The Impact of Undergraduate Degrees on Lifetime Earnings: Online Appendix

Jack Britton, Lorraine Dearden, Laura van der Erve and Ben Waltmann

February 2020

A Sample Selection and Data Quality

A.1 HESA Data

The HESA data set presents a number of challenges. One is lower data quality in the early years of the data: crucial information is sometimes missing or inconsistent across years. As far as characteristics ought to be constant across years, we have dealt with missing data by taking the earliest record as authoritative.[1] In cases where information on subject studied was missing, we have filled in missing data from earlier years on the same course of study.[2] We also harmonised the birth year of each individual, taking data from undergraduate degrees as the most authoritative for cohorts for which no NPD data are available.[3]

A second challenge is changing variable definitions over time. The most important example of this for our purposes is a radical change in the way HESA classifies courses that occurred between the 2006/07 and 2007/08 academic years.[4] As far as possible, we have attempted to follow HESA's own coarser classification of degrees that includes a category for first degrees.[5] However, we have diverged slightly from this classification in the interest of continuity with the pre-2007/08 classification scheme.

As a result, we classify the following course codes as 'undergraduate degrees' for our purposes:

- the **courseaim** codes 18 to 24
- the **qualaim** codes H00, H11, H12, H16, H18, H22, H23, H24, H50, I00, I11, I16.

[1] This was done for gender, ethnicity, POLAR index and home region.

[2] For our main analysis, subject mix in the final year of study was taken as authoritative.

[3] As we only observe age in the year each student started their course, and for students in the earliest years of data that year might not be in our data set, we do not have completely reliable age information on the earliest cohorts in the data. As far as possible, we have determined these students' ages on the basis of further study they may have undertaken; as a last resort, we have imputed their age based on common degree lengths.

[4] At that time, the HESA classification variable changed from 'qualaim' to 'courseaim'.

[5] For example, as reflected in the 'xqlev501' derived variable.

Table 1: HESA Data Sample Selection

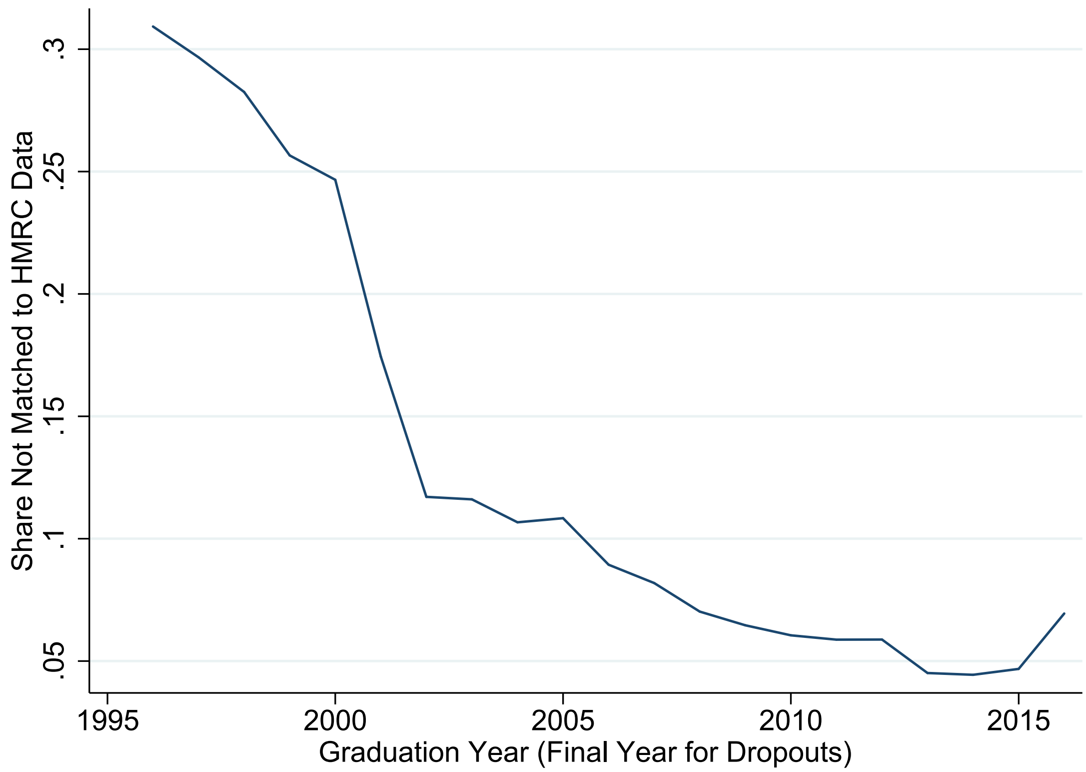
	Women			Men		
Birth Year	**Raw Data**	**Sample**	**Matched**	**Raw Data**	**Sample**	**Matched**
1975/1976	126,397	80,318	40,439	107,370	72,725	64,892
1976/1977	131,694	84,813	45,271	113,557	79,611	70,804
1977/1978	145,087	106,704	61,030	126,905	99,764	89,360
1978/1979	154,790	116,292	76,020	133,448	105,980	96,510
1979/1980	158,619	119,775	97,853	135,187	107,897	100,132
1980/1981	157,742	120,957	102,707	132,663	107,268	100,547
1981/1982	157,382	122,997	106,186	132,319	108,479	101,869
1982/1983	160,803	128,796	112,527	134,822	112,383	105,684
1983/1984	162,793	133,027	117,739	135,214	114,167	107,773
1984/1985	165,776	137,206	123,698	136,958	116,257	110,092
1985/1986	164,607	138,138	126,376	136,770	117,113	111,192

Note: HESA data from different birth cohorts. 'Raw Data' gives the number of all first-degree students for whom we observe gender and to whom we can assign an academic year of birth. 'Sample' lists the number of students from each birth cohort who started studying full time between the ages of 17 and 21. 'Matched' gives the number of students in the resulting sample who can be matched to an HMRC tax record.

Table 1 shows our HESA sample in different birth cohorts. The column labelled 'Raw Data' gives the number of all first-degree students for whom we observe gender and to whom we can assign an academic year of birth. 'Sample' lists the number of students from each birth cohort who we would like to include in the analysis; some are excluded either because they did not start studying between the ages of 17 and 21 or because they started studying part time. 'Matched' gives the number of students in the sample who we can match to an HMRC tax record.

Three aspects of these numbers are worth noting. First, student numbers in the data have increased significantly between the 1975/76 and 1985/86 cohorts, especially for women. Part of the reason for this is that we are missing some dropouts from the earliest cohorts, but a large part will also be due to an actual increase in the number of students in each cohort attending university. Second, the number of students in our sample has increased by even more, as larger numbers of students have attended university full time and soon after finishing secondary school. Third, and again especially for women, a much smaller share of the earlier cohorts can be matched to HMRC tax records. One consequence is that in our matched sample, there are three times as many women in the 1985/86 cohort as in the 1975/76 cohort.

Figure 1: Share of HESA Student Records Not Matched to HMRC Data



Note: Share of students for whom we have a HESA student record but no matching HMRC tax record. The data are organised by graduation year, or the final year observed in the case of dropouts.

While some of the difference in HMRC match rates will be due to the increased labour force participation of women, the scale and speed of the improvement in the match rates suggest that this is largely an artefact of the data. Figure 1 shows the unmatched share organised by graduation year. While there is a steady increase in match rates throughout this period, much of the difference between earlier and later cohorts is attributable to a steep change in match rates between the 2000 and 2002 graduation years.

In order to address potential bias associated with the low match rates in the early cohorts, we have imputed earnings based on gender, year of birth, subject studied, institution attended and POLAR score, which proxies for socio-economic status. This imputation removes bias insofar as whether an individual's HMRC record can be matched is independent of earnings conditional on these observables. Reassuringly, including imputed earnings changes the patterns in the data only modestly, suggesting that whether or not HESA and HMRC record can be matched is mostly random.

Figures 1 and 2 of the report and Figures 2 and 3 in this appendix are based on data including these imputed earnings. Our main estimates do not include imputed earnings but are based on simulations that are as far as possible done separately by subject studied, institution attended and

POLAR score. Consequently, bias should be minimal insofar as whether there is an HMRC–HESA match is random conditional on these observables. However, the low match share in the early cohorts should be noted as a risk to our earnings forecast, especially for women.

A.2 Linked NPD Data

Table 2 shows details of sample selection from the linked NPD–HESA–HMRC data. All data are for the 2002 GCSE cohort, as this is the only cohort used in our analysis; the vast majority of students in this GCSE cohort were born in the 1985/86 academic year. 'Raw Data' lists all English students in the 2002 GCSE cohort with a usable Key Stage 4 record. '5 A*–C & KS5' shows the number of students who obtained at least five A*–C marks in their GCSE exams and went on to Key Stage 5. 'Sample' gives the number of students who fall into our sample because they either never had any contact with the HE system or started a full-time first-degree course between the ages of 17 and 21. 'Matched' lists the number of students from this sample for whom we have HMRC records. 'Final Selection' gives the number of these for whom we can identify the institution and who did not study one of the subjects excluded from the analysis.[6]

Table 2: 2002 GCSE Cohort Sample Selection

Gender	Raw Data	5 A*–C & KS5	Sample	Matched	Final Selection
Women	261,859	116,764	108,171	100,632	98,202
Men	259,294	97,694	90,835	87,792	85,046

Note: Sample sizes in the linked NPD–HMRC–HESA data for the 2002 GCSE cohort. 'Raw Data' lists all English students in the 2002 GCSE cohort with a usable Key Stage 4 record. '5 A*–C & KS5' shows the number of students who obtained at least five A*–C marks in their GCSE exams and went on to Key Stage 5. 'Sample' gives the number of students who fall into our sample because they either never had any contact with the HE system or started a full-time first-degree course between the ages of 17 and 21. 'Matched' lists the number of students from this sample for whom we have HMRC records. 'Final Selection' gives the number of these for whom we can identify the institution and who did not study one of the subjects excluded from the analysis.

Less than half of those with a Key Stage 4 record achieved the threshold of five A*–C GCSEs and went on to Key Stage 5, highlighting that those fulfilling the minimum criteria for inclusion in our sample are already a selected group. Among people fulfilling these two criteria, a large majority are in our sample; few enrolled in an 'other undergraduate' degree or went to university as mature students. Almost everyone from this sample can be matched to an HMRC tax record, as match rates are no longer a major issue for the 1985/86 cohort. Among those who can be matched, very nearly all end up in our final sample; most of the difference is due to the exclusion of students of sports science, which was not commonly taught at the time the earliest cohorts in our data went to university.

The linked sample is substantially smaller than the matched HESA–HMRC sample, as can be seen from a comparison of Tables 1 and 2. The main reason for this is that the NPD data cover

[6]These subjects are Celtic studies, combined studies, humanities not further specified, sports science and veterinary science.

only English students, whereas the HESA data include students from elsewhere in the UK. Other relevant factors include incomplete Key Stage 4 data, immigration and the fact that a small number of students from the 2002 GCSE cohort entered university without both five or more A*–C GCSEs and a Key Stage 5 record.[7]

B Details on Counterfactual Employment Simulation

The effect of a degree on the probability of being employed is an important part of the overall return of HE. Attending university depresses participation at the time of attendance and, depending largely on subject choice, in some cases raises it thereafter. This section presents a more detailed description of our simulation of counterfactual employment than is given in Subsection 3.3.2 of the report.

As a first step, we estimate the probit model

$$P(E_{i,a} = 0 | x_i) = \Phi(x_i' \gamma_a^E)$$

at each age and for each gender on data only on individuals who did not attend university. Using this model, we can calculate a counterfactual worklessness probability for all individuals in our sample who *did* attend HE. That probability will reflect what we would have expected their worklessness probability to be if they had not attended HE given their background conditions.

In a second step, we take the average over counterfactual worklessness rates by subject and gender. Mathematically, these are given by

$$\pi_{as} = \frac{1}{N^s} \sum_{i \in I^s} \Phi(x_i' \hat{\gamma}_a^E)$$

where π_{as} is the counterfactual worklessness probability for subject s at age a, I^s is the set of indices relating to individuals who studied subject s, N^s is the number of such individuals and $\hat{\gamma}_a^E$ is the estimated parameter vector from the estimation of the probit model above.[8]

As a third step, we create a simulated panel of counterfactual employment statuses that matches the counterfactual subject worklessness rates π_{as}. This is achieved as follows. First, we asssign everyone their actual observed or simulated employment status as a default counterfactual employment status. Second, we determine whether the counterfactual worklessness rate is higher or lower than the actual worklessness rate of each subject group.

If the counterfactual worklessness rate is *lower* than the actual worklessness rate, we calculate the weight

$$w_{as}^- = \frac{u_{as} - \pi_{as}}{u_{as}(1 - \pi_{as}^U)}$$

where u_{as} is the actual worklessness rate for subject group s at age a and π_{as}^U is the mean counter-

[7]This was rare at the time but has become much more widespread since.

[8]Gender subscripts are dropped for readability.

factual worklessness rate for subject group s at age *a conditional on not working at age a*. Then we change counterfactual employment statuses to

$$E^*_{i,a} = I\left\{\nu_{ia} < w^-_{as}\left[1 - \Phi(x_i\hat{\gamma}^E_a)\right]\right\}$$

for all individuals whose observed or simulated employment status was employed, i.e. for whom $E_{i,a} = 1$. $I\{.\}$ is the indicator function and ν_{ia} is a systematically sampled draw from the uniform distribution.

If the counterfactual worklessness rate is *higher* than the actual worklessness rate, we calculate the weight

$$w^+_{as} = \frac{\pi_{as} - u_{as}}{(1 - u_{as})\pi^E_{as}}$$

where u_{as} is the actual worklessness rate for subject group s at age a and π^E_{as} is the mean counterfactual worklessness rate for subject group s at age *a conditional on working at age a*. Then we change counterfactual employment statuses to

$$E^*_{i,a} = I\left\{\nu_{ia} < 1 - w^+_{as}\Phi(x_i\hat{\gamma}^E_a)\right\}$$

for all individuals whose observed or simulated employment status was workless, i.e. for whom $E_{i,a} = 0$. Again, $I\{.\}$ is the indicator function and ν_{ia} is a systematically sampled draw from the uniform distribution.

Using this procedure, we can match counterfactual worklessness rates by subject.[9] Assigning counterfactual employment or worklessness according to a modelled probability of being out of work allows us to deal with selection issues. At the same time, preserving actual/simulated employment as a default means that we can mostly retain the residual component of earnings in counterfactual simulation that our earnings model does not capture.

C Mid-Career Earnings of the 1975/76 Cohort

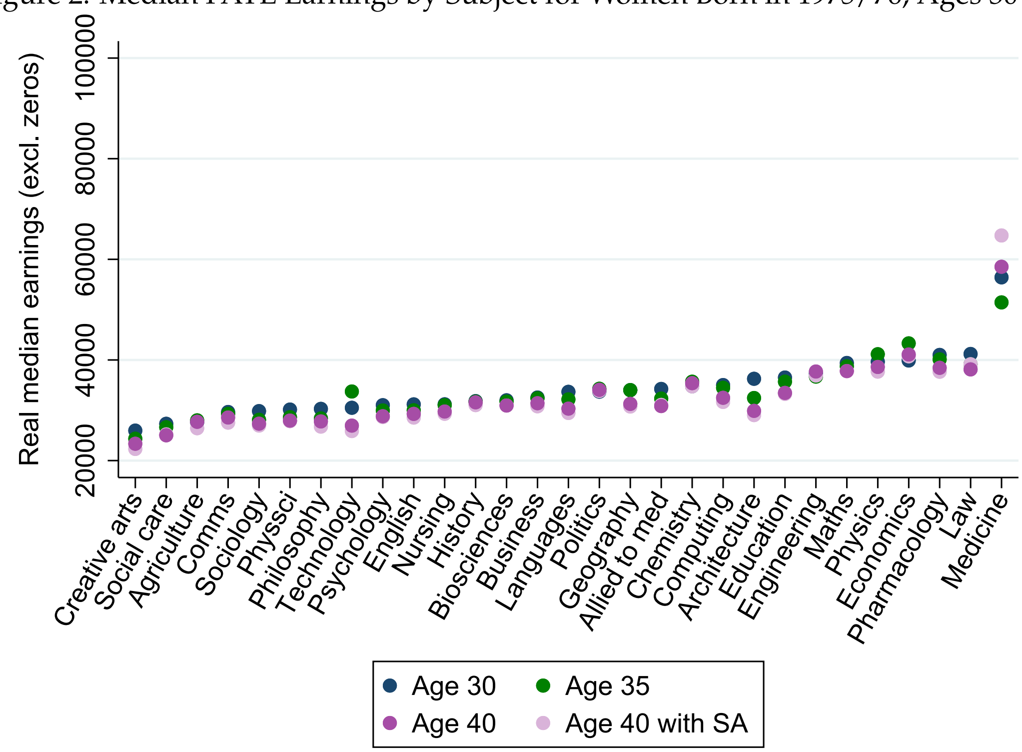
In Section 2.3, we presented evidence on earnings at different ages *in the 2016/17 tax year*, which is the latest tax year we observe. This appendix complements that section by presenting earnings at different ages *for the 1975/76 birth cohort*. In contrast to the data presented in the main report, this has the advantage that it does not pick up *cohort effects*, i.e. influences on earnings that affect some cohorts but not others. For instance, if people studying economics in the 1985/86 cohort earned more than those in the 1975/76 cohort at the same age because of an improved curriculum, this would lead to a narrower gap between age 30 and age 40 earnings in our diagrams in Section 2.3, even though the age profile of earnings for economists might be unchanged. The same would not

[9] Proofs are available on request. A technical condition is that $0 < w^-_{as}\left[1 - \Phi(x_i\hat{\gamma}^E_a)\right] < 1$ or $0 < w^+_{as}\Phi(x_i\hat{\gamma}^E_a) < 1$ when counterfactual unemployment is respectively lower or higher than actual worklessness. A sufficient (but not necessary) condition for this to hold is that $w^-_{as}, w^+_{as} \in [0, 1]$ for all a and s. We observe that this stronger condition generally holds in practice.

be the case in the diagrams presented in this appendix, as they show earnings for the 1975/76 cohort only.

The main downside of following one cohort through time is that the diagrams also incorporate *time effects*, i.e. influences on earnings that affect all individuals across cohorts at a particular time. An important example of a time effect is the Great Recession, which severely affected all cohorts. Another downside is that we only observe earnings for the self-employed from the 2013/14 tax year onwards, so for the 1975/76 cohort, self-employment earnings are only available from age 37.[10]

Figure 2: Median PAYE Earnings by Subject for Women Born in 1975/76, Ages 30–40



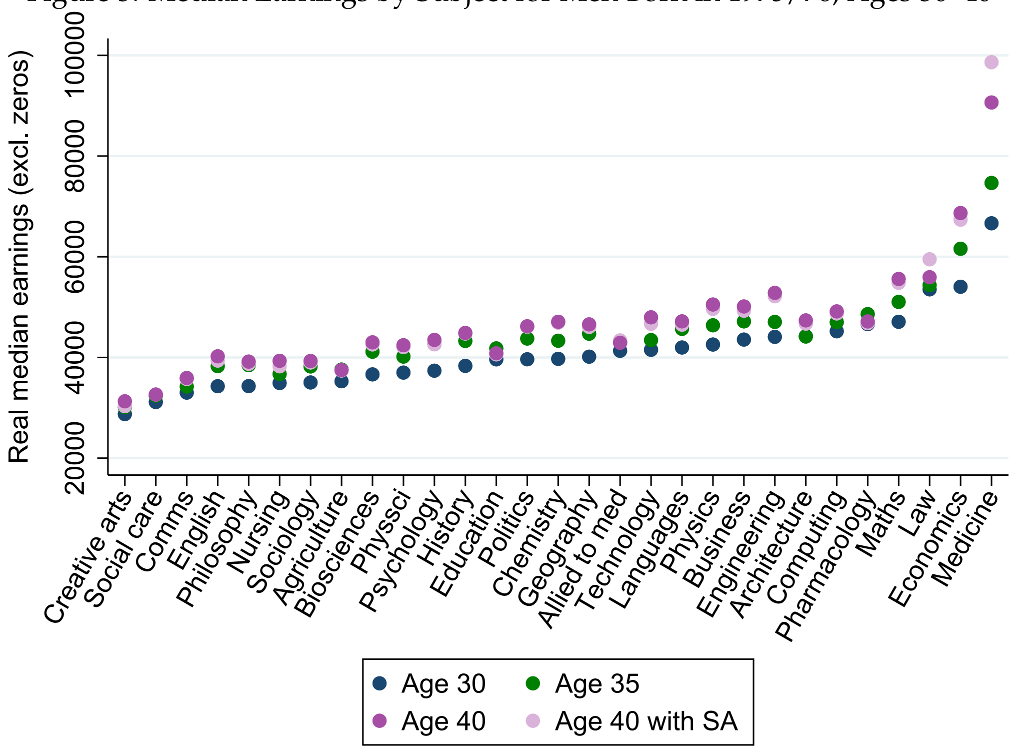
Note: Median pre-tax earnings by subject for female graduates from the 1975/76 cohort at different ages. All values are conditional on positive earnings, in 2018 prices, and Winsorised at the 1st and 99th percentiles. 'Age 40 with SA' includes self-assessment earnings in addition to PAYE earnings.

Figure 2 shows the growth of median PAYE earnings for women born in the 1975/76 academic year, conditional on having positive earnings. For age 40, we also include data points that incorporate self-employment earnings. Overall, the picture is much less optimistic than when looking across cohorts. Across nearly all subjects, women in this cohort saw zero or negative earnings

[10] A third concern is data quality, which is lower for earlier cohorts. For the 1975/76 cohort, only about two-thirds of university students can be matched to an earnings record. We adjust for this issue in all our figures by including imputed earnings based on gender, year of birth, subject studied, institution attended and POLAR score, which proxies for socio-economic status. Reassuringly, whether or not imputed earnings are included makes very little difference to the figures.

growth in their 30s. Notably, this is also true for law, where there are strong earnings differences in the 2016/17 tax year. The only exception is medicine. Medicine graduates saw a modest amount of earnings growth, which looks especially strong considering that few will have been self-employed at age 30.

Figure 3: Median Earnings by Subject for Men Born in 1975/76, Ages 30–40



Note: Median pre-tax earnings by subject for male graduates from the 1975/76 cohort at different ages. All values are conditional on positive earnings, in 2018 prices, and Winsorised at the 1st and 99th percentiles. 'Age 40 with SA' includes self-assessment earnings in addition to PAYE earnings.

Figure 3 shows that, as in the cross-cohort comparison, earnings differences between age 30 and age 40 were larger for men than for women in the 1975/76 cohort. However, the differences are much smaller overall than when comparing across cohorts. As for women, medicine stands out as the subject with both the highest median earnings at age 30 and the strongest earnings growth between age 30 and age 40.

The main reason for the disappointing growth in median real earnings is likely to be the Great Recession: this cohort had relatively high pre-recession earnings at age 30, and then disappointing earnings growth throughout their 30s as the recession hit. All graduates in subsequent cohorts, but especially men, have seen much lower real earnings at age 30, explaining the larger differences between ages 30 and 40 in the cross-cohort comparison.[11] It should also be noted that, especially for men, mid-career growth in *average* earnings for the 1975/76 cohort was much stronger than

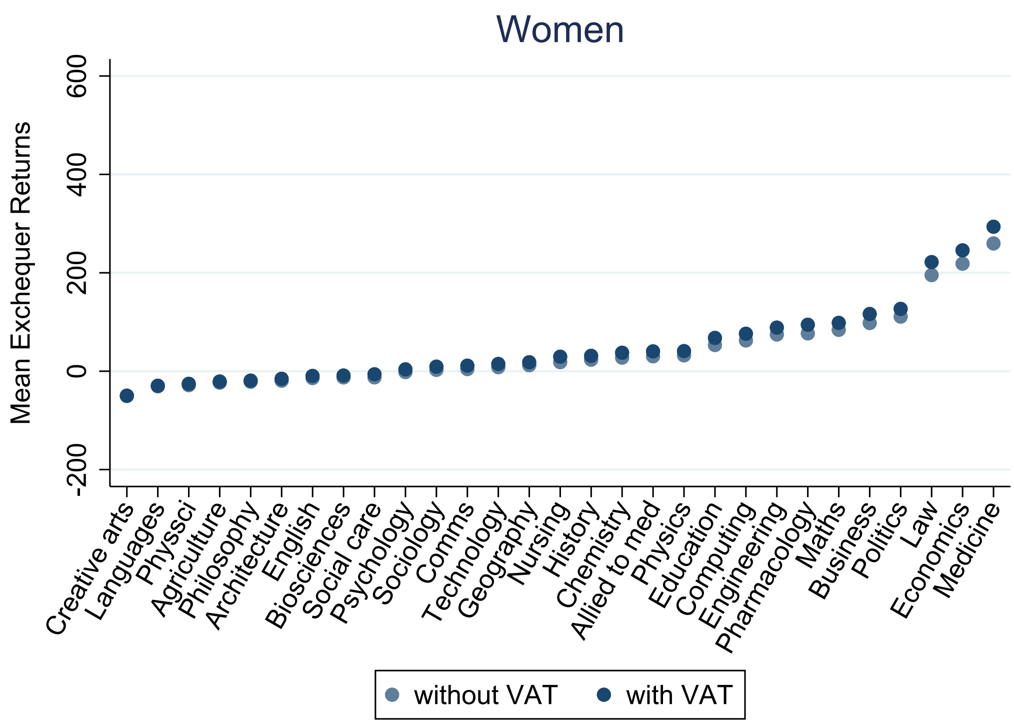
[11]See Figures 35 and 36 in the report for a direct comparison of earnings trajectories across cohorts.

growth in *median* earnings, as those at the top end of the earnings distribution gained the most.

D The Effect of HE on VAT Payments

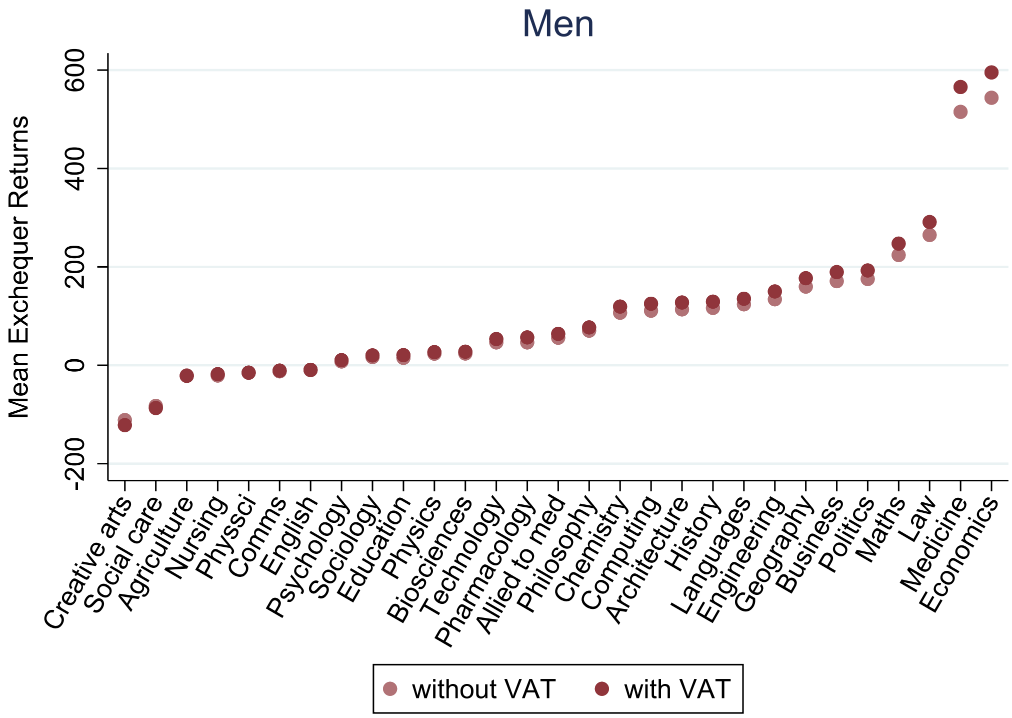
This section addresses the effect of higher education on VAT receipt. VAT was left out of the exchequer returns calculations in the main report in order to ensure that net returns and exchequer returns would add up to total returns. As VAT payments are included in the net returns figures, the sum of exchequer returns and net returns would otherwise include VAT payments twice.

Figure 4: Lifetime Exchequer Returns with and without VAT by Subject, Women



Note: All figures are shown in 2018 prices in £k and are discounted using Green Book discounting.

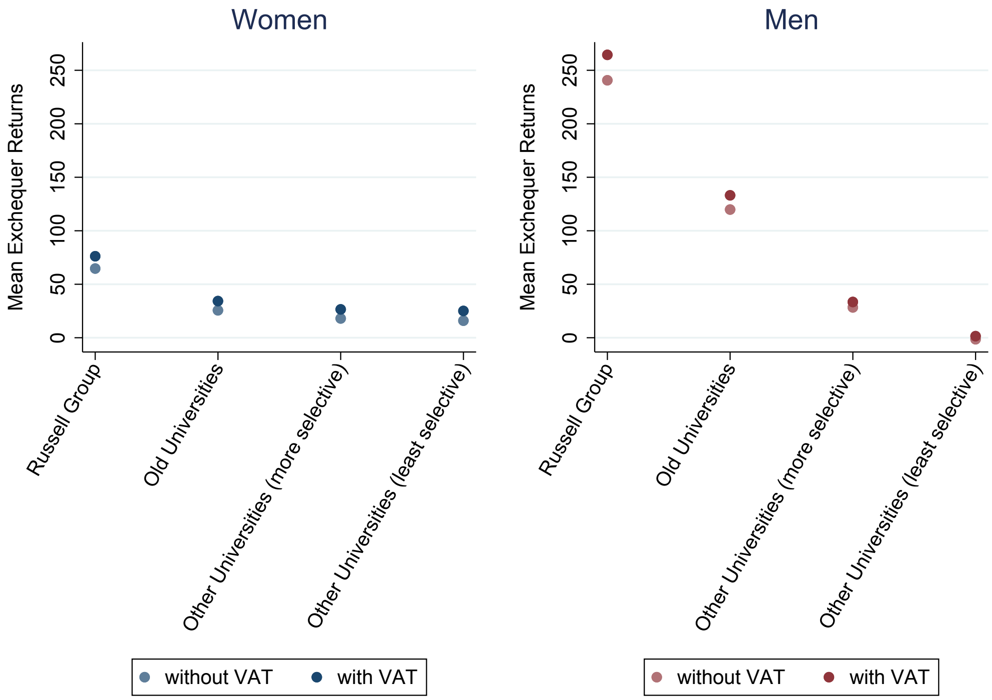
Figure 5: Lifetime Exchequer Returns with and without VAT by Subject, Men



Note: All figures are shown in 2018 prices in £k and are discounted using Green Book discounting.

Figures 4 and 5 show the impact of including VAT in the lifetime exchequer returns calculations by subject. VAT only makes an appreciable difference to the lifetime returns to the highest-earning subjects. For these subjects, taking account of VAT raises the exchequer return by up to around £50k for men and £30k for women. For subjects with negative net lifetime returns, taking into account VAT further lowers the estimated lifetime exchequer returns.

Figure 6: Lifetime Exchequer Returns with and without VAT by HEI Type



Note: All figures are shown in 2018 prices in £k and are discounted using Green Book discounting.

Figure 6 shows the effect of counting VAT on lifetime exchequer returns by HEI type. For women, including VAT raises lifetime exchequer returns appreciably in relative terms, with cash returns increasing by around half for women who studied at 'other' universities. For men, including VAT only makes a meaningful difference to the exchequer returns of men attending Russell Group and other pre-1992 ('old') universities; with VAT included, the exchequer returns from men educated at these universities appear even higher.

E The Effect of HE on Benefit Receipt

Given the data that we have, estimating the effect of pursuing an undergraduate degree on benefit receipt is a complex task. Benefits are assessed at the family level, but the LEO data only contain information about individuals and not about their families. Furthermore, benefit eligibility usually depends on factors other than income, such as housing tenure or hours worked, which we do not observe. Hence we cannot give estimates for the impact of HE on benefit receipt at the same degree of confidence as our other estimates, which are already based on assumptions of what we expect the future to look like. For this reason, we have not included any impact on benefits in the main results of this report.

However, it is possible to simulate these missing data using information from other data sets.

While the assumptions underlying such simulations are strong, we include estimates from such a simulation exercise here to give the reader a very rough sense of the difference that the inclusion of benefits might make to our main results. We include six different benefits, which make up the bulk of working-age, non-disability benefit spending: income support,[12] housing benefit, council tax reduction, working tax credit, child tax credit and child benefit.[13] Our calculations are based on the 2019 benefits system, but we ignore the introduction of universal credit.

For the simulation, we proceed as follows. First we split our analysis data set into groups by employment status, gender and age. Then we match all individuals within these groups with individuals from the Labour Force Survey who fall into the same categories. In this way, we can simulate whether the individuals in our analysis data set have a partner, if so the partner's earnings, and the number of children.[14] Then we match in an analogous way by number of children, age, employment and partner's employment[15] with data from the Family Resources Survey to obtain all other relevant variables for the benefit calculation.[16] This is done twice: once with the simulated actual incomes of people from the 2002 GCSE cohort with HE, and once with their estimated counterfactual earnings if they had not attended HE. For each of the two resulting data sets, we calculate benefit entitlements using the Institute for Fiscal Studies's TAXBEN model of the tax and benefit system.[17]

Table 3 shows the average change in the discounted lifetime benefit entitlements of individuals' households as a result of attending HE, calculated as the average difference between lifetime benefits with predicted earnings and with counterfactual earnings. Overall, we estimate that for the families of women in HE, benefit entitlements over their lifetimes would be around £5k higher if they had not attended HE; for men, the same figure is roughly zero. These figures are small in comparison with our estimated average net lifetime returns to HE of £100k for women and £130k for men. They also appear modest from the point of view of the exchequer, compared with a lifetime exchequer return without VAT and benefits of £30k per student for women and £110k per student for men, especially considering that take-up of benefit entitlements tends to be below 100%.

[12]We include income-based employment and support allowance (ESA) and jobseeker's allowance (JSA) in this category.

[13]Child benefit is calculated net of the high-income child benefit tax charge.

[14]In order to reflect the degree of assortative mating on earnings, we assign partner's earnings as a proportion of own earnings if both partners are employed.

[15]We do not match on partnership status for single men with children or for single women with three or more children, because these combinations are rare. For 19-year-olds living with a partner, and under-24-year-olds with a partner and children, we do not match by partner's employment for the same reason.

[16]These are: partner's age, hours worked, partner's hours worked, region, council tax band, housing tenure, rent, number of bedrooms and ESA entitlement status.

[17]As this is an extremely computationally expensive process, we can only calculate benefits for a subsample of individuals. The data presented here were obtained using predicted and counterfactual life-cycle earnings profiles for 2000 individuals.

Table 3: Average Effect of HE on Lifetime Benefit Entitlement

	Women	**Men**
Income support	-1500	400
Housing benefit	-700	300
Council tax reduction	-300	100
Working tax credit	-700	-700
Child tax credit	-1000	0
Child benefit	-400	0
Total	-4500	200

Note: Average effects of HE on lifetime benefit receipt are shown in 2018 prices, and are discounted using Green Book discounting. As benefits are paid to families and not individuals, the figures shown are differences at the family level. 'Income support' includes income-based employment and support allowance (ESA) and jobseeker's allowance (JSA). Effects on child benefit are calculated net of the high-income child benefit tax charge. Individual components may not sum to the total due to rounding.

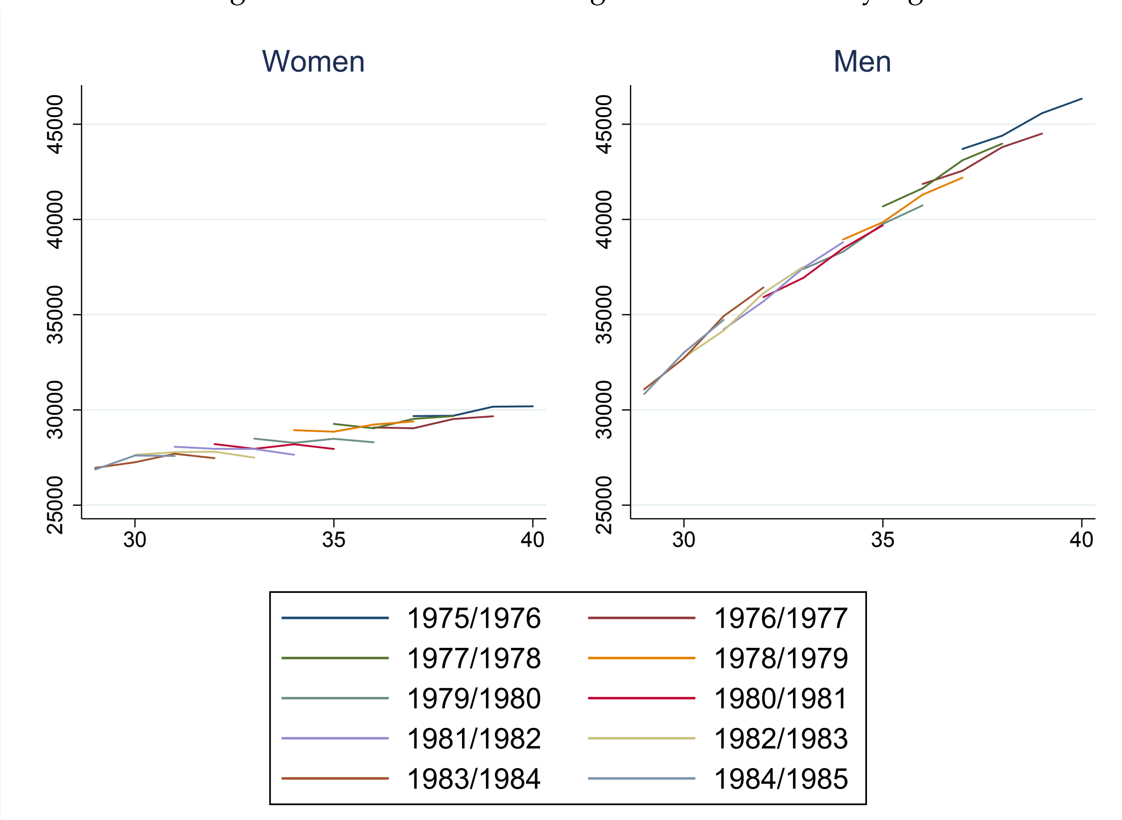
Women's average benefit income is reduced by much more than that of men as a consequence of HE. There are three main reasons for this. First, women without HE generally have lower incomes than men, making them more likely to be eligible for benefits in the first place. Second, women have higher and less dispersed returns to HE than men in their 20s and 30s, leading to a more positive effect on income compared with men. Third, women are much more likely than men to be single parents, making it much more likely that they will be eligible for means-tested benefits relating to children. The exception is working tax credit, for which a substantial number of men without HE are eligible.

F Robustness: Approaches to the Age–Period–Cohort Problem

As noted in Subsection 3.1.3 of the report, extracting measures of earnings growth by age from observational data is in general impossible without making substantive assumptions. The reason is that, in addition to age, we might expect birth cohort and observation period to have independent effects on earnings. For example, earnings of later *cohorts* might be lower if they contained more students with lower school marks. Earnings in a given *period* might be lower because of poor macroeconomic conditions, e.g. as in the Great Recession. Unfortunately, it is mathematically impossible to separate out all three types of effects without making some assumption about them. This general difficulty is known in the economics literature as the age–period–cohort problem. This problem is especially salient in our case, as our data cover both a large expansion in the HE sector, which may have led to cohort effects, and substantial macroeconomic turbulence due to the Great Recession, which will have caused period effects.

The applied economics literature facing the age–period–cohort problem mostly falls into one of two camps (cf. Schulhofer-Wohl, 2018): the *period view* and the *cohort view*, which emphasise period and cohort effects respectively. The simplest version of each view is to assume that there are no other effects, i.e. no cohort effects on the period view and no period effects on the cohort view. On each view, a mathematically more complex alternative is to assume that other effects are orthogonal to a time trend; the first papers in the literature using these techniques were Deaton and Paxson (1994) for the cohort view and Chamon and Prasad (2010) for the period view.[18] Our estimates in the main text of the report, following previous IFS work as well as other work in the economics of education literature, rely on the simple version of the period view: we assume that there are no cohort effects. In this section, we show how our results in Section 5 of the report would have changed if we had instead used the method of Chamon and Prasad (2010), or taken the cohort view and used either the simple method or the method of Deaton and Paxson (1994).[19]

Figure 7: Median Total Earnings of HE Attendees by Age



Note: Median total earnings for the 1975/76 to 1984/85 cohorts by age. HE attendees with positive earnings only. Earnings are in 2018 prices.

Figure 7 shows median total earnings for women and men from the 2013/14 tax year to the 2016/17 tax year, which are the years of data that are actually used for our earnings growth forecasts.[20] While median earnings of older women are higher than those of younger women, earn-

[18]In fact, there are multiple versions of each method that differ in terms of the weighting of observations from different cohorts. These details do not materially affect our results, so we gloss over them here.

[19]For the cohort view methods, we remove the economy-wide trend in overall earnings by adjusting for earnings growth using the OBR definition.

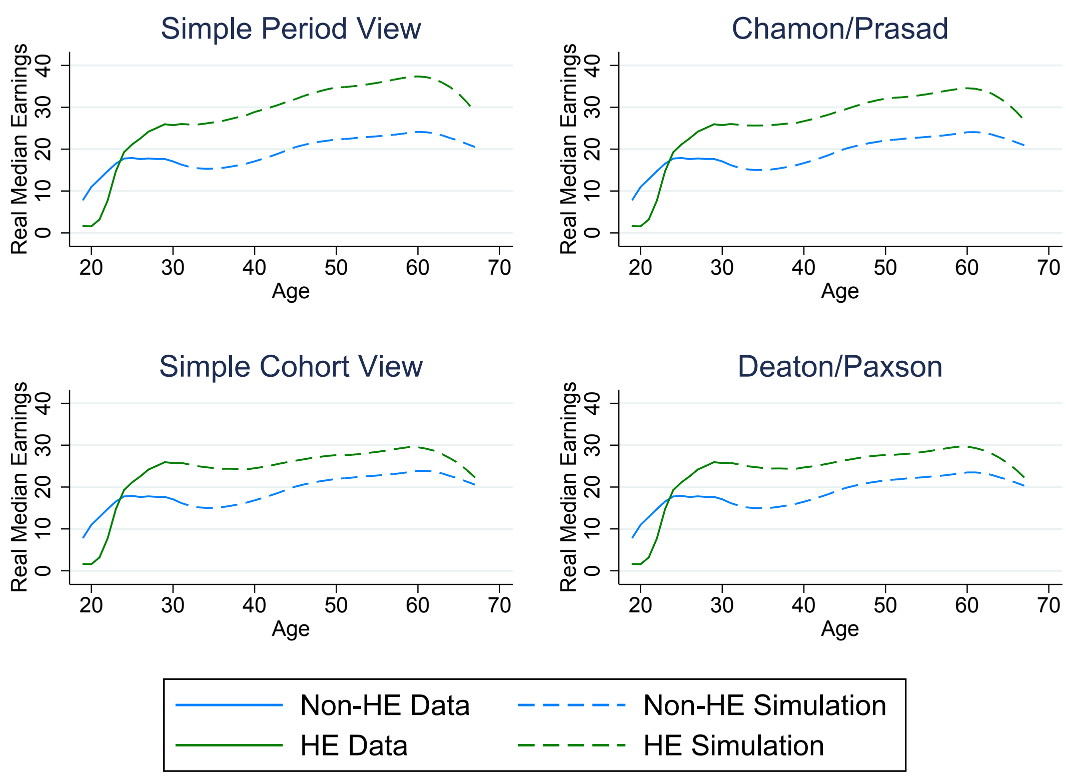
[20]We do not make use of data from earlier years in our earnings growth forecasts, because we do not observe self-employment earnings for earlier years. Furthermore, data from earlier years are likely to be strongly affected by the

ings of women in each cohort have generally stayed roughly constant over the four years; in some of the earlier cohorts, we even see declines. For men, overall earnings growth is clearly much stronger, and differences over time are closer in line with differences across cohorts. However, especially for the earliest cohorts, earnings growth within a given cohort is somewhat lower than differences across cohorts.

How one interprets these patterns hinges crucially on whether one takes the period or the cohort view. On the period view, we interpret the weak growth in earnings over time as a temporary phenomenon, and take the differences across cohorts as representing the true age effects. On the cohort view, one would assume that the low earnings growth seen for each cohort represents the true age effect, and the differences between cohorts represent not age effects but cohort effects. Which view is correct is impossible to tell from the data. However, especially for men, data from the most recent cohorts are encouraging; they suggest that these cohorts may well achieve higher earnings growth than their immediate predecessors in their 30s and thus achieve similar incomes to earlier cohorts by the time they reach 40. These differences in interpretation of the same pattern in the data result in different simulated earnings profiles, and therefore different simulated returns using the different methods of resolving the age–period–cohort problem.

Great Recession.

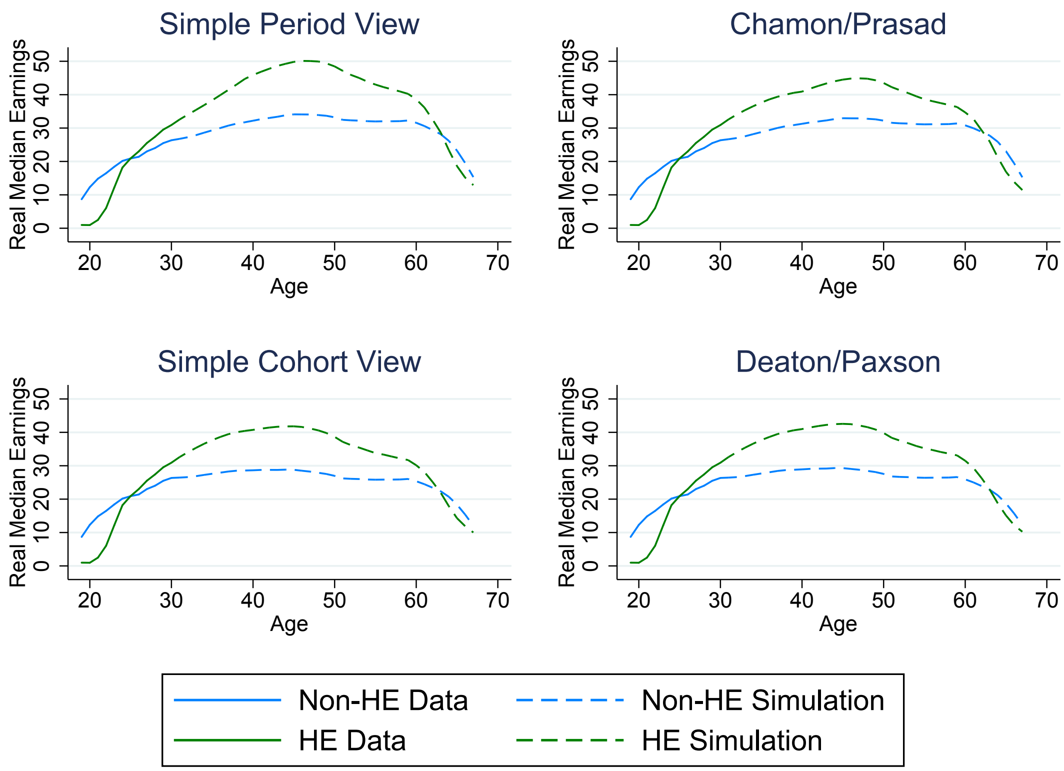
Figure 8: Median Life-Cycle Earnings of the 2002 GCSE Cohort, Women



Note: All results are estimated from separate OLS regressions, where the non-HE group only includes those with at least five A*–C GCSEs and a Key Stage 5 record and excludes those with other HE below undergraduate level. The impact of initial conditions is fixed at age 30 to help deal with the fact that the later-life estimates are based on simulated data. The dashed line shows the returns at age 29, in line with the estimates in Belfield et al. (2018). The 95% confidence intervals only capture sampling uncertainty regarding the difference in conditional means between HE and non-HE groups; they do not account for simulation uncertainty or uncertainty in the estimation of the dependence of earnings on initial conditions.

Figure 8 shows median projected life-cycle earnings for HE and non-HE women using the four different methods. Projected earnings for non-HE women are very similar across all four methods. For HE women, the pattern is somewhat different with different methods, with our preferred method (top left panel) yielding the most optimistic forecast and the methods that take the cohort view (bottom two panels) giving the least optimistic forecast. This is consistent with the pattern observed in Figure 7 that earnings differences with age for HE women appear to be larger when looking across cohorts than when considering the same cohort over time.

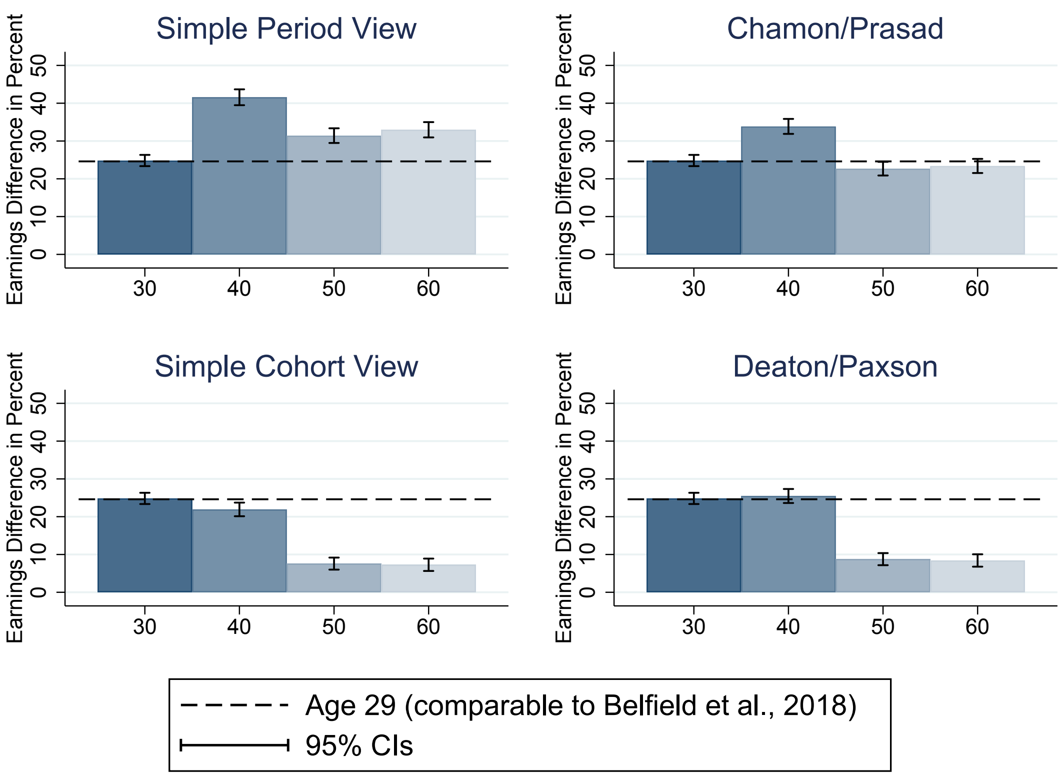
Figure 9: Median Life-Cycle Earnings of the 2002 GCSE Cohort, Men



Note: All results are estimated from separate OLS regressions, where the non-HE group only includes those with at least five A*–C GCSEs and a Key Stage 5 record and excludes those with other HE below undergraduate level. The impact of initial conditions is fixed at age 30 to help deal with the fact that the later-life estimates are based on simulated data. The dashed line shows the returns at age 29, in line with the estimates in Belfield et al. (2018). The 95% confidence intervals only capture sampling uncertainty regarding the difference in conditional means between HE and non-HE groups; they do not account for simulation uncertainty or uncertainty in the estimation of the dependence of earnings on initial conditions.

Figure 9 gives the analogous comparison for men. Again the period view (top two panels) yields a more optimistic forecast for university graduates than the cohort view (bottom two panels). However, in contrast to the forecast for women, we see the same pattern for non-HE men. For both men and women, the Chamon/Prasad method yields a somewhat less optimistic forecast than the comparison of means method that we employ in the main part of the report.

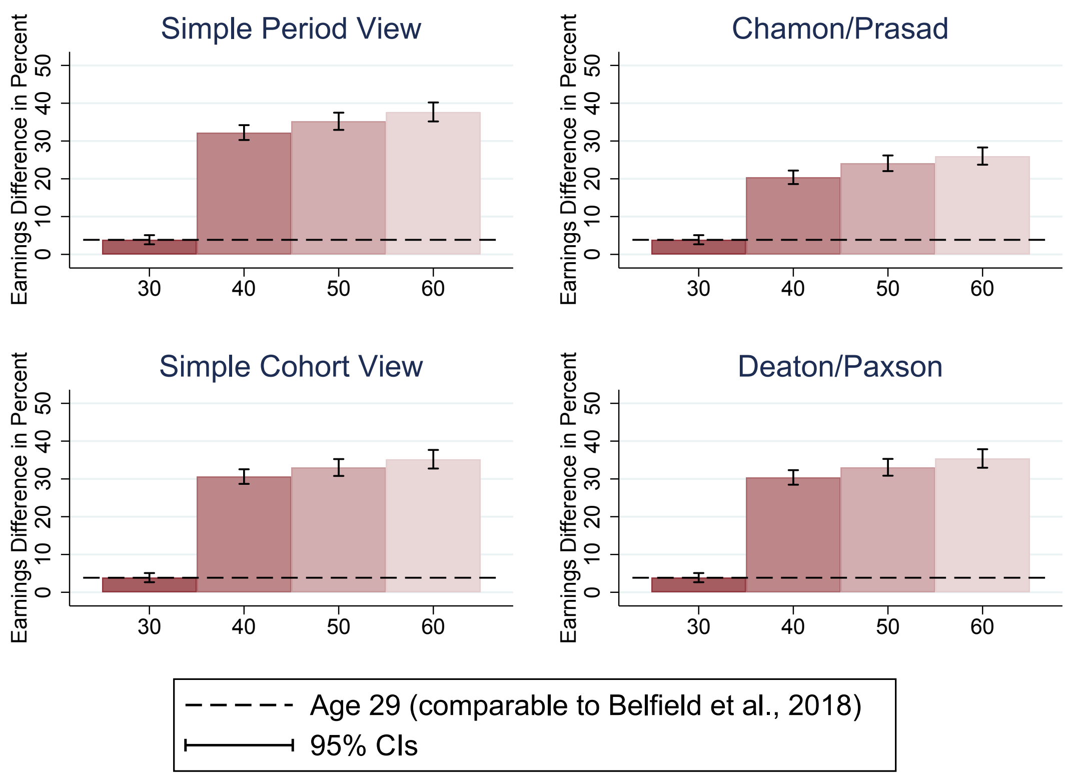
Figure 10: Average Returns to HE for Women in Work



Note: Women's median projected earnings in 2018 prices. Includes zero earnings. 'Non-HE' conditions on having at least five A*–C GCSEs and a Key Stage 5 record.

These differences in earnings forecasts are reflected in expected returns. Overall returns for women are presented in Figure 10. Again we see that our preferred method (top left panel) presents the most optimistic picture. While the Chamon/Prasad method yields broadly similar results to our preferred method, projected percentage returns are much lower for the two methods representing the cohort view. These differences highlight the large uncertainty attending our foecasts, especially with regard to women's earnings. As highlighted in the report, large changes in the labour market participation of women over the past two decades and the earnings impact of childcare responsibilities lead to large uncertainties in our forecast.

Figure 11: Average Returns to HE for Men in Work

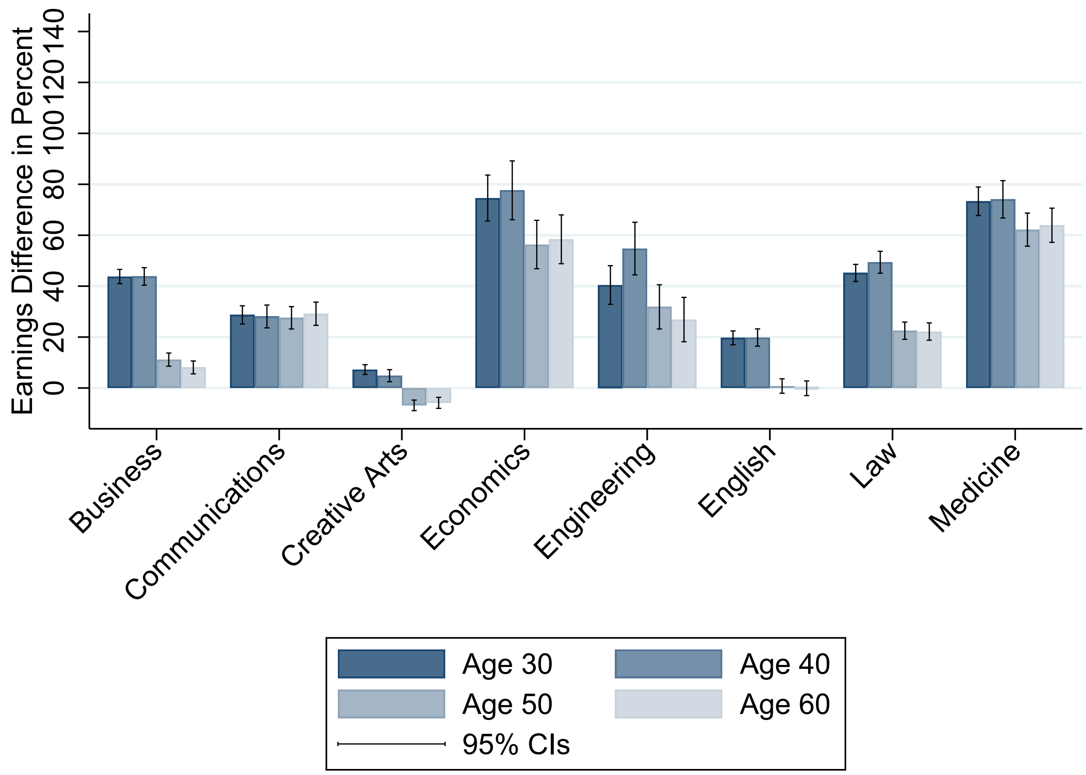


Note: Men's median projected earnings in 2018 prices. Includes zero earnings. Non-HE conditions on having 5 A*-C GCSEs and a Key Stage 5 record.

Figure 11 is the equivalent figure for men. For men, the differences in percentage returns between the four methods are smaller. The reason is that, even though the earnings forecasts differ quite significantly, forecasts for non-HE and HE men change largely in tandem, leaving estimated percentage returns approximately unchanged.

We now provide more details for the Deaton/Paxson method (bottom right panel in the comparison figures). We have singled out this method because it is widely used in the wider economics literature. It also provides a suitably large contrast to our preferred method.

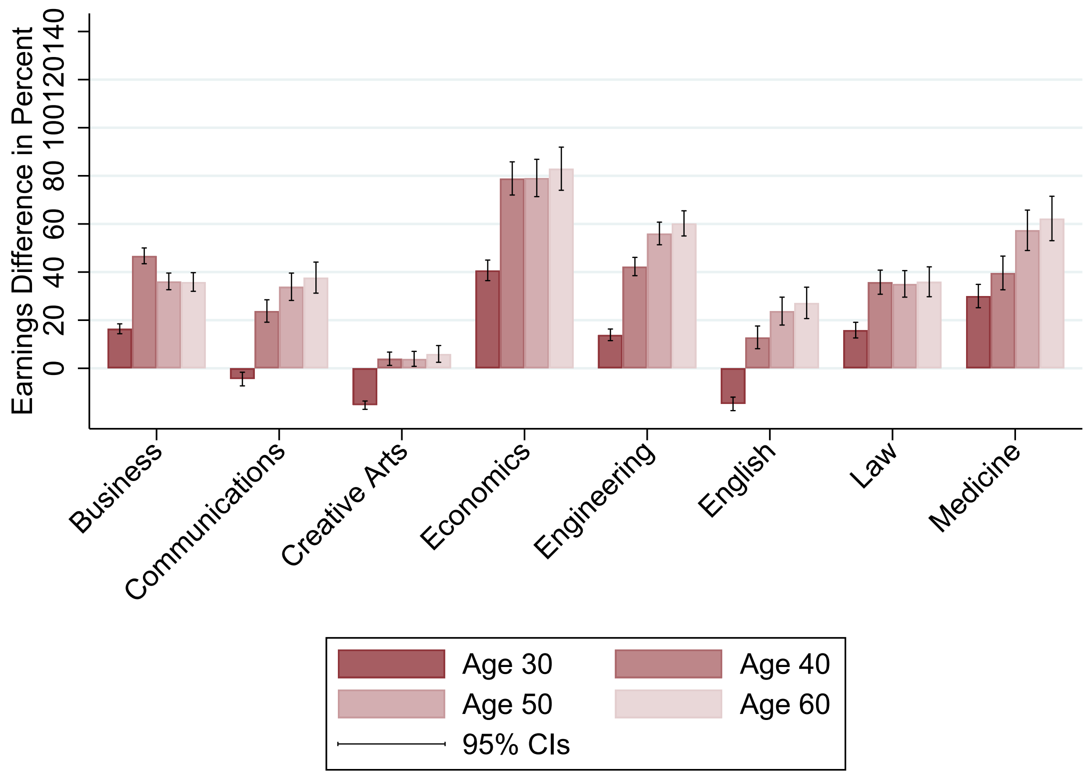
Figure 12: Returns to HE for Women in Work by Subject and Age, Deaton/Paxson Method



Note: All results are estimated from separate OLS regressions, where the non-HE group only includes those with at least five A*–C GCSEs and a Key Stage 5 record. The impact of initial conditions is fixed at age 30 to help deal with the fact that the later-life estimates are based on simulated data. The 95% confidence intervals only capture sampling uncertainty regarding the difference in conditional means between graduates of a given subject and the non-HE group; they do not account for simulation uncertainty or uncertainty in the estimation of the dependence of earnings on initial conditions.

Figure 12 provides subject returns for women; it is directly analogous to Figure 14 in the main report. In keeping with the overall returns shown in Figure 10, returns for all subjects are either similar to or lower than those given in the main report. Differences are especially large in medicine and law. For medicine, this is likely related to low public sector earnings growth in the last few years. Recent law graduates may have been particularly hard-hit by the Great Recession.

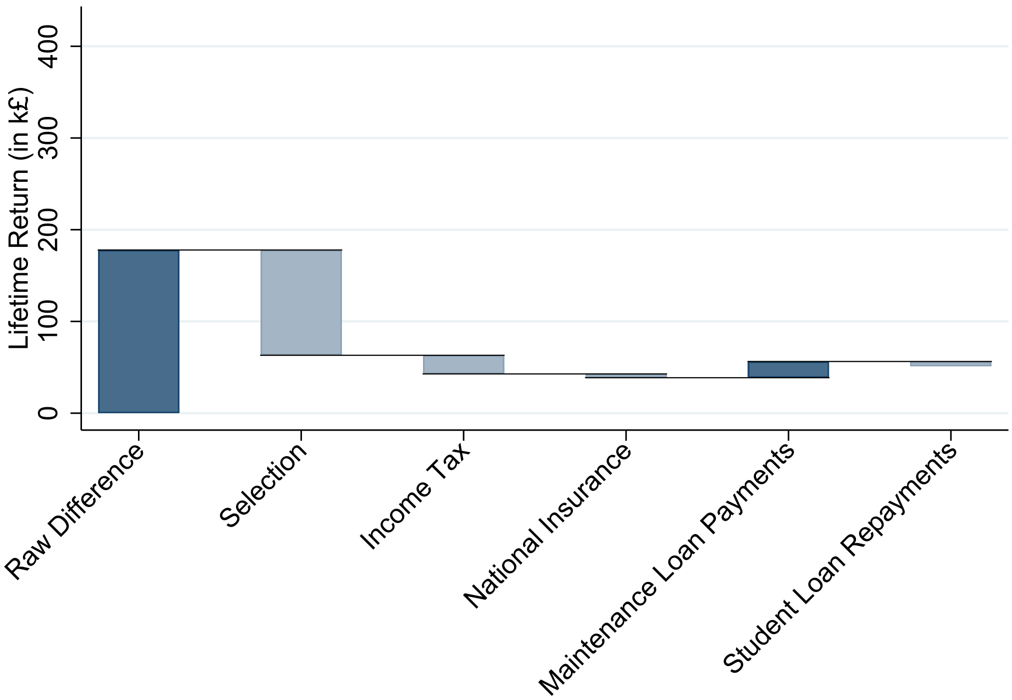
Figure 13: Returns to HE for Men in Work by Subject and Age, Deaton/Paxson Method



Note: All results are estimated from separate OLS regressions, where the non-HE group only includes those with at least five A*–C GCSEs and a Key Stage 5 record. The impact of initial conditions is fixed at age 30 to help deal with the fact that the later-life estimates are based on simulated data. The 95% confidence intervals only capture sampling uncertainty regarding the difference in conditional means between graduates of a given subject and the non-HE group; they do not account for simulation uncertainty or uncertainty in the estimation of the dependence of earnings on initial conditions.

Some similar patterns hold for men, as shown in Figure 13: law and medicine again look much less lucrative on the cohort view. However, an interesting feature of the data for men is that the lower-earning subjects in the graph – communications, creative arts and English – actually see higher returns on the cohort view than on the period view. Economics stands out as by far the most lucrative subject for men.

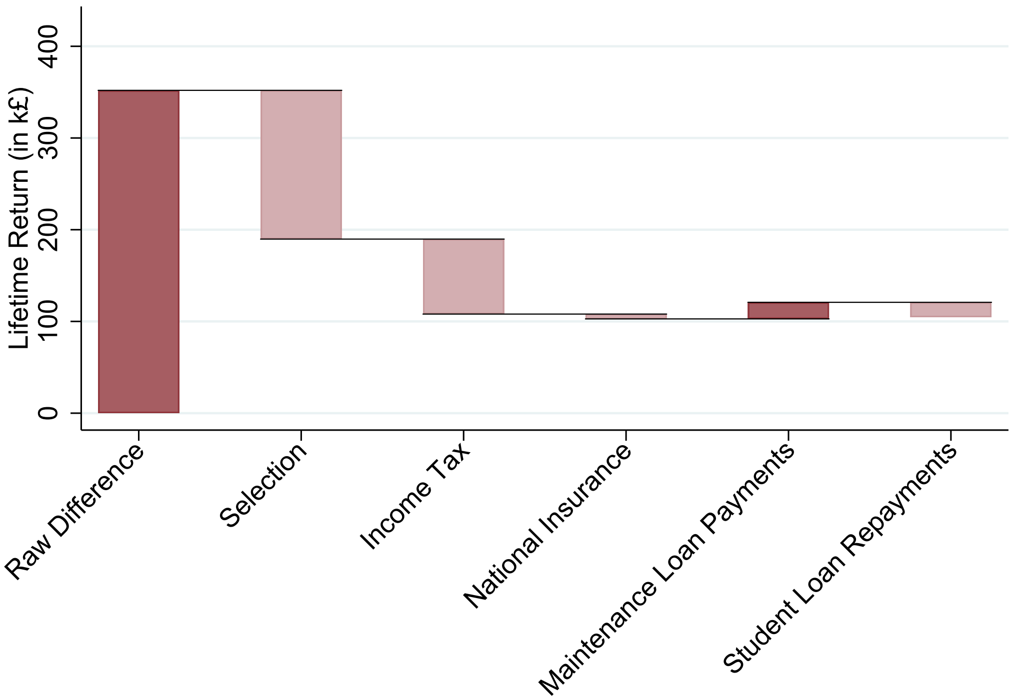
Figure 14: Overall Average DPV Lifetime Returns to HE for Women, Deaton/Paxson Method



Note: All figures are shown in 2018 prices and are discounted using Green Book discounting. The first bar shows the difference in raw earnings between those who did not attend HE, but have a KS5 record and at least five A*–C GCSEs, and those who started a first degree. The second bar shows how much of this difference in earnings is accounted for by differences in prior attainment and background characteristics. We then account for the extra income tax and National Insurance payments from graduates. The penultimate bar adds on the net present value of the maintenance loans payments received by students, and finally the last bar takes into account the net present value of student loan repayments over the life cycle. Dark blue bars indicate additions and light blue bars reductions.

As the tax and student loan system is highly non-linear, the implications of these differences for net private returns and exchequer returns are not straightforward. Figure 14 shows the net discounted lifetime return for women; it is the analogous graph to Figure 17 in the report. The overall discounted lifetime return is estimated to be about £50k, which is around half of the figure we arrived at taking the period view.

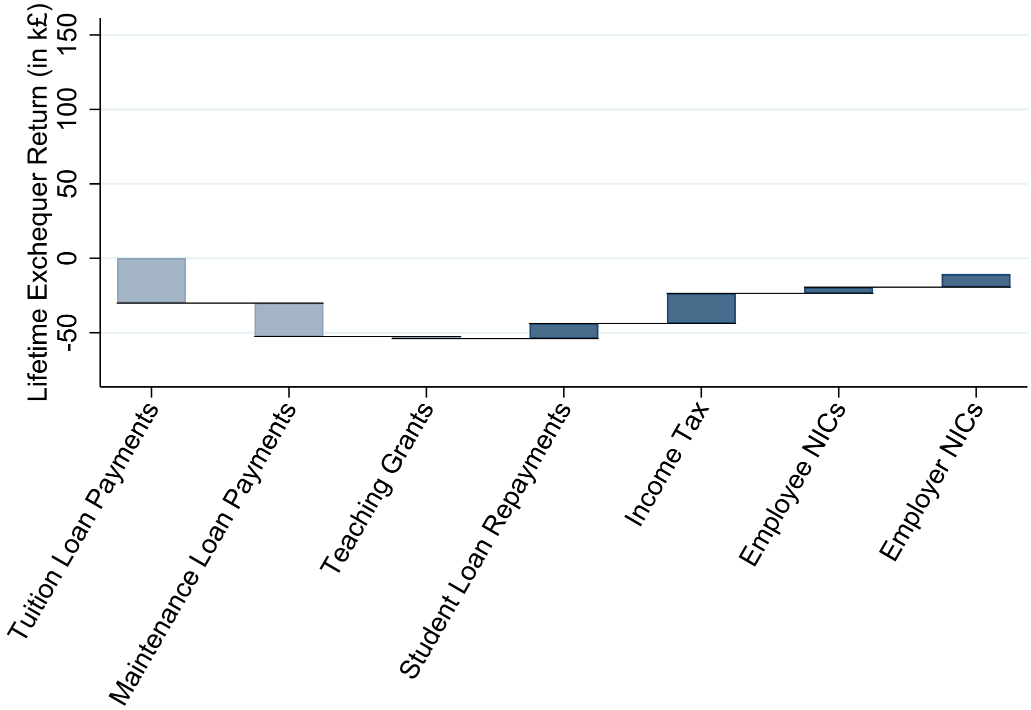
Figure 15: Overall Average DPV Lifetime Returns to HE for Men, Deaton/Paxson Method



Note: All figures are shown in 2018 prices and are discounted using Green Book discounting. The first bar shows the difference in raw earnings between those who did not attend HE, but have a KS5 record and at least five A*–C GCSEs, and those who started a first degree. The second bar shows how much of this difference in earnings is accounted for by differences in prior attainment and background characteristics. We then account for the extra income tax and National Insurance payments from graduates. The penultimate bar adds on the net present value of the maintenance loans payments received by students, and finally the last bar takes into account the net present value of student loan repayments over the life cycle. Dark red bars indicate additions and light red bars reductions.

Figure 15 shows the same information for men. Overall returns are about £100k or around double the amount for women. While this is somewhat less than in the main report, the difference is much less dramatic than for women.

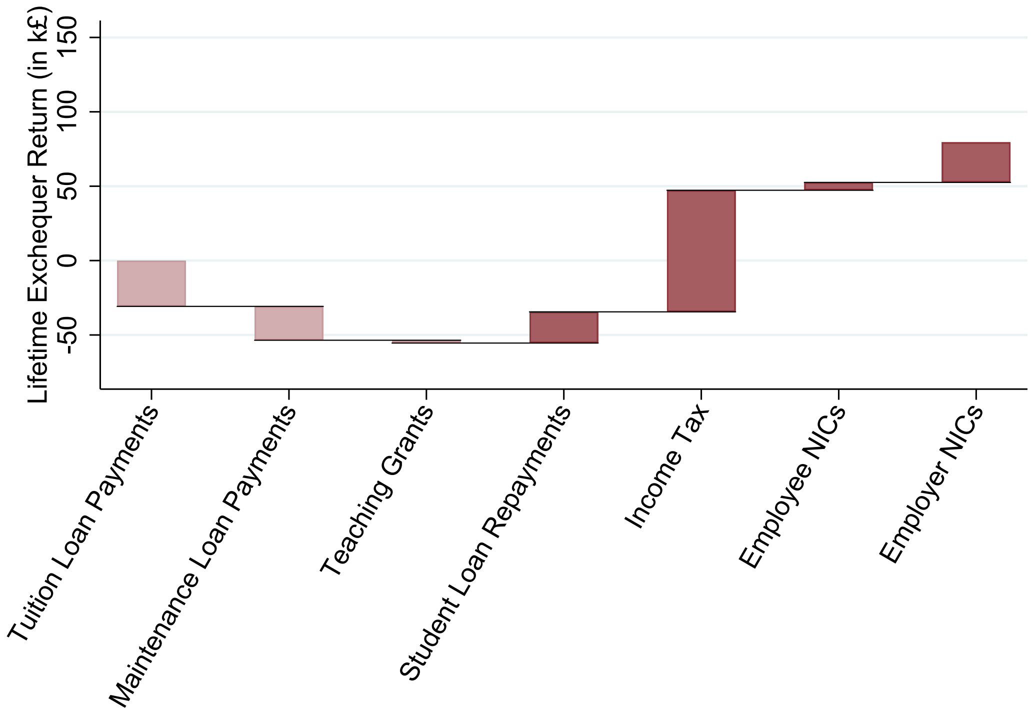
Figure 16: Overall Average DPV Lifetime Exchequer Returns to HE for Women, Deaton/Paxson Method



Note: All figures are shown in 2018 prices in £k and are discounted using Green Book discounting. The first two bars show the net present value of the tuition and maintenance loan payments to students. The next bar shows the net present value of teaching grants for high-cost subjects. Subsequent bars then show the net present value of government receipts in terms of student loan repayments and higher income tax and National Insurance payments over the life cycle from graduates compared with non-graduates. Dark blue bars indicate additions and light blue bars reductions.

The implications of the less optimistic forecast are on the whole larger for exchequer returns given that the tax system is progressive. Figure 16 presents average discounted lifetime exchequer returns for women. Overall returns are slightly negative at around -£10k, compared with the modest positive returns we found taking the period view. It should be noted, however, that this result depends heavily on the choice of a relatively high real discount rate, which is unlikely to represent the government's true cost of funding.

Figure 17: Overall Average DPV Lifetime Exchequer Returns to HE for Men, Deaton/Paxson Method



Note: All figures are shown in 2018 prices in £k and are discounted using Green Book discounting. The first two bars show the net present value of the tuition and maintenance loan payments to students. The next bar shows the net present value of teaching grants for high-cost subjects. Subsequent bars then show the net present value of government receipts in terms of student loan repayments and higher income tax and National Insurance payments over the life cycle from graduates compared with non-graduates. Dark red bars indicate additions and light red bars reductions.

Figure 17 is the equivalent for men. Due to the less optimistic earnings forecast on the cohort view, the resulting net exchequer returns are somewhat lower than in the main report, at around £80k.

On the whole, these results highlight that our lifetime earnings figures are subject to a considerable amount of forecasting uncertainty, especially for women. It should be noted, however, that forecasts on the period view can be strongly affected by short-term developments in the economy. In our case, the large depreciation of sterling in 2016, generally anaemic wage growth in the aftermath of the Great Recession, and an increase in working hours among non-HE women amidst a tightening labour market are all likely to have been important factors. However, the comparison with different methods of resolving the age–period–cohort problem does indicate that the risks to our earnings and returns forecasts might be skewed to the downside, especially for women.

G Further Robustness Checks

G.1 Testing the Copula Method

In this subsection, we test the performance of our model with regard to the evolution of earnings over time. This is particularly important for student loan accounting, as, due to the non-linearity of the system, less stable earnings will lead to higher payments on average. A direct measure of this stability is the **distribution of the difference in earnings ranks between different ages**.

Figure 18: Difference in Percentile Rank, Women Born in 1980/81, Simulation from Age 29/30

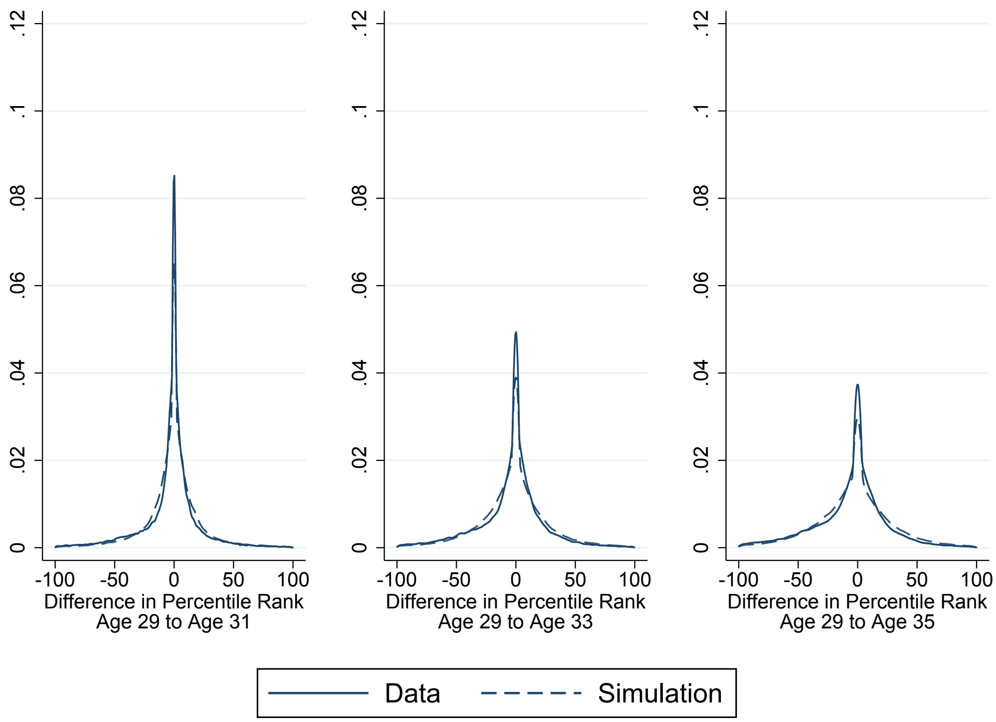


Figure 18 plots the density of the distribution of differences in percentile ranks for HE women from the 1980/81 cohort between age 29 and ages 31, 33 and 35. The solid line shows a density estimated from the data, whereas the dashed line gives the density from the simulation.[21] Zero earnings are counted as a percentile rank of zero.

For all age pairs, this density has a pronounced spike at zero, as most people's percentile rank in the income distribution changes little year by year. The spike becomes less pronounced as more distant age pairs are considered, as relative rises or falls in income are more common over longer time horizons. The model captures the patterns in the data well. The only noticeable difference is that the model modestly underestimates the probability mass near zero.

[21] In each case, we estimate the density using an Epanechnikov kernel with the bandwidth selected according to Silverman's rule of thumb. Percentile ranks are calculated from the distribution of all graduates who fulfil the criteria for inclusion in our sample (see above).

Figure 19: Difference in Percentile Rank, Men Born in 1980/81, Simulation from Age 29/30

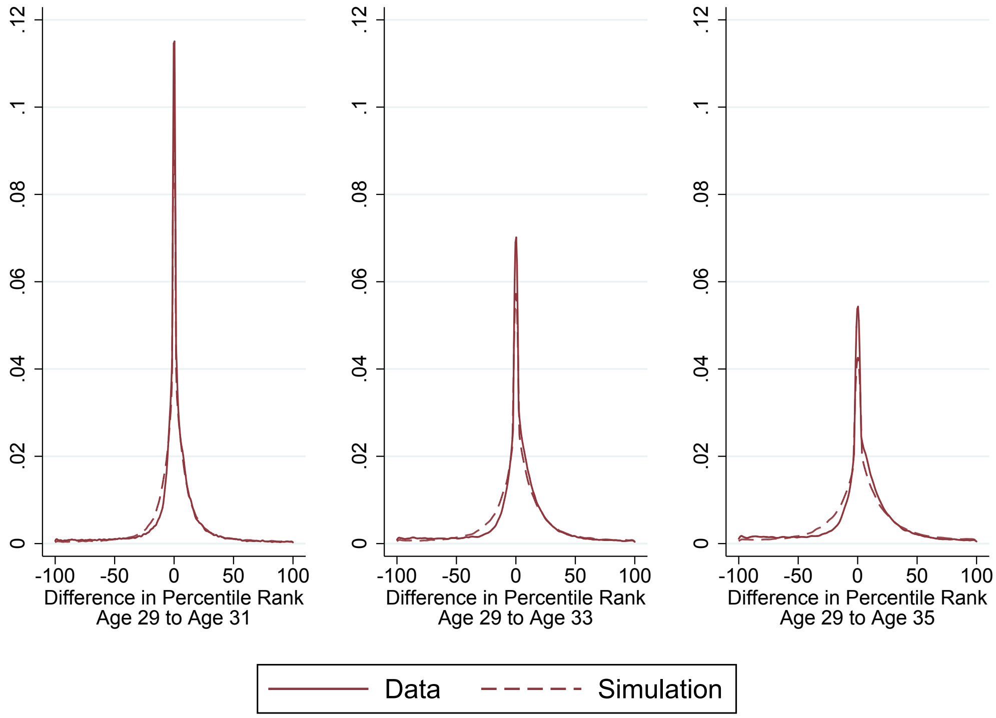


Figure 19 is the equivalent figure for men. The spikes at zero are even more pronounced for men, reflecting greater stability in men's earnings; otherwise the patterns in the data are similar. Again the model captures the patterns in the data well except for a modest underestimate of probability mass near zero. More noticeably than for women, the model also slightly overstates the probability of modest declines in earnings rank, which are rare in the data.

A second test for our earnings model is whether the model can roughly capture the **dependence of earnings on initial conditions**. While our returns estimates themselves do not directly rely on this, as we hold the parameters on initial conditions constant, a very poor fit might still lead to some bias in our calculations of student loan repayments and the distribution of lifetime earnings.[22]

[22]There are two reasons for this. First, student loan sizes directly depend on parents' socio-economic situation. Second, large errors in the dependence of later-life earnings on initial conditions would worsen the fit between people's pre-30 and post-30 earnings histories.

Figure 20: Earnings Difference between Top and Bottom POLAR Quintile, 1980/81 Cohort

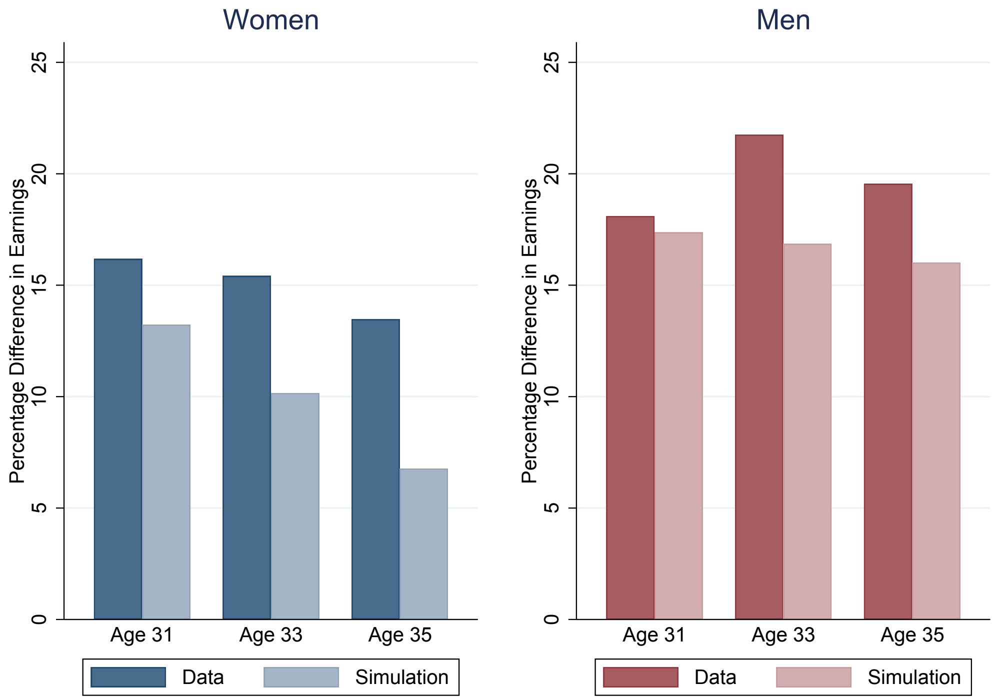


Figure 20 shows the percentage earnings difference between earners from the top and bottom POLAR quintiles for men and women of the 1980/81 cohort (we use POLAR quintiles as very few other background characteristics are available in the HESA data). At age 31, women from the 1980/81 cohort in employment earned around 16% more than those from the bottom quintile on average, while men earned around 18% more.[23]

It is clear that while the percentage earnings difference between those from low and high POLAR quintiles stayed roughly constant in the early 30s for the 1980/81 cohort, the simulation predicts a fall in dependence over time for women and, to a lesser extent, for men. This 'washing out' of initial conditions is attributable to the reliance of the model on data from the past two periods only, whereas in reality initial conditions such as socio-economic status are predictive of future earnings even conditional on observed earnings in the past two periods.

This observation provides an important justification for our decision to hold coefficients on initial conditions constant at age 30 values. However, it should be noted that this process of 'washing out' is slow enough so as to be unlikely to materially affect our results. Five years into the simulation, the model still predicts substantial dependence of earnings on initial conditions, even though initial conditions do not explicitly feature in the model.

A third test is a different specification of the earnings model. Instead of using the first and second lags of the earnings rank in the copula and unemployment models, we have run the model

[23] In keeping with our estimation methodology, all averages are average log differences on Winsorised data, converted into percentage terms.

using the **first and third lags of the earnings rank** to test whether our results are robust to this change in specification.

The results are encouraging. The overall returns in net present value terms are very similar to our main estimates. Using the third instead of the second lag increases the overall return for men by around £5,000 and lowers it by around £1,000 for women. The RAB charge, which we would expect to be sensitive to changes in the persistence of earnings, is virtually unaffected when we use the third lag instead of the second lag. Overall, none of these changes are economically significant. We conclude that other modelling choices are much more important than the precise specification of the copula and unemployment models.

G.2 Fixed Parameters on Background Conditions

In the main estimates of our report, we fix the parameters on background conditions at their age 30 levels (see Section 3.3). In this subsection, we compare this assumption with three alternative assumptions: fixing the parameters on background conditions at their age 29 levels, not fixing them at all but relying on the earnings model to capture dependence on background conditions, and not controlling for background conditions at all.

Figure 21: Overall Returns with Alternative Assumptions on Background Conditions, Women

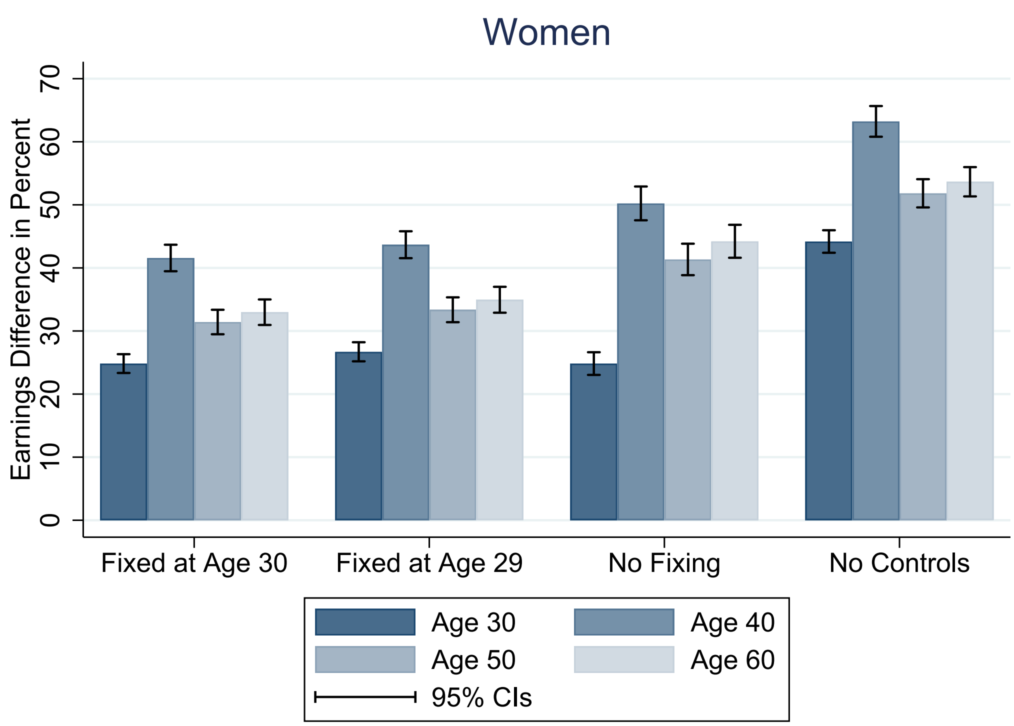


Figure 21 shows estimated returns by age for women under our main assumption and these three alternative assumptions. Estimated returns with parameters fixed at age 29 are slightly larger, which is consistent with the general trend that the importance of controlling for initial conditions rises throughout students' 20s. Estimated returns are significantly higher if background conditions are not fixed at all, and rise towards estimated returns from the model without any controls over the life cycle. However, even at age 60, estimated returns without fixed background conditions are still significantly lower than those without any controls.

While fixing the parameters on background conditions at age 30 is clearly a substantive assumption that is unlikely to hold precisely in reality, fully relying on the copula model to replicate dependence on background conditions also requires a strong assumption – namely, that current earnings y_{it} only depend on past earnings $y_{i,t-1}$ and $y_{i,t-2}$ and on shocks that are independent of background conditions. In particular, this implies that $\{y_{it}|y_{i,t-1}, y_{i,t-2}\} \perp x_i$, where x_i is the vector of initial conditions. Intuitively, this means that with a gender/subject/institution-type group, two people with the same earnings in any two consecutive years have the same expected future earnings independent of their background conditions. This assumption is likely to be more problematic for women than for men: for instance, two women might have the same earnings in two years in which they take maternity leave, but have radically different earnings paths thereafter that are attributable to their different initial conditions.

Figure 22: Overall Returns with Alternative Assumptions on Background Conditions, Men

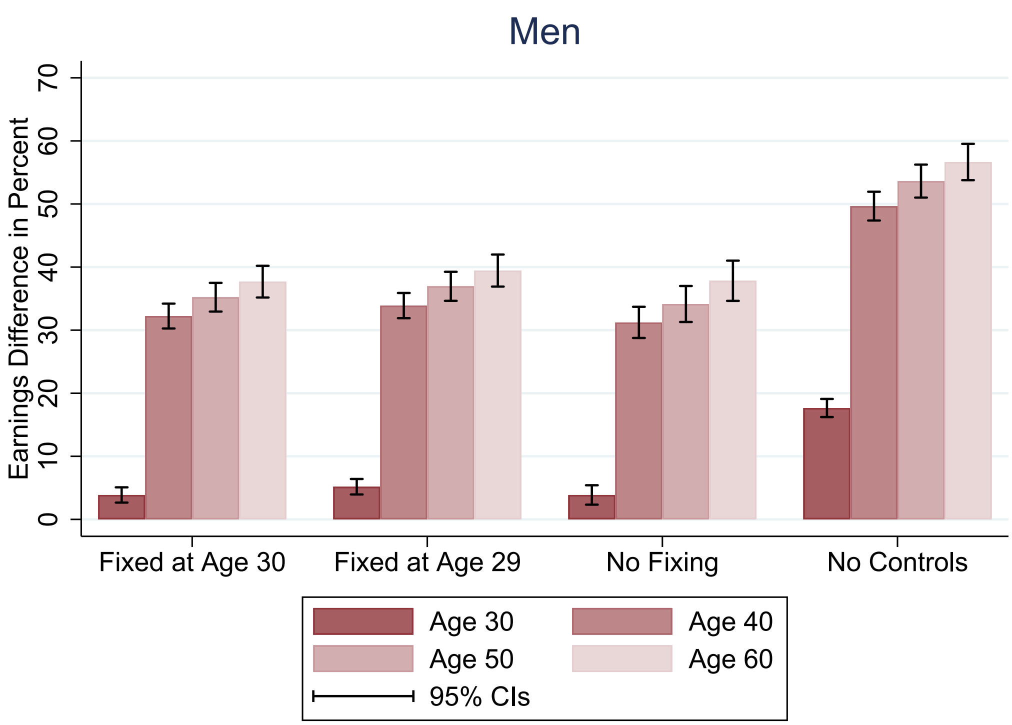


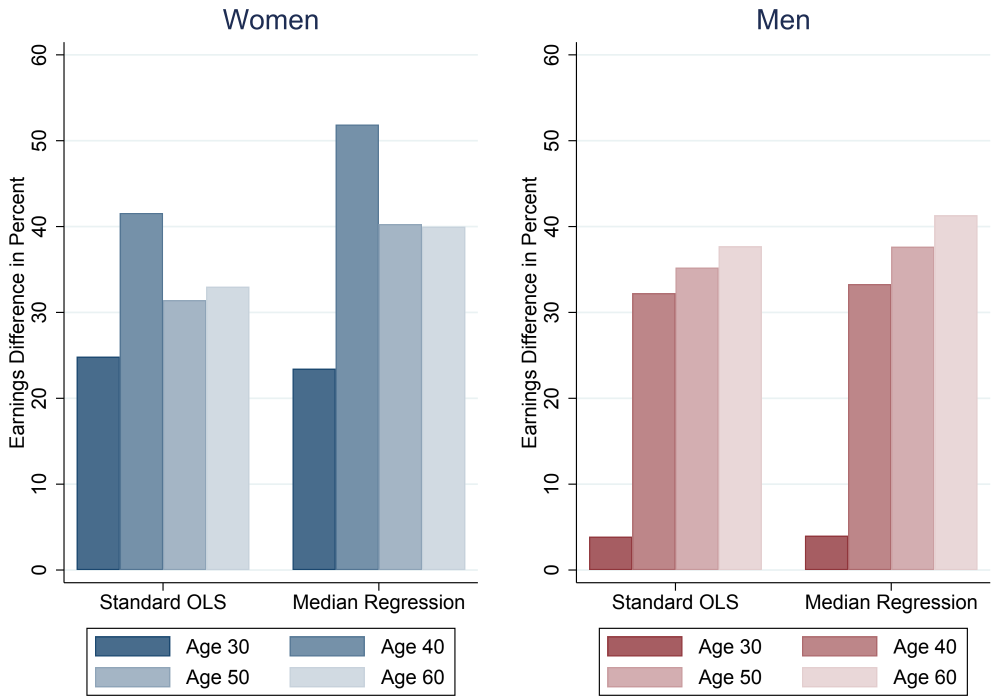
Figure 22 shows the equivalent results for men. As for women, returns with parameters on background conditions fixed at age 30 levels are slightly lower than when they are fixed at age 29 levels. In contrast to the results for women, results for men are essentially identical if parameters on background conditions are not fixed at all. This supports the hypothesis that our earnings model can capture the earnings dynamics of men much better than those of women due to their greater attachment to the labour market. Overall, fixing parameters at age 30 appears to be the most reasonable course of action to us, as it avoids the upward bias in returns that we would expect, especially for women, if we relied on the copula model alone.

G.3 Median Regression Results

Another concern is that our results might be driven by outliers in the distribution, so that the average returns may not represent the return of a typical student. In particular, those at the top of the (potential) earnings distribution tend to benefit the most in absolute terms from a university education. The standard way of addressing this concern, which we follow throughout our analysis, is to take the natural logarithm of earnings. If university has roughly the same relative effect on earnings throughout the distribution, percentage returns available to all students will be well approximated by the average effect on log earnings.

However, if relative returns are not the same across the distribution, this may fail to hold exactly, and our estimated earnings returns may differ from the returns available to the median earner. In order to test for this possibility, we have estimated median regressions, a version of linear regression that estimates the conditional *median* of the dependent variable as a linear function of the covariates. Figure 23 gives a comparison of returns estimates using standard OLS and median regression for both women and men.

Figure 23: Overall Returns, Standard OLS and Median Regression



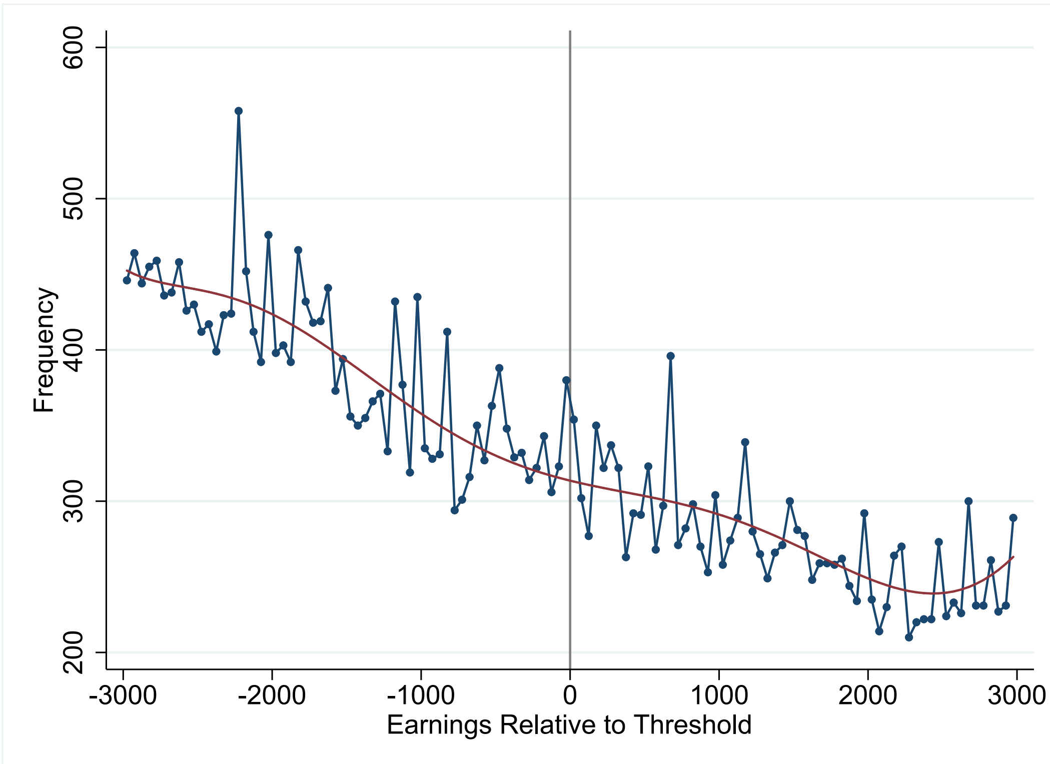
Using median regression, women's returns are significantly higher later in life than when standard OLS is used. This may indicate that women near the middle of the potential earnings distribution may gain the most from going to university. However, the differences in these estimates should not be over-interpreted; they may well be the results of differential patterns of part-time work between HE and non-HE women that do not represent differences in potential earnings. For men, median regression yields virtually identical results to standard OLS, indicating that results for typical working men are similar to those for average male earners.

G.4 Labour Supply Responses to the New Student Loan System

A final reason to be concerned about the robustness of our estimates is that changes in the tax and student loan system might have affected the labour supply and thus the pre-tax earnings of recent cohorts. In particular, as has been widely documented, only a small share of current graduates are expected to pay off their student loan in full. For all other students, the new student loan system effectively imposes an additional tax on earnings above the repayment threshold. Insofar as this new tax discourages students from work, we would expect our projection of pre-tax earnings for previous cohorts to overstate the expected pre-tax earnings for more recent cohorts.

However, this problem only arises if students think of student loan repayments as a tax and adjust their labour supply behaviour accordingly. As we do not observe hours worked, we cannot directly assess labour supply. However, economic theory suggests that insofar as they affect labour supply, discrete thresholds in the tax system should lead to 'bunching', i.e. a clustering of taxpayers with annual incomes near the threshold. The more bunching is observed, the larger is likely to be the labour supply elasticity, i.e. the percentage effect on hours worked of a 1% rise in the wage (or, equivalently, of a 1% fall in the tax rate).

Figure 24: Binned 2016 Earnings around the Repayment Threshold for the 2012 Matriculation Cohort



Note: Students were included in the graph if they earned within £3,000 either side of the earnings threshold (£21,000). Earnings have been sorted into one of 120 bins with a width of £50 each, and the number of individuals in each bin is shown on the vertical axis. The vertical grey line indicates the repayment threshold. The red line indicates the counterfactual distribution estimated using the method of Chetty et al. (2011).

Figure 24 shows 2016 earnings data for the 2012 matriculation cohort, the first cohort entering university after a significant increase in the cap on tuition fees, which led to a large rise in average loan sizes. Students were included in the graph if they earned within £3,000 either side of the earnings threshold (£21,000). Earnings have been sorted into one of 120 bins with a width of £50 each, and the number of individuals in each bin is shown on the vertical axis. No perceptible bunching or clustering near the threshold is observed.

Table 4: Labour Supply Elasticities Estimated Using the Bunching Estimator of Chetty et al. (2011)

Overall	**Women**	**Men**	**Low-Earning Subjects**	**Low-Earning Institutions**
0.010	0.006	0.016	-0.001	-0.016
(0.005)	(0.006)	(0.007)	(0.008)	(0.009)
N = 39,683	*N = 21,854*	*N = 17,829*	*N = 14,369*	*N = 8,783*

Note: Standard errors in parentheses were obtained using 500 bootstrap samples for each estimate. Subjects and institutions are ranked by average pre-tax lifetime earnings (Green Book discounting). Institutions with fewer than 100 students in the sample are excluded. Individuals who studied education as their main subject are excluded from the calculation, as teachers are paid at nationally standardised payscales, which leads to bunching that is unrelated to the tax and student loan system.

Table 4 shows estimated labour supply elasticities calculated using the bunching estimator of Chetty et al. (2011). Results are shown overall, and separately for men and women. We also separately show results for students of the lowest-earning third of subjects and for those who studied at the lowest-earning third of institutions.[24] Estimated elasticities are very small and largely not significantly different from zero. Although the estimated overall elasticity and the overall elasticity for men are significantly greater than zero at the 95% confidence level, they are too small in magnitude to be economically significant. We conclude that labour supply effects of changes in the student loan system are likely to be small and do not pose an important risk to our forecast.

[24] Subjects and institutions are ranked by average pre-tax lifetime earnings (Green book discounting). Institutions with fewer than 100 students in the sample are excluded. Individuals who studied education as their main subject are excluded from the calculation, as teachers are paid at nationally standardised payscales, which leads to bunching that is unrelated to the tax and student loan system.

H List of Universities by Group

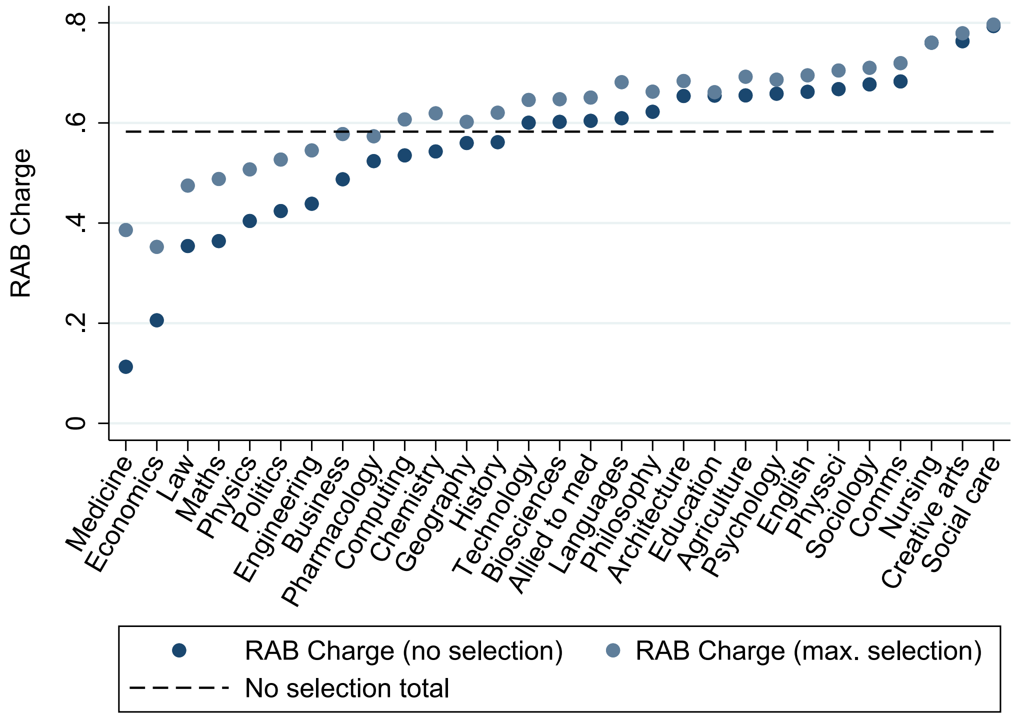
Table 5: List of Universities by Group

Russell Group	**Pre-1992 Universities**	**Other (More Selective)**	**Other (Least Selective)**
Cardiff University	Aston University	Arts Institute at Bournemouth	Anglia Ruskin University
Imperial College London	Bangor University	BPP University	Buckinghamshire New University
King's College London	Birkbeck College	Bath Spa University	Canterbury Christ Church University
London School of Economics	Brunel University	Birmingham City University	De Montfort University
Oxford University	City University	Bishop Grosseteste University	Edge Hill University
Queen Mary, University of London	Goldsmiths College	Bournemouth University	Edinburgh Napier University
Queen's University Belfast	Heriot-Watt University, Edinburgh	Cardiff Metropolitan University	Glasgow Caledonian University
University College London	Keele University	Central School of Speech and Drama	Glyndwr University
University of Birmingham	Lancaster University	Conservatoire for Dance and Drama	Kingston University
University of Bristol	Loughborough University	Courtauld Institute of Art	Leeds Trinity University College
University of Cambridge	Prifysgol Aberystwyth	Coventry University	Liverpool Hope University
University of Durham	Royal Holloway	Glasgow School of Art	London Metropolitan University
University of Edinburgh	Royal Veterinary College	GSM London Ltd	London South Bank University
University of Exeter	School of Oriental and African Studies	Guildhall School of Music and Dance	Middlesex University
University of Glasgow	St George's Hospital Medical School	Harper Adams University College	Newman University, Birmingham
University of Leeds	Swansea University	Heythrop College	Robert Gordon University
University of Liverpool	University of Aberdeen	Leeds City College	Roehampton University
University of Manchester	University of Bath	Leeds College of Art and Design	Southampton Solent University
University of Newcastle Upon Tyne	University of Bradford	Leeds Metropolitan University	St Mary's University, Twickenham
University of Nottingham	University of Buckingham	Liverpool Institute for Performing Arts	Staffordshire University
University of Sheffield	University of Dundee	Liverpool John Moores University	University Campus Suffolk
University of Southampton	University of East Anglia	Manchester Metropolitan University	University College Birmingham
University of Warwick	University of Essex	Norwich University College of the Arts	University for the Creative Arts
University of York	University of Hull	Nottingham Trent University	University of Abertay Dundee
	University of Kent	Oxford Brookes University	University of Bedfordshire
	University of Leicester	Plymouth College of Art	University of Bolton
	University of London	Queen Margaret University, Edinburgh	University of Central Lancashire
	University of Reading	Ravensbourne	University of Derby
	University of Salford	Richmond, The American Intl University	University of East London
	University of St Andrews	Rose Bruford College	University of Greenwich
	University of Stirling	Royal Academy of Music	University of Hertfordshire
	University of Strathclyde	Royal Agricultural College	University of Northampton
	University of Surrey	Royal College of Music	University of St Mark and St John
	University of Sussex	Royal Conservatoire of Scotland	University of Sunderland
	University of Ulster	Royal Northern College of Nursing	University of Teesside
		Scotland's Rural College	University of the West of Scotland
		Sheffield Hallam University	University of Wales Trinity Saint David
		Stranmillis University College	University of West London
		Trinity Laban Conservatoire	University of Westminster
		University College Falmouth	University of Wolverhampton
		University of Brighton	
		University of Chester	
		University of Chichester	
		University of Cumbria	
		University of Glamorgan	
		University of Gloucestershire	
		University of Huddersfield	
		University of Lincoln	
		University of Northumbria at Newcastle	
		University of Plymouth	
		University of Portsmouth	
		University of Winchester	
		University of Worcester	
		University of the Arts London	
		University of the Highlands and Islands	
		University of the West of England	
		Writtle College	
		York St John University College	

Note: 'Other (Least Selective)' contains the 40 least selective universities by total GCSE score of students from the 2004 to 2007 GCSE cohorts (excluding universities with very few observations).

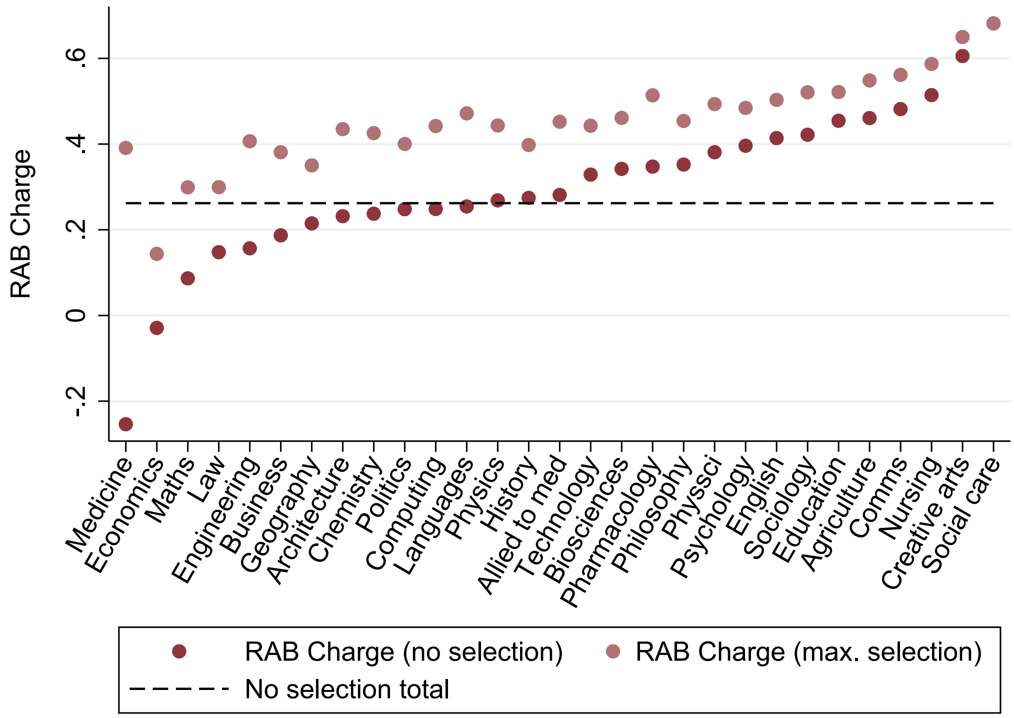
I Implied RAB charges by subject

Figure 25: RAB Charge for Women



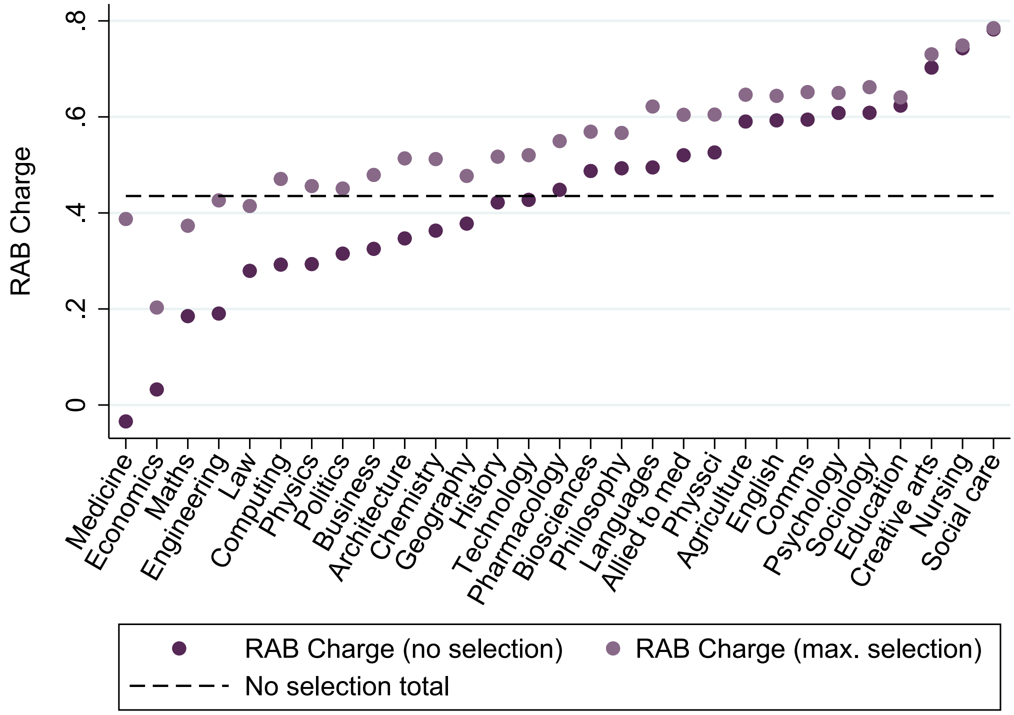
Note: Implied RAB Charges for women by subject, using a discount rate of 0.7%. 'RAB Charge (no selection)' indicates the estimated RAB charge if a random 10% of students do not take out loans (in order to match the take-up rate of loans in the data), and all others take up the full amount they are entitled to. 'RAB Charge (maximum selection)' is the RAB charge if the 10% of students with the most favourable future repayment profiles do not take out any loan. As student loans are not our focus in this report, these results should not be taken as conclusive.

Figure 26: RAB Charge for Men



Note: Implied RAB Charges for men by subject, using a discount rate of 0.7%. 'RAB Charge (no selection)' indicates the estimated RAB charge if a random 10% of students do not take out loans (in order to match the take-up rate of loans in the data), and all others take up the full amount they are entitled to. 'RAB Charge (maximum selection)' is the RAB charge if the 10% of students with the most favourable future repayment profiles do not take out any loan. Negative RAB charges can occur because interest rates on student loans are generally higher than the discount rate used to calculate the RAB charge. As student loans are not our focus in this report, these results should not be taken as conclusive.

Figure 27: RAB Charge Pooling across Genders



Note: RAB Charges by subject pooling across genders, using a discount rate of 0.7%. 'RAB Charge (no selection)' indicates the estimated RAB charge if a random 10% of students do not take out loans (in order to match the take-up rate of loans in the data), and all others take up the full amount they are entitled to. 'RAB Charge (maximum selection)' is the RAB charge if the 10% of students with the most favourable future repayment profiles do not take out any loan. Negative RAB charges can occur because interest rates on student loans are generally higher than the discount rate used to calculate the RAB charge. As student loans are not our focus in this report, these results should not be taken as conclusive.

References

Belfield, Chris, Jack Britton, Franz Buscha, Lorraine Dearden, Matt Dickson, Laura van der Erve, Luke Sibieta, Anna Vignoles, Ian Walker, and Yu Zhu, "The impact of undergraduate degrees on early-career earnings," *Department for Education Report*, 2018.

Chamon, Marcos D and Eswar S Prasad, "Why are saving rates of urban households in China rising?," *American Economic Journal: Macroeconomics*, 2010, *2* (1), 93–130.

Chetty, Raj, John N Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: evidence from Danish tax records," *Quarterly Journal of Economics*, 2011, *126* (2), 749–804.

Deaton, Angus S and Christina Paxson, "Saving, growth, and aging in Taiwan," in "Studies in the Economics of Aging," University of Chicago Press, 1994, pp. 331–362.

Schulhofer-Wohl, Sam, "The age-time-cohort problem and the identification of structural parameters in life-cycle models," *Quantitative Economics*, 2018, *9* (2), 643–658.