. Whilst the report is not of direct relevance to my research, one of the report's conclusions is that school quality has no statistically significant impact on student performance on standardised test scores (Coleman, 1966). It is this finding that has motivated several related studies, given the influence of Coleman Report at the time and the subsequent policy implication that increased funding for schools gives a poor return on investment (Card & Krueger, 1992).

Eric Hanushek (1986) offers a highly influential review focused on this particular conclusion of the Coleman Report. He notes that two-thirds of school expenditures are made up of instructional (teaching) expenditures, which are determined by class sizes and teacher salaries. Teacher salaries are, in turn, determined by years of teaching experience and completed education. He conducts a meta-analysis of 147 separately estimated production functions based on these three signals of school quality (class size, teacher experience and teacher education) and concludes that the research generally agrees with Coleman. The effect of class sizes on outcomes is inconclusive with very few estimates offering a statistically significant positive effect, and a similar number giving statistically significant negative effects. Teacher education has similar results, though teacher experience does appear to be generally positive. His overall conclusion agrees with Coleman: there is no systematic relationship between school expenditure and student performance.

An alternative view is given by David Card and Alan Krueger in two papers (1992; 1996). In the former, using US census data from 1980, they show how school quality factors, including class sizes and teacher education, increase the returns to education for men born between 1920 and 1949. In the latter, they criticise Hanushek's 1986 review, and an updated version (Hanushek, 1996). They cite Hedges, Laine, and Greenwald's (1994) own meta-analysis of the magnitudes of the estimates used by Hanushek, which concludes that the data are more consistent with a positive effect than that of no or negative effects. Card and Krueger further argue that if the likelihood of any given estimate being positive or negative is fifty percent, the chances of there being twice as many positives than negatives, as was the case for Hanushek's analysis of expenditure per student, are "less than one in a million". They concede that, at the time of their work, the effect of expenditure on outcomes is ambiguous, with the literature generally suggesting it has no effect on test scores, but a possible positive effect on incomes. The latter is of greater relevance to my research. That said, Angrist and Lavy (1999) use instrumental variables estimates and find that class size does have an effect on test scores. Their method exploits Israel's 40 student class-size cap and the discontinuity which means the 41st student enrolment causes the (average) class size to fall to 20.5. Schools with a cohort of just over 40 (or a multiple of 40) students have significant better results than those with cohorts of just under the cap.

More contemporary studies side with these conclusions and find positive effects of expenditure. Jackson, Johnson and Perisco (2016) use American school finance reforms throughout the 1970s and 1980s as exogenous changes in schools' expenditures and compare adult outcomes of students based on when and where they went to school, using a difference-in-difference method as the reforms were implemented in different parts of the country at different times. They conclude that this growth in school spending increased future earnings, particularly for poor students, with significant, robust results. They give a specific estimate of a 10% increase in spending for all school-age years yielding a 7.7% increase in wages. Hyman (2017) takes a similar approach, focusing specifically on Michigan's 1994 reform, and relates an exogenous increase in spending with university enrolment and university completion by using individual-level student data, the first study to do this beyond high school (Hyman, 2017). He concludes a 10% increase in spending causes a 7% increase in enrolment and an 11% increase in completion rates. Strayer (2002) also highlights the indirect effects of schooling quality on wages. He first finds that better schooling quality increases the likelihood of attending university, and specifically attending better universities. Secondly, and more relevant to my research, he finds that attending university increases incomes and in turn attending better universities has further income gains. His analysis suggests that this indirect effect of improved schooling is at least as important as the direct effect of higher quality schooling increasing skills. A paper which looks at more modern data is working paper by Jackson, et al. (2018). They take states' differing exposure to spending cuts during The Great Recession, such as whether a state relies on revenue from taxes which are more sensitive to the business cycle, as exogenous differences and find that cohorts more exposed to spending cuts had lower test scores and lower college-going rates: evidence that expenditure does matter.

The school-focused literature leans towards quality having a significant effect on student outcomes, particularly in contemporary papers, and, whilst sparser, the literature regarding university quality and outcomes follows the same pattern. One of the most significant studies focused on university quality and incomes is by James, et al. (1989), and it is this paper that I draw upon most in this study. The authors run regressions on future earnings using explanatory variables including sets of student characteristics, institutional characteristics (including expenditure), higher educational experience variables (for example, what subject a student studies), and labour market variables. Relevant to my research, they found institutional characteristics to be largely insignificant, explaining 1-2% of the variance. In all their regressions they found expenditure to be insignificant, with selectivity and an indicator of prestige being more significant, the latter of which they put down to peer group and/or halo effects.

There are several papers, however, that find positive effects of expenditure. Dale & Krueger (2002) use a matching technique by comparing students with the same pattern of college acceptances and rejections to control for unobserved characteristics relating to ability. They claim that for students with the same pattern, attending a more selective university has little impact on future earnings, but a higher expenditure per student does, contradicting James, et al. (1989)'s findings. Elsewhere, Behrman, et al. (1996) study twins to conclude that expenditure per student is generally significant across a range of regressions. This study particularly highlights the effect specification has on results with the significance of variables varying greatly depending on their model set-up. They also comment on the effect differing tuition fees has on graduates' incomes (p.682): this variation in fees is much less prevalent in the UK, with implications for my study which I mention in the relationships section. Finally, an older but influential study by Wachtel (1976), which jointly considers an individual's school and university experiences finds significant positive effects of expenditure at school, undergraduate and postgraduate levels, with the effect of school expenditures being greatest, followed by postgraduate then undergraduate spending.

The research into this quality-outcome relationship is poorly studied at a university level with no major papers examining the issue in the last decade. The importance of developing an understanding of this relationship is elevated due the current political climate: higher education funding is increasingly under scrutiny, with major changes seen in the UK in the 2010's; and inequality is an increasing prominent issue in politics, with higher education having the potential to be part of the problem and/or solution. Not only does my research look to provide a contemporary update to this literature, but it also seeks to add a different perspective. Almost all the aforementioned studies have used American data. Giving a British perspective might uncover a different relationship between quality and outcomes, for example due to intangible cultural factors or due to differences in the countries' education systems. Finally, unlike most these studies, I use individual universities as my unit of measure, as opposed to individual students. There are strengths and weaknesses to this approach. It will be harder to control for student characteristics, as I discuss below, but I will be able to give a better overview of the whole population of universities in the UK, which is perhaps of more relevance to prospective students of these universities and those associated with them. Throughout my study I draw on the lessons of the literature, especially regarding my model specification.

Methodology

Data Sources

I calculate each university's average income from the Higher Education Statistics Agency's (HESA) Higher Education Graduate Outcomes data, published in June 2020. This cross-sectional data includes information on the salaries of the 2017/18 graduating cohort by which university they attended, collected via surveys fifteen months after graduation. I took several steps to prepare this data for analysis, which I overview here (a more detailed and formalised explanation of how the dataset was prepared and how average income was calculated can be found in Appendix A). Firstly, a range of different details to categorise graduates are listed. I filtered this data to give only the salaries of graduates who previously studied full-time and were now in paid employment. This is for simplicity and comes with the caveat that my research is only applicable to such graduates. The implications of this are further considered in the conclusion.

Secondly, the data does not provide a precise salary for each individual, rather it counts the number of students per institution with a salary in each £3000 bands. These bands start at <£15,000 followed by £15,000-£17,999 and continuing up to £39,000+. I have taken the middle value of each band, for example, the £27,000 - £29,999 is given as £28,500, and taken the outer bands as £15,000 and £39,000. This value is multiplied by the number of observations in each band to give the total income[3] for that band. The total income of each band is added together to get the total income of all bands. This is then divided by the total number of observations to give the average income of graduates at that university.

It is not a perfect measure. HESA's banding of the salaries masks variation within the bands, especially in the end bands where it is not possible to obtain a middle value. My method for calculating each university's average income has two implicit assumptions for it to be an accurate measure. Firstly, I am assuming that within each band, the mean income is the centre-point of the band and secondly, that the distributions within the end bands are identical so they cancel each other out. The latter assumption is likely too strong: minimum wages will keep pay close to the upper limit of the lowest band, but salaries could far exceed the lower limit of the highest band. However, counts in these bands (especially the highest band) are relatively low, reducing the effect of any inaccuracies this causes.

The former assumption is more reasonable, especially with larger sample sizes within each university where the mean will converge to the centre-point. Because of this, I have limited the population of universities to those with a sample of 100 or more. This results in the original population of 305 reducing to 132. Of those to drop out of the population, very few are universities; rather they are mostly further education colleges. These are not the subject of my research. The few universities that do drop out are generally highly specialised institutions focusing on postgraduate study, for example The University of Law. I do not believe such dropouts constitute a selection bias as I still have a full population of the UK's primarily undergraduate universities which are the focus of my study.

I have manually gathered data on universities' expenditures from their financial statements for the year ending 31st July 2018, thus the data relates to the graduates' final year and is implied as a proxy for expenditure throughout their degree. In my analysis I consider four

[3] Whilst salaries and incomes are two different concepts, I take them to be interchangeable here, referring to salaries when discussing the raw data and incomes once the salary bands are aggregated/averaged. I assume the difference between graduates' salaries and incomes to be negligible as low asset ownership amongst recent graduates will mean most of their income will come from labour.

measures of expenditure (all per student). Total expenditure and staff costs are the two broadest measures and are measured and reported in a standard manner for all 132 observations. Academic department expenditure and academic staff expenditure are more targeted measures but are not available for all universities and even where reported, they are less standardised and less comparable between universities. I have chosen staff costs per student as my variable of choice after trading off three factors: relevance, sample sizes, and model performance. The academic-specific measures capture teaching quality best and, in my models, had marginally better R-squared values, but they have smaller sample sizes which inflate the standard errors. Furthermore, there is likely to be a selection bias with the academic measures because whether a university provides such a detailed breakdown of is expenditures appears non-random: smaller universities were less likely to report such data. Total expenditure and staff costs do not have this problem as they make use of the full population of universities, and their models perform comparably well. Within the two, staff costs seem more targeted as it does not include a large proportion of university spending, such as marketing, that has no direct impact on quality. Because of this, I use staff costs per student as my main variable for expenditure, but still use the three other measures as a robustness check.

Supplementary data is gathered from a variety of sources. Some data, such as the location of a university or whether it is a Russell Group university, is simply factual. Data on student-staff ratios and entry standards are taken from The Guardian's 2018 university league table (The Guardian, 2017). Whether a subject is offered at a university is taken from The Guardian's subject-specific league tables: if there is an entry for that university in a subject table, it is taken that the subject is offered there. Data on student numbers in the 2017/18 academic year, used in per student calculations and the proportion of postgraduates, come from a different HESA dataset (Higher Education Statistics Authority, 2017).

A further dataset is used to calculate graduate incomes in robustness checks for subject studied. The Longitudinal Education Outcomes (LEO) graduate outcomes data (Department for Education, 2020) reports the median income of graduates by subject and university; the strengths and weaknesses of which are discussed later.

Relationships

Figure 1 shows a scatterplot of average income by staff costs per student and reveals an upwards sloping relationship as well as three universities (Cambridge, Oxford, and Imperial) having substantially greater per student expenditures than others. The fit line is weighted by the sample size which has gone into generating the average income estimate for each university (and is so in all figures). The presence of these three outliers reduces the linearity of the relationship, so I will also estimate a model with a quadratic term on expenditure, visualised in Figure 2. The relationship between average income and student-staff ratio is plotted in Figure 3, where a quadratic term appears less justified. Finally, in Figure 4, the relationship between these two quality measures is visualised. As expected, there is a negative relationship between staff costs per student and student-staff ratios: to lower the student-staff ratio, more teachers need to be hired, increasing staff costs. This may cause problems regarding multicollinearity: the two quality variables have a correlation of -0.64. I re-run the regressions in my primary model using each of these quality variables without the other. The results are largely unaffected and so I do not report them.

Figure 1 – Scatter Plot of Average Income and Staff Costs (Linear Fit)

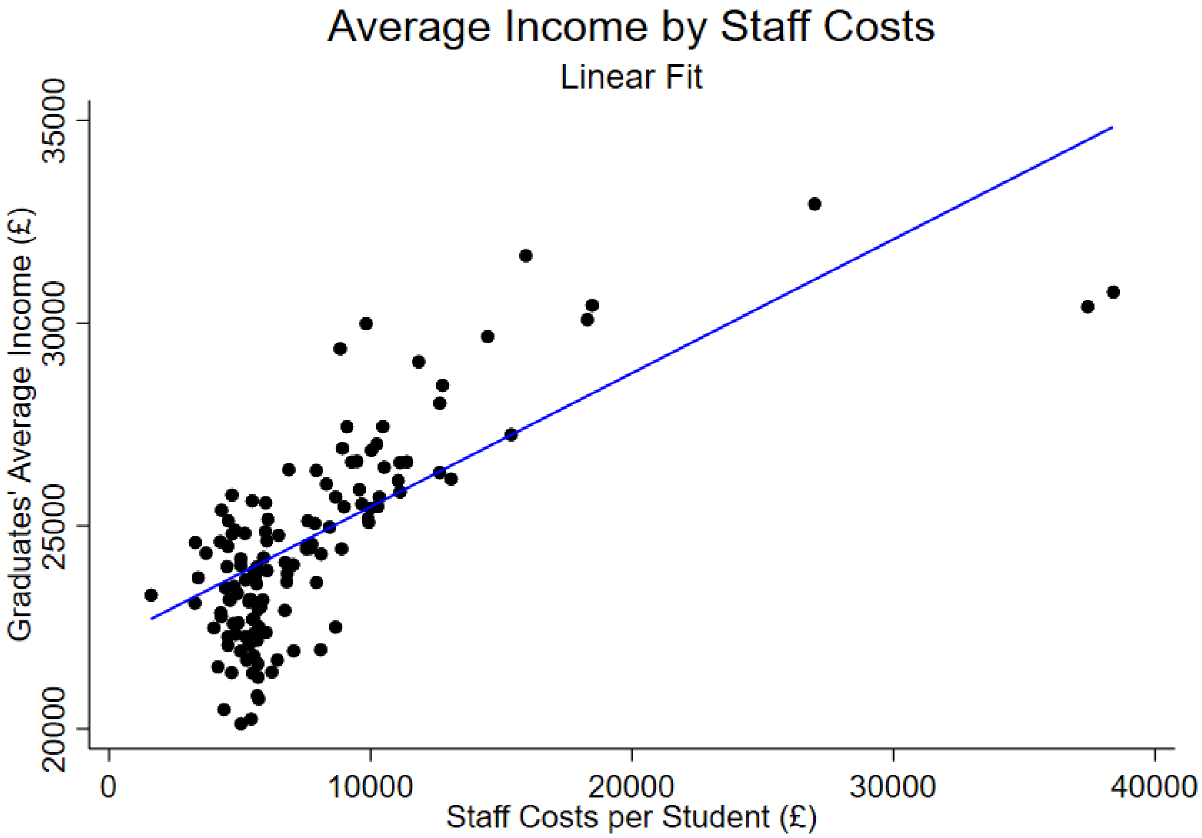

Figure 2 – Scatter Plot of Average Income and Staff Costs (Quadratic Fit)

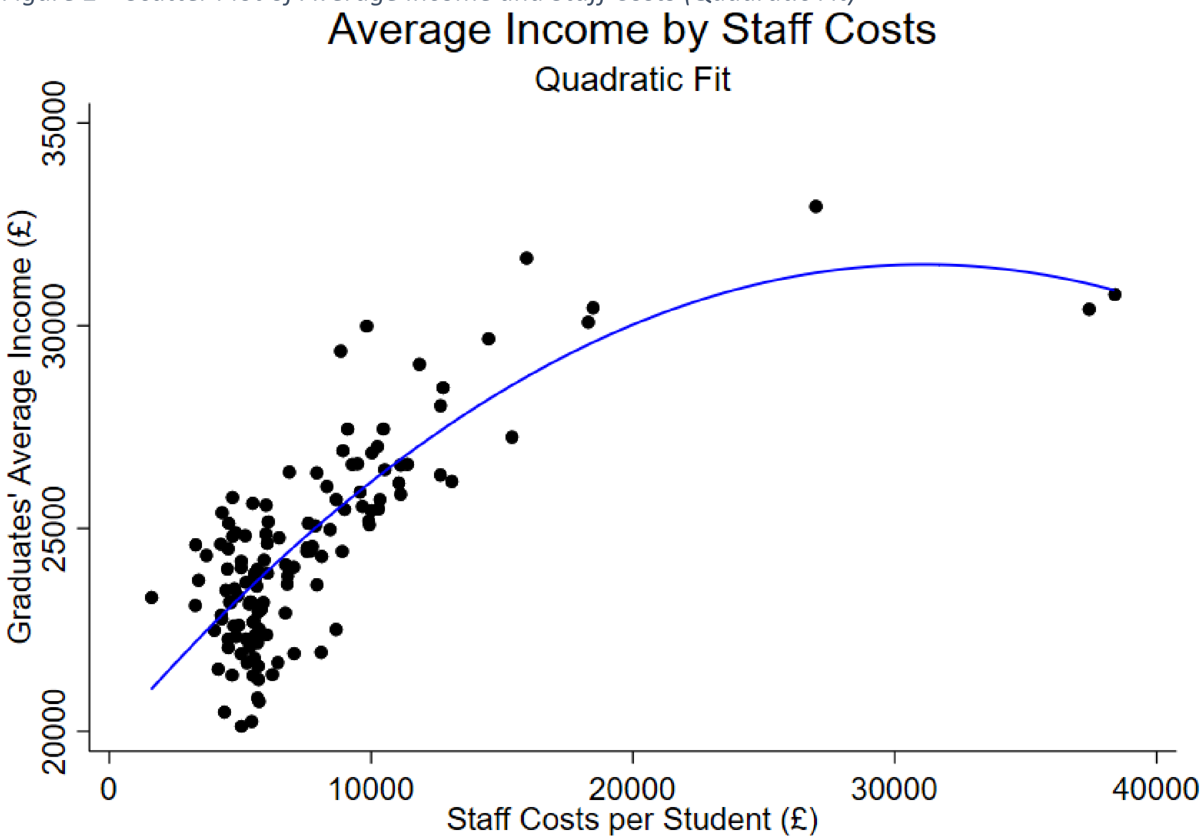

Figure 3 – Scatter Plot of Average Income and Student-Staff Ratio (Linear Fit)

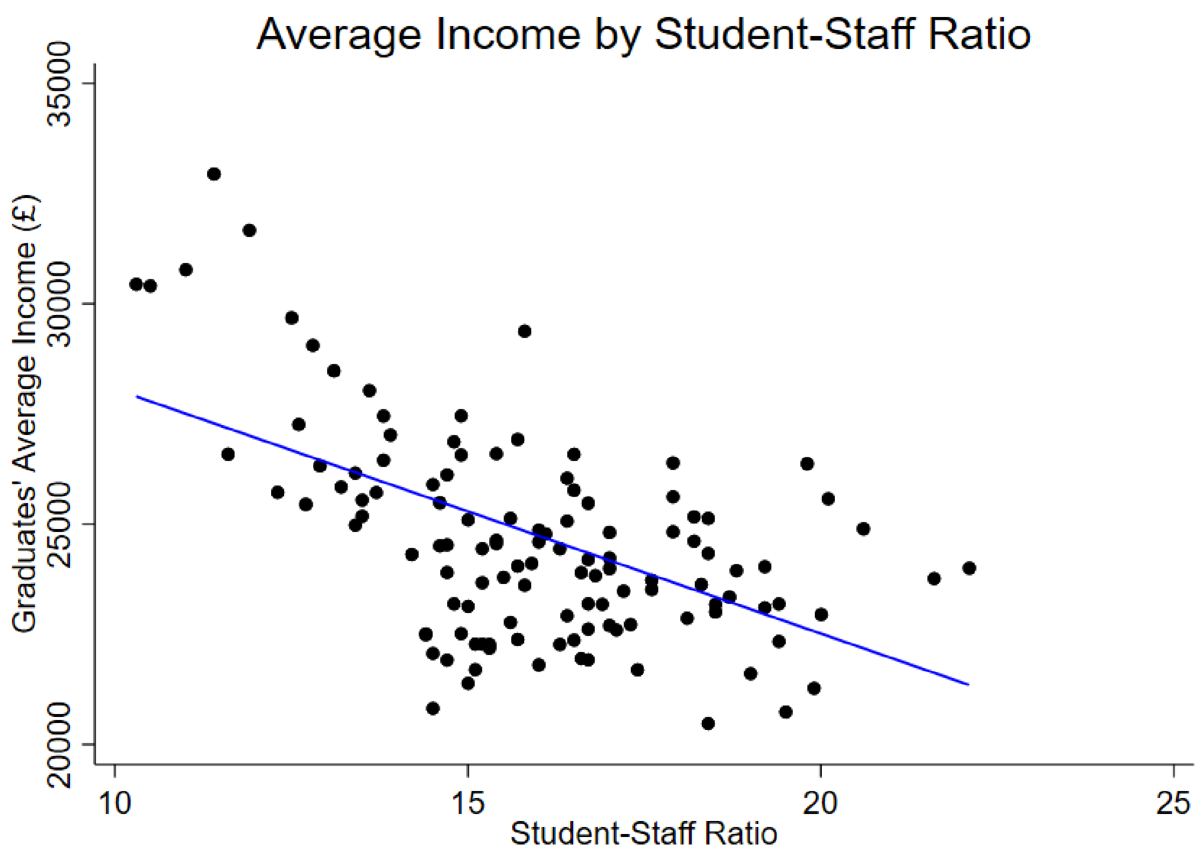

Figure 4 – Scatter Plot of Staff Costs and Student-Staff Ratio (Linear Fit)

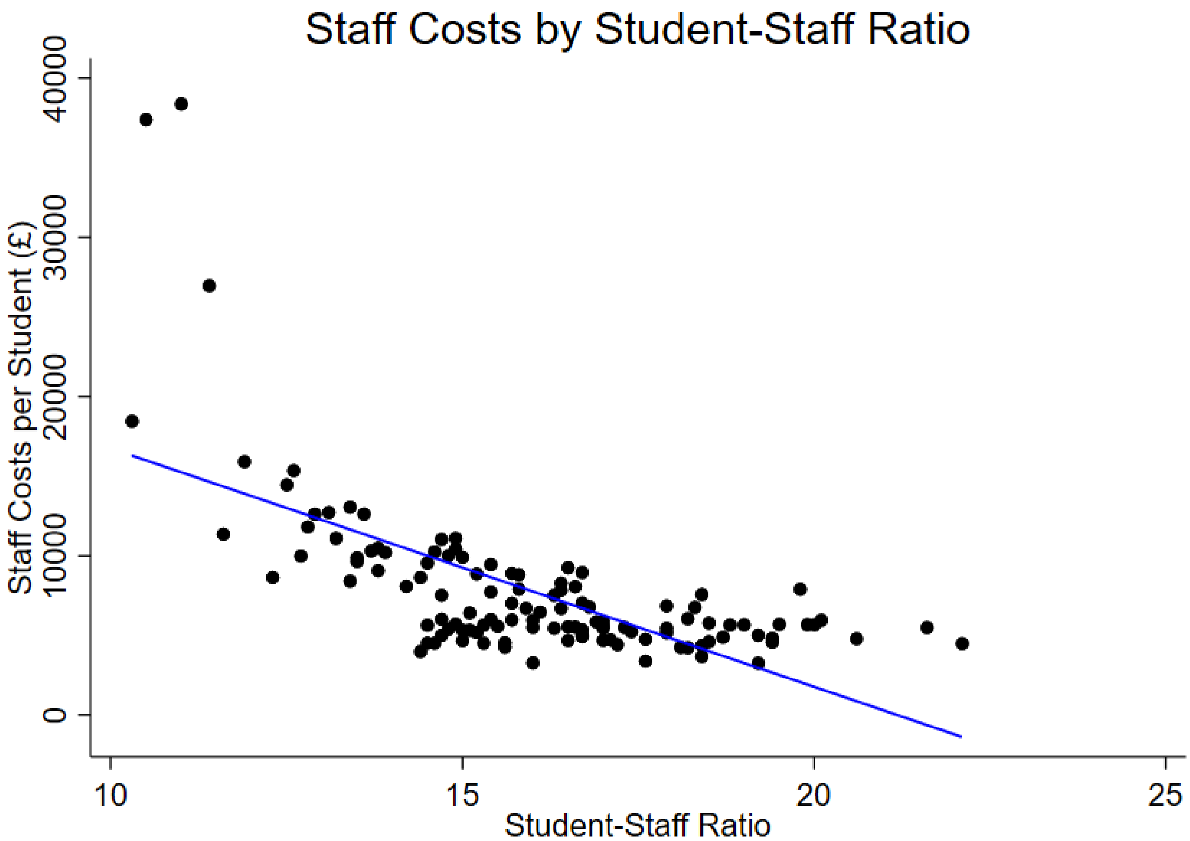

Figure 4 shows that there are the same three major outliers on the expenditure variable, suggesting these universities pay their staff substantially more, as well as having proportionally more of them. A brief comparison of these universities' income structures, shown in Table 1, offers some insight into why they are so different.

Table 1 – 'Outlier' University Income Analysis

University	Total Income	Tuition Fees	Funding Grants	Research Grants	Other Income	Investment Income	Endowments
Warwick	£0.62B	51%	10%	20%	18%	0%*	0%*
UCL	£1.43B	36%	16%	33%	12%	0%*	1%
Durham	£0.36B	54%	11%	15%	17%	2%	2%
Basket Average		**47%**	**12%**	**23%**	**16%**	**1%**	**1%**
Cambridge	£1.82B	16%	10%	28%	42%	1%	3%
Oxford	£2.01B	16%	9%	29%	40%	1%	5%
Imperial	£1.03B	29%	15%	35%	14%	1%	6%
Outlier Average		**20%**	**11%**	**31%**	**32%**	**1%**	**5%**

* rounding

Compared to a basket of similar (roughly next best) universities, the outliers receive a larger proportion of their incomes away from tuition fees. The most dramatic cause for this difference is Cambridge and Oxford's large publishing and examination revenues (included under "other income"). Imperial is less obviously different with only its donation and endowment income distinguishing it from UCL. It is worth noting that the typical university generally has a much larger proportion of its income coming from tuition fees than all those listed here. In light of these three universities being substantially different, I re-run all my regressions without those observations included. This makes little difference to any of the results and so I do not report them in this paper.

Another issue this raises is the potential for reverse causality, which would occur if universities with the highest earning graduates generate higher incomes and therefore have higher expenditures. Relative to doing this study is America, for example, using British data has an advantage here. Most undergraduate tuition fees are identical for British students across universities (currently £9,250/year), so the channel of wealthier students, who would potentially earn more anyway, paying higher tuition fees does not operate here. That channel could still exist with endowments replacing tuition fees, but outside the three outlier universities, this effect would be small.

Model

The model I will estimate takes the following form:

$$Average\ Income_i = \alpha + \beta_1 Expenditure_i + \beta_2 Student - Staff\ Ratio_i + \beta_3 \boldsymbol{X}_i + \mu_i$$

where $\boldsymbol{X}_i$ is a set of control variables, i is a university, and μ is the error term. The dependent variable is the average income of graduates from university i and the explanatory variables of interest are university i's expenditure per student (measured by staff costs per student) and its student-to-staff ratio (calculated by The Guardian (2017) for staff with significant teaching time). I am primarily interested in the direction of β_1 and β_2 and whether they are statistically significant or not. The exact value of the coefficients is not important in answering my research question.

I estimate the model using a simple ordinary least squares (OLS) regression, controlling for a range of university-specific factors that might otherwise bias the results. My choice of control variables and my reasoning for including them is largely motivated by James, et al.'s (1989) research. I control for the number of students at each university to allow for (dis)economies of scale; the proportion of students that are postgraduate (in case postgraduate teaching either 'distracts' from undergraduate teaching, or alternatively offers a pool of cheap teaching labour), and whether or not a university is in the Russell Group of 'elite' research-intensive universities to control for either prestige or research detracting from the teaching function. The most important control, especially with regards to controlling for endogeneity, is entry standards, which I measure by average UCAS entry tariff[4], as students with high-achievement pre-university will likely earn relatively more regardless of where or if they go to university. I also include two dummy variables for university location: one for London as

[4] More detailed explanations of how entry standards and student-staff ratios are calculated is available at https://www.theguardian.com/education/2016/may/23/methodology-behind-the-guardian-university-guide-2017.

graduates of London-based universities are more likely to work there, where wages are higher, after graduating; and one for Scotland due to a different environment regarding tuition fees potentially affecting the higher education dynamics there. Finally, I attempt to control for subject studied by including dummy variables for whether engineering, economics and/or medicine are offered at each university. These subjects are chosen as their graduates' have generally higher incomes than other subjects on average (Department for Education, 2016). All regression analysis is performed with robust standard errors to avoid issues relating to heteroskedasticity.

Summary Statistics

Summary statistics for the variables I am using are detailed in Table 2.

Table 2 – Summary Statistics

Variable	Mean	Std. Dev	Min	Max	Observations
Average Income	24,500	2450	20,130	32,950	132
Observations (per University)	706	437	100	1975	132
Staff Costs per Student	7720	5100	1601	38,390	132
Student-Staff Ratio	15.9	2.3	10.3	22.1	120
Number of Students	16,500	8960	1190	40,140	132
Entry Standards	140	28	95	224	120
Proportion Postgraduate	0.226	0.092	0.022	0.584	132
London	0.174	0.381	0	1	132
Scotland	0.114	0.319	0	1	132
Russell Group	0.189	0.393	0	1	132
Medicine Offered	0.258	0.439	0	1	132
Economics Offered	0.629	0.485	0	1	132
Engineering Offered	0.720	0.451	0	1	132

The only variable listed in Table 2 yet to be mentioned is observations per university. This refers to the number of observations that have been used in estimating each university's average income. Having more observations results in a more accurate estimate for average income and so in the regression analysis I weight the observations to account for the differing confidence in the accuracy of average incomes. The datapoints for University of Nottingham (n=1975) will be given more weight than those for SOAS (n=100).

Note that the student-staff ratio and entry standards variables have fewer observations than the full population. This is due to The Guardian's league table not offering these figures for all universities. As such, my sample size for the regressions is only 120 universities. The twelve universities which drop out are mostly specialised universities (arts, veterinary, agriculture), which were large enough to get through the first filtering. However, some universities which I would have preferred to include also drop out, such as Ravensbourne University London and the University of the Highlands and Islands. These other dropouts are likely non-random: they are generally smaller universities. This may cause a selection bias; however, the effect is likely to be small as there are only a few of these universities.

Results

Primary Model

The model is estimated four times under slightly different specifications. All models have average income as the dependent variable and staff costs per student and student-staff ratio as the independent variables of interest. The control variables listed above are included in all models, except Model 1, which does not include the subject dummies. Model 3 adds a quadratic term on staff costs, and Model 4 adds another one for student-staff ratio. Table 3 provides the coefficient on each variable for all four of the models.

Table 3 – Primary Results

Variable	Model			
	1	2	3	4
Staff Costs per Student	0.081** (0.036)	0.076** (0.035)	0.386*** (0.129)	0.364*** (0.130)
Student-Staff Ratio	54.6 (65.3)	27.6 (69.1)	74.6 (69.4)	-540.8 (494.5)
(Staff Costs)^2			$-6.31e^{-6}$** ($2.51e^{-6}$)	$-6.30\ e^{-6}$** ($2.48\ e^{-6}$)
(SSR)^2				18.3 (14.3)
R^2	0.7701	0.7910	0.8026	0.8041
Includes Subject Dummies?	No	Yes	Yes	Yes

*Standard errors in parentheses; n = 120; * = p < 0.10, ** = p < 0.05, *** = p < 0.01*

Staff costs per student are statistically significant across all models and the coefficients are positive. These preliminary findings suggest a greater expenditure per student increases students' incomes post-university. Student-staff ratios, on the other hand, have no statistically significant effect in any of the models. As expected, the quadratic term for staff costs is important whilst the one on student-staff ratio is less so. The full results table, including results on the control variables can be found in Appendix B.

Expenditure Variable Robustness

As discussed previously, whilst staff costs per student is the most useful measure given my data, there are alternative measures for expenditure which I consider. Re-running Model 4 with staff costs per student and its quadratic term replaced with the alternative measures (all per student) generates similar results, shown in Table 4.

Table 4 – Expenditure Variables Robustness Results

Expenditure Variable	Coefficient	P-value	R^2	Sample Size	% Covered
Total Expenditure***	0.210	0.000	0.8115	120	100
Staff Costs***	0.364	0.006	0.8041	120	100
Academic Department Expenditure***	0.458	0.002	0.8362	88	73
Academic Staff Costs*	0.874	0.077	0.9154	40	33

Total expenditure and academic department expenditure, like staff costs, are statistically significant at the 1% level, whilst academic staff costs are also significant, albeit only at the 10% level. This is likely due to the smaller sample size inflating standard errors with this measure.

An initial conclusion from this model is that university expenditure per student has a positive effect on students' earnings as graduates, robust to different measures of expenditure. However, this model has a major weakness regarding controlling for subject. The subject studied has a major effect on income and different universities have different proportions of students studying different subjects. James, et al. (1989) conclude that what you study is, under certain circumstances, more important than where you study. In my primary model I have only controlled for subjects studied by including three dummy variables for whether a university offers a course in that subject. This is clearly inadequate, as suggested by the smaller-than-expected increase in R-squared between Models 1 and 2. The dummy variables do not distinguish between universities that simply offer one of these high-paying subjects and those that specialise in them; consider a large generalist university offering economics amongst a range of other subjects and the London School of Economics, which specialises in it. The failure to fully control for subject studied will result in biased estimates on the explanatory variables of interest.

LEO Model

Ideally, the primary model would be adjusted to include data on the proportion of students at each university studying certain subjects. This data is hard to find and is often not publicly reported. I instead make use of a different dataset to correct for this problem. As mentioned in the data section, I use LEO graduate outcomes data for the same year group as the HESA data to provide my dependent variable, average income. This dataset reports the median income for different subjects within each university, overcoming the previous issues regarding subject studied. However, there are several weaknesses which explain why I am only using this data for supplementary analysis. Firstly, I cannot directly answer my research question using this dataset: universities report expenditure for the whole university rather than breaking it down by subject. Unless expenditure per student is uniformly distributed across subjects, it will be an inaccurate measure. Secondly, the sample sizes are smaller both across the model as a whole, due to specific subjects only being offered by some universities;

and within each university, with fewer observations contributing to generating the average income estimate. There are further differences between the HESA and LEO data, for example LEO data is gathered one year after graduation versus HESA's fifteen months, but these are not significant differences.

To make use of the LEO data, I chose eight different subjects from the LEO dataset. These were chosen as they are studied at many universities and so have relatively large sample sizes; they corresponded well to The Guardian's subject-specific league tables; and covered a broad range of subjects to allow for heterogenous effects of quality to emerge, for example is quality was only important for more practical subjects. I replaced the overall average income for each university with a subject-specific average income for these eight subjects. The student-staff ratio and entry standards variables were also changed as The Guardian's subject-specific league tables reported these figures by subject, as well as for the university as a whole.

Initially, I re-run the regression for Model 4 eight times, once for each subject. Despite it likely being a weak measure, I include university-wide staff costs to see if it has an effect despite not being subject-specific; excluding it makes little difference to student-staff ratio results. The results on these quality variables are shown in Table 5.

Table 5 – LEO Model Results

Subject	Measure	Coefficient	P-value	R^2	Sample Size
Biological Sciences	Staff Costs per Student	0.034	0.832	0.6121	92
	Student-Staff Ratio	-26.7	0.928		
Computing	Staff Costs per Student*	0.802	0.057	0.6758	96
	Student-Staff Ratio	428	0.287		
Economics	Staff Costs per Student	0.205	0.555	0.8042	54
	Student-Staff Ratio	28.3	0.621		
Engineering	Staff Costs per Student	0.626	0.174	0.5547	60
	Student-Staff Ratio	30.7	0.961		
English Studies	Staff Costs per Student	0.275	0.172	0.4591	94
	Student-Staff Ratio	121	0.694		
Languages	Staff Costs per Student	0.219	0.417	0.5178	52
	Student-Staff Ratio	-629	0.546		
Law	Staff Costs per Student	0.034	0.871	0.6982	91
	Student-Staff Ratio	-2187	0.147		
Nursing	Staff Costs per Student	-0.086	0.787	0.5505	66
	Student-Staff Ratio	-1263	0.170		

As when using the HESA data, there is no significant effect of student-staff ratios on incomes. More interestingly, using subject-specific LEO data results in the staff costs variable becoming generally insignificant. This could be due to several reasons including a non-uniform distribution across subjects making it a weak measure, smaller sample sizes inflating standard errors, or as a result of better controlling for subjects and providing a less biased estimate than the primary model.

REF Model

Staff costs per student's strength as a measure relative to student-staff ratios is that it captures the number of staff, like SSR, but also captures their quality, assuming a higher wage is paid to higher quality staff. A suitable subject-specific replacement for staff costs could be Research Excellence Framework (REF) ratings. The REF is a research impact evaluation for British universities to determine future research funding allocation. Staff submit recent research for evaluation which is graded along a five-point scale ranging from 4* (world-leading) and 3* (internationally excellent) through to unclassified, which is then aggregated at department level to give different proportions of research at each grade. I use the proportion of a department's research at these highest grades as a measure for staff quality, which student-staff ratio fails to capture. Furthermore, De Fraja, et al. (2019) find a robust positive relationship between REF scores and professorial wages, strengthening the case for REF being a strong substitute for staff costs per student as a higher REF scores will be associated with higher average staff costs.

Replacing staff costs in the above regressions with the proportion of research graded at 4* returns the following results in Table 6a. Running the model with the proportion graded at 4* or 3* returns similar results, which can be found in Appendix B, Table 6b.

Table 6a – LEO Model REF (4) Results*

Subject	Measure	Coefficient	P-value	R^2	Sample Size
Biological Sciences	REF 4	-2272	0.164	0.7888	41
	SSR**	-1075	0.010		
Computing	REF 4**	8877	0.021	0.7184	80
	SSR	633	0.141		
Economics	REF 4	368	0.916	0.8680	25
	SSR	468	0.396		
Engineering	REF 4	9027	0.404	0.7454	15
	SSR	-404	0.800		
English Studies	REF 4	-1581	0.233	0.4483	81
	SSR	-153	0.650		
Languages	REF 4	-3403	0.198	0.5670	37
	SSR	-1423	0.363		
Law	REF 4	1763	0.474	0.7504	61
	SSR***	-5994	0.000		
Nursing	REF 4	600	0.709	0.5715	56
	SSR	-561	0.448		

Despite some significant results, neither quality measure has a consistent effect on incomes. This is perhaps due to small sample sizes: not all departments had REF results and those that did were generally research-focused universities whose graduates earn more regardless, decreasing dependent variable variation. Alternatively, research quality might be unrelated to teaching quality, making REF a poor measure. The more targeted Teaching Excellence Framework is not an appropriate measure as it is not graded at department level. Subject heterogeneities could also explain why some results are significant and others are not. However, based on this data alone it is hard to make such a case: there is no obvious reason why student-staff ratio would matter for Biology and Law but not for other subjects. Overall, using OLS on these subject-specific LEO data returns no consistent evidence to support the findings of my primary model.

Fixed Effects Model

A final robustness test involves exploiting the structure of the LEO data to control for subject fixed effects. This allows me to study the effect of quality differences between universities within a given subject, without losing power in the regressions from having smaller sample sizes. By combining the subjects together and treating them as static panel data, the differences between university observations within each subject allow inferences to be drawn without needing to control for the subject-makeup of different universities; a problem which biased my primary model.

Using the same control variables, with SSR and REF quality measures, returns the results shown in Tables 6a and 6b, the full versions of which are reported in Appendix B, Tables 6c and 6d.

Table 7a – Fixed Effects Results (4)*

Variable	Coefficient	Std. Error	t	P>|t|
SSR	-13.0	34.0	-0.38	0.714
REF 4* Proportion	2216	1544	1.44	0.194
Observations	396			
Groups (subject)	8			
R^2 (within)	0.5480			
rho	0.90			

Table 7b – Fixed Effects Results (4 & 3*)*

Variable	Coefficient	Std. Error	t	P>|t|
SSR	-13.5	31.0	-0.44	0.676
REF 4/3* Proportion	1546	1374	1.13	0.298
Observations	396			
Groups (subject)	8			
R^2 (within)	0.5481			
rho	0.90			

This final robustness test also fails to support the initial findings of a significant positive effect of quality on graduates' incomes. Both student-staff ratios and REF scores have statistically insignificant effects on graduates' earnings: within the eight subjects in my dataset, these quality measures cannot explain differences in average income from different universities. The rho values are worth commenting on; they show 90% of the total error variance comes from variance of the subject fixed effects. This implies subject studied has a proportionally large effect, relative university attended, which provides further evidence that the weakness of subject controls in the primary model is likely to make its results biased.

Conclusion

The existing literature is divided as to whether quality factors affect student outcomes. Using data about the incomes of the 2017/18 graduating cohort from UK universities, I fail to find conclusive evidence that universities' expenditures affect graduates' incomes. Whilst my primary model finds a statistically significant effect, robust to different measures of expenditure, there is good reason to believe this result is a product of omitted variable bias relating to subjects studied. Whilst the robustness checks do not involve results on a suitable measure of university expenditure, the REF scores measure incorporates the key features of staff costs per student and acts as a suitable enough substitute to cast serious doubt on my initial results. Once subject-studied is fully controlled for I find no evidence of a significant effect of university expenditure on graduates' incomes.

The primary focus of my research has been studying the effect of university expenditure on graduates' incomes. This is a specific form of the wider question: does education quality have an effect on student outcomes? As I have only examined a specific form of this question my research only offers partial evidence to this wider question. However, combined with the consistent insignificance of student-staff ratios, my research fits into the existing literature on the side of education quality having no effect on student outcomes.

It is perhaps unsurprising that my research falls on this side of the debate. My methodology, particularly the variables included in my regressions, is most similar to James, et al. (1989) and I produce similar findings to them, including mirroring their finding on the importance of subject studied, given the rho values in my fixed effect models. More generally, research using simple econometric techniques tends to find no effect of quality, like I have done. Other methods, most significantly using natural experiments and difference-in-differences techniques, find significant effects of quality more frequently.

There are some significant limitations to my study. With the initial model there was likely omitted variable bias due to failing to control for subjects studied. Whilst I believe my robustness checks overcome this problem, the lack of subject-specific expenditure data means I have not been able to directly test the university expenditure data in a sufficiently unbiased model. As such, my research could be improved upon with better data: either by including better control variables for subjects in the primary model, or by including subject-specific expenditures in the fixed effects model. Such data do exist, however, not all universities publicly report these figures and when they do, they are not easy to find. Further research should look to develop and use these data sources further.

Within the data I do have, there are further limitations. Firstly, this data only considers the salaries of graduates who previously studied full-time and were now in paid employment. This mean it does not account for part-time students nor does it factor in graduates' chances of being employed. Focusing only on full-time students is not a major weakness: most undergraduate students study full-time, and any inferences drawn about them can likely be applied to part-time students as well, though specific research into the quality-outcome relationship for part-time students would be needed to confirm it. Only considering graduates in paid employment, however, is a significant limitation of this study. If there were a positive relationship between university expenditure and graduates' outcomes, this would be captured not only in higher incomes, as studied here, but also in higher employment rates. There are ways to account for this, for example by multiplying average income by the proportion of graduates in employment or further education, however, such measures bring up further considerations such as how to account for graduates who work as volunteers or are voluntarily taking time out. For simplicity I have not accounted for employment prospects, but it should be considered that average income is not a perfect measure for student outcomes, even when only considering financial ones.

My finding that university expenditure does not significantly affect graduates' earnings raises several implications. Perhaps prospective students focused on their future earnings should not be concerned with the measures of quality reported in league tables, such as student-staff ratios, and focus instead on the location of potential universities or the number of students at a university (these variables had significant effects in the fixed effects model). These findings also question whether the marketisation of the university sector benefits students. With revenue generation a more pressing priority for universities under the new policy environment it could be argued that students will benefit from subsequent higher spending. My findings suggest this will not be the case. Instead, it could be argued that growing universities, given the removal of the student numbers cap, could decrease earnings, given the statistically significant ($P<0.01$) negative coefficient on total number of students. On the other hand, given student-staff ratios are generally insignificant throughout my results, larger within subject cohort sizes will be unlikely to negatively impact students' future earnings.

However, these conclusions are only as strong as the data and methodologies used to draw them. More research is needed into UK higher education outcomes to verify these findings. Future research could focus on more detailed subject data or school-specific expenditure data to improve the models using the methodology I have used here; or by using a different methodology, such as by generating individual-level data or by using time-series analysis. It will be interesting to revisit this analysis once future waves of HESA data are released.

References

Angrist, J. & Lavy, V., 1999. Using Maimonides' Rule To Estimate The Effect of Class Size on Scholastic Achievement. *The Quarterly Journal of Economics,* 114(2), pp. 533-575.

Behrman, J. R., Rosenzweig, M. R. & Taubman, P., 1996. College Choice and Wages: Estimates Using Data on Female Twins. *The Review of Economics and Statistics,* 78(4), pp. 672-685.

Card, D. & Krueger, A. B., 1992. Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy,* 100(11), pp. 1-40.

Card, D. & Krueger, A. B., 1996. School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina. *The Journal of Economic Perspectives,* 10(4), pp. 31-50.

Coleman, J. S., 1966. *Equality of Educational Opportunity,* Washington: Government Printing Office.

Dale, S. B. & Krueger, A. B., 2002. Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables. *Quarterly Journal of Economics,* 117(4), pp. 1491-1527.

De Fraja, G., Facchini, G. & Gathergood, J., 2019. Academic salaries and public evaluation of university research: Evidence from the UK Research Excellence Framework. *Economic Policy,* 34(99), pp. 523-583.

Department for Education, 2016. *Employment and earnings outcomes of higher education graduates: experimental statistics using the Longitudinal Education Outcomes (LEO) data: further breakdowns.* [Online]
Available at:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/573831/SFR60_2016_LEO_main_text_v1.1.pdf

Department for Education, 2020. *Graduate outcomes (LEO): 2017 to 2018.* [Online]
Available at: https://www.gov.uk/government/statistics/graduate-outcomes-leo-2017-to-2018

Hamilton, N., 2016. The Coalition's higher education reforms in England. *Oxford Review of Education,* 42(3), pp. 330-345.

Hanushek, E. A., 1986. The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature,* 24(3), pp. 1141-1177.

Hanushek, E. A., 1996. School Resources and Student Performance. In: G. Burtless, ed. *Does Money Matter? The Effect of School Resource on Student Achievement and Adult Success.* Washington, D.C.: Brookings Institution, pp. 43-73.

Hedges, L. V., Laine, R. & Greenwald, R., 1994. Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes. *Education Researcher,* 23(3), pp. 5-14.

Higher Education Statistics Authority, 2017. *HE student enrolments by HE provider, domicile, level of study, mode of study, first year marker, sex and academic year.* [Online]
Available at: https://www.hesa.ac.uk/data-and-analysis/students/table-1-(2017-18).csv

Higher Education Statistics Authority, 2020. *HE Graduate Outcomes Data.* [Online]
Available at: https://www.hesa.ac.uk/data-and-analysis/graduates/salaries

Hyman, J., 2017. Does Money Matter in the Long Run? Effects of School Spending on Educational Attainment. *American Economic Journal: Economc Policy,* 9(4), pp. 256-280.

Jackson, C. K., Johnson, R. C. & Perisco, C., 2016. The Effects of School Spending on Educational Outcomes: Evidence from School Finance Reforms. *The Quarterly Journal of Economcs,* 131(1), pp. 157-218.

Jackson, C. K., Wigger, C. & Xiong, H., 2018. Do School Spending Cuts Matter? Evidence From The Great Recession. *NBER Working Paper No.24203.*

James, E., Alsalam, N., Conaty, J. C. & To, D.-L., 1989. College Quality and Future Earnings: Where Should You Send Your Child to College?. *The American Economic Review,* 79(2), pp. 247-252.

Strayer, W., 2002. The Returns to School Quality: College Choice and Earnings. *Journal of Labor Economics,* 20(3), pp. 475-503.

The Guardian, 2017. *The Guardian University League Tables 2018.* [Online]
Available at: https://www.theguardian.com/education/ng-interactive/2017/may/16/university-league-tables-2018
[Accessed 2021].

Wachtel, P., 1976. The Effect on Earnings of School and College Investment Expenditures. *The Review of Economics and Statistics,* 58(3), pp. 326-331.

Appendices

Appendix A: Converting HESA Raw Data to Average Income Data

The HESA dataset originally had 112,680 observations, whilst my analysis only has 132. The original dataset was trimmed down in several stages. The first stage involved filtering the data so that only one observation per salary band per institution was given. The original dataset reported the number of graduates in each £3000 salary band who had a certain set of characteristics, resulting in multiple observations per band. These characteristics are listed below with the option filtered for in bold.

- Country of provider
  - England
  - Scotland
  - Wales
  - Northern Ireland
  - **All**
- Mode of former study
  - **Full-time**
  - Part-time
  - All
- Skill group
  - High-skilled
  - Medium-skilled
  - Low-skilled
  - **All**
- Work population marker
  - **Paid employment is an activity**
  - Paid employment is most important activity (a subset of the above)

After this stage, each institution had ten observations, one for each salary band. The second stage involved totalling the number of observations across the ten salary bands. Institutions with fewer than 100 observations were removed from the dataset. As discussed, very few of these were universities.

The third stage involved converting the number of observations in each band for each university into one single number for each university's average income. As discussed, I took the maximum of the lowest band, the minimum of the highest band, and the centre-point of the other bands for income point of each band. This income point was then multiplied by the number of observations to give the total income in that band. These total incomes were summed across bands and then divided by the total number of observations for that university to give it its average income. This process for each university is formalised below, where x_i is the income point for band i, n_i is the number of observations in the band, y_i is the total income in that band, and the capitalised versions are the cross-band values, including X which is the average income for that university.

$$x_i \times n_i = y_i$$

$$\sum_{i=1}^{10} y_i = Y$$

$$\sum_{i=1}^{10} n_i = N$$

$$\frac{Y}{N} = X$$

Appendix B: Full Results Tables

Table 3b – Full Primary Results

Variable	Model			
	1	2	3	4
Staff Costs per Student	0.081** (0.036)	0.076** (0.035)	0.386*** (0.129)	0.364*** (0.130)
Student-Staff Ratio	54.6 (65.3)	27.6 (69.1)	74.6 (69.4)	-540.8 (494.5)
(Staff Costs)^2			$-6.31e^{-6}$** ($2.51e^{-6}$)	$-6.30\ e^{-6}$** ($2.48\ e^{-6}$)
(SSR)^2				18.3 (14.3)
Entry Standards	45.9*** (12.0)	40.9*** (11.9)	34.3*** (10.6)	35.4*** (10.8)
Number of Students	0.004 (0.018)	-0.21 (0.20)	-0.016 (0.020)	-0.017 (0.021)
London	1430*** (326)	1580*** (322)	1390*** (330)	1360*** (333)
Scotland	-918* (475)	-773 (471)	-721* (427)	-823* (424)
Russell Group	396 (480)	489 (443)	196 (458)	180 (459)
Proportion Postgraduate	4390** (1900)	4470** (1930)	3220 (2090)	2916 (2082)
Medicine Offered		126 (290)	-213 (307)	-221 (308)
Economics Offered		292 (300)	246 (298)	270 (303)
Engineering Offered		996*** (365)	1000*** (378)	1020*** (380)
R^2	0.7701	0.7910	0.8026	0.8041

Table 6b – LEO Model REF (4 & 3*) Results*

Subject	Measure	Coefficient	P-value	R^2	Sample Size
Biological Sciences	REF 4+3**	-2272	0.037	0.8032	41
	SSR***	-1087	0.005		
Computing	REF 4+3*	3046	0.061	0.7110	80
	SSR	540	0.211		
Economics	REF 4+3	1091	0.782	0.8687	25
	SSR	419	0.460		
Engineering	REF 4+3	2061	0.696	0.7166	15
	SSR	-22	0.943		
English Studies	REF 4+3	103	0.936	0.4409	81
	SSR	-167	0.608		
Languages	REF 4+3	-927	0.762	0.5487	37
	SSR	-1339	0.388		
Law	REF 4+3	-3002	0.135	0.7612	61
	SSR***	-5801	0.000		
Nursing	REF 4+3	141	0.902	0.5702	56
	SSR	-590	0.417		

Table 7c – Full Fixed Effects Results (4)*

Variable	Coefficient	Std. Error	t	P>|t|
SSR	-13.0	34.0	-0.38	0.714
REF4	2216	1544	1.44	0.194
Entry Standards***	63.6	13.6	4.67	0.002
Number of Students***	-0.056	0.011	-5.06	0.001
London*	1070	481	2.22	0.062
Scotland***	-1670	324	-5.16	0.001
Russell Group	365	245	1.49	0.180
Proportion Postgraduate	5520	3060	1.81	0.114
Observations	396			
Groups (subject)	8			
R^2 (within)	0.5480			
rho	0.90			

Table 7d – Full Fixed Effects Results (4 & 3*)*

Variable	**Coefficient**	**Std. Error**	**t**	**P>|t|**
SSR	-13.5	31.0	-0.44	0.676
REF43	1546	1374	1.13	0.298
Entry Standards***	63.0	13.5	4.66	0.002
Number of Students***	-0.061	0.013	-4.64	0.002
London*	1082	461	2.35	0.051
Scotland***	-1670	318	-5.35	0.001
Russell Group	490	279	1.75	0.123
Proportion Postgraduate	6000	2640	1.87	0.103
Observations	396			
Groups (subject)	8			
R^2 *(within)*	0.5481			
rho	0.90			