Would additional investment in skills benefit areas of the country that are poorer performing economically?

Skills and Productivity Board

Arun Advani, Jack Cornish and Claire Crawford

May 2022

Advani, A.: Assistant Professor, Department of Economics, University of Warwick and SPB Member. Cornish, J.: Central Analysis Division, Department for Education. Crawford, C.: Associate Professor, Centre for Education Policy and Equalising Opportunities, University College London and SPB Member. The authors would like to thank the Department for Education LEO Strategic Analysis team, the Skills and Productivity Board (SPB) Secretariat, and other members of the SPB for their insights, support and helpful comments

Executive Summary

There are very high returns to qualifications, in terms of both employment and earnings. Reaching Level 2 (the equivalent of 5 GCSEs at A*-C) – something almost one in six people under 27 do not do – increases the probability of employment for women (men) by around 19 (10) percentage points and earnings by around 22% (13%).

The returns to qualifications are higher in areas of the country that are economically poorer performing. Reaching Level 2 increases employment prospects for women (men) by around 25 (13) percentage points for individuals living in areas in the bottom quartile based on the Index of Multiple Deprivation (IMD, the 25% most deprived areas of the country), but only around 18 (8) percentage points in other areas. Differences are smaller, but still present, at higher levels of qualification.

Mobility across labour markets is relatively low, particularly among non-graduates. Just one in six people whose highest qualification is below Level 6 (degree level) has moved commuting area by the age of 27 compared to around one in three with an undergraduate or postgraduate degree. Local moves are more common for non-graduates, with more than half of both graduates and non-graduates moving neighbourhoods by age 27.

Mobility across labour markets is lower for individuals from areas of the country that are economically poorer performing. Among non-graduates, individuals who grow up in areas in the bottom quartile based on the IMD, are around 40% less likely to move commuting zone than those who grow up in the top quartile. Among graduates the difference is smaller but still sizeable, with those growing up in the bottom quartile around 25% less likely to move than those from the top quartile.

Low mobility among non-graduates might be driven by low (short-term) returns to moving. Among those non-graduates who do move, men see a small increase in employment prospects but no change in earnings, while women are less likely to be employed but have slightly higher earnings if they are employed. By contrast, for graduates there are strong positive effects of moving on both employment and earnings (although these are reduced somewhat if we account for differences in the cost of living across areas). For both graduates and non-graduates, these effects likely underestimate the career returns to moving, being based on only what we observe by age 27.

Investment in skills alone is unlikely to be sufficient to 'level up' economically poorer performing areas. Our findings suggest that the returns to education in poorer-performing areas are strong, with the benefits of upskilling highly likely to remain within those areas, highlighting the importance of skills investments for the levelling up agenda. However, around a third of the difference in earnings between areas cannot be explained by individual characteristics such as education or skills, highlighting that other features of the areas, aside from the individuals that live there, play an important role in driving earnings. This suggests that complementary investments to improve those features of poorer-performing areas may be required to fully realise the benefits of skills investments, and truly 'level up'.

Contents

List of tables

List of figures

1. Introduction

The UK has large spatial disparities compared to most OECD countries, with differences in productivity mirrored by differences in earnings and education (HM Government, Levelling Up White Paper, 2022). It is natural, therefore, to think that improving education or skills in poorer performing areas would lead to higher earnings and productivity in those areas, thus reducing inequalities across areas. An alternative view is that differences in education or skills across regions arise not because of a lack of investment in education or skills in those areas, but because of a lack of opportunities to use those skills, leading to an exodus of skilled workers. In this case, the differences we observe across areas might reflect 'sorting': individuals moving to where they can get the highest returns for their skills. This would imply that differences in education levels across regions are a symptom, not a cause, of regional disparities.

This report explores the potential for investments in education or skills in poorer-performing areas to improve productivity in those areas.[1] In particular, we look at the benefits of investments, first from the perspective of the individual receiving education or training, and then from an area perspective, taking into account the possibility that individuals move.

Looking first at the individual benefits to gaining qualifications, we find evidence of strong returns to the acquisition of higher levels of education, particularly for women. Perhaps surprisingly, these returns are very similar across areas: the percentage point increase in the likelihood of being in work or the percentage increase in earnings from undertaking additional qualifications is the broadly the same in poorer-performing and better-performing areas, although the level of earnings at each qualification level tends to be lower in poorer-performing areas, leaving open the possibility that individuals could earn more by moving.

In fact, this does not appear to lead to significant outflows from poorer-performing areas. Indeed, our findings suggest that a substantial proportion of the benefits from the acquisition of higher qualifications – at least up to degree level – is likely to accrue to the areas in which individuals gained them. While more than half of individuals move across small geographic areas between age 16 and age 27, the vast majority only move short distances (less than 10km). Internal migration across commuting zones (travel to work

[1] Together with Keep (2022) and Mayhew (2022), this paper forms the SPB's response to the third of the questions set to the Board by the former Secretary of State, Gavin Williamson: "What is the role of skills and the skills system in promoting productivity growth in areas of the country that are poorer performing economically?". Advani, Crawford, Keep and Mayhew (2022) draws together the lessons from across these pieces of work.

areas, TTWAs) within the UK is relatively rare, especially amongst non-graduates: only one in six people whose highest qualification is below degree level has moved TTWA by the age of 27, compared with more than one in three with an undergraduate or postgraduate degree. This may be because the benefits associated with moving across areas are relatively low in both employment and earnings terms for non-graduates – at least in the short-term – although it could also be that the barriers to moving are high. Moreover, individuals in poorer-performing areas are, if anything, less likely to move, and amongst those who do move, there is migration across all types of areas, rather than clear movement from poorer to better-performing areas.

Taken together, these findings underscore the potentially positive role that education and skills investments can play in levelling up poorer performing regions of the UK – although it should be noted that this does not imply that investment in education or skills is *sufficient* to improve outcomes across locations. One way to see this is to compare what proportion of the differences in earnings across areas can be explained by differences between the areas themselves rather than differences between the people who live there. While the details of this are subtle, and explained in detail below, the key finding is that around one-third of the differences in earnings across areas cannot be attributed to individual characteristics such as qualifications. Hence, investment in education and skills alone will not be sufficient to eliminate inequalities across areas, and complementary investments in other types of 'capital' may be required (see also Advani, Crawford, Keep and Mayhew, 2022).

2. Data and context

2.1 Data sources

The analysis in this report uses the Longitudinal Education Outcomes (LEO) dataset[2], which combines administrative data on:

- Employment and earnings data from HMRC, and home address records from DWP;
- School records from the National Pupil Database; and
- Records of Further and Higher Education qualifications.

We follow five cohorts of school leavers in England, who took their GCSEs (were aged 16) between the 2001/02 academic year and the 2005/06 academic year. We have data up to the 2017/18 tax year, meaning that we observe earnings from employment for all of our cohorts between ages 22 and 27 inclusive, and for the two oldest cohorts (those taking GCSEs in 2001/02 and 2002/03) up to age 30. We do not have access to earnings from self-employment. We use annualised earnings and adjust them to 2017/18 prices using CPIH. When considering employment probabilities, we regard an individual as being in employment if they have non-zero days worked in a given tax year.

In parts of the analysis, we adjust earnings for the local cost of living. This methodology follows Britton et al. (2021b), extended across more years of data: data on price levels, rent and housing costs[3] are weighted using Consumer Prices Index including Owner Occupiers' Housing (CPIH) weights to produce a cost-of-living index at Local Authority (LA) level, for each year between 2011 and 2018**.** Individuals' earnings can be adjusted using this index based on the LA in which they live to account for these geographic differences in cost of living. This index is imperfect and should be treated as indicative of patterns of living costs, rather than authoritative. Price level data is the most tenuous part of the index, as this is only available at regional level. To avoid over-reliance on this data, in robustness checks we also use a more conservative index where price levels are assumed to be equal across the country, and cost of living variation is entirely driven by rent and housing costs.

Location at age 16 is derived from the school census, and location for ages 22 onwards is based on address records held by DWP. Limitations in the DWP data mean we only have location data prior to 2012 in cases where the individual still lived at that address in

[2] See the LEO privacy notice for useful information on LEO, fuzzy matching and Subject Access Requests Longitudinal Education Outcomes (LEO): privacy notice - GOV.UK (www.gov.uk)

[3] ONS - Private Rental Market Statistics ; ONS - Mean house prices for administrative geographies ; ONS – Relative regional consumer price levels of goods and services (both 2011 and 2016)

October 2012. That leaves six years of complete location records from 2012 through to 2017/18. We restrict our sample to individuals who lived in England in each year for which we have location data.

For analysing differences across areas, we consider two different geographies: Middle layer Super Output Areas (MSOAs), each of which contain between 5,000 and 15,000 individuals (or 2,000-6,000 households) and of which there are around 6,800 in England, and 149 Travel To Work Areas (TTWAs), which are a geography created to approximate labour market areas based on census data.[4] MSOAs allow us to explore granular geographic variation, whilst TTWAs are larger and represent relatively more coherent economic areas.

We run most of our analysis separately by gender. We also account for a parsimonious set of individual controls, including free school meal eligibility at age 16 (as an indicator of an individual's socio-economic status in childhood), ethnicity (categorised as White British, White Other, Black African, Black Caribbean, Black Other, Bangladeshi, Chinese, Indian, Pakistani, and Other), and attainment at age 11 (continuous points score at Key Stage 2).

The LEO data follows the vast majority of individuals in the GCSE cohorts of interest, and therefore gives a very large sample size. Broken down by cohort, the number of individuals with earnings and location data available at age 27, and location data at age 16, are:

Table 1: Sample size by KS4 cohort

Key Stage 4 Cohort	Number of Individuals
2001-02	483,005
2002-03	516,580
2003-04	534,998
2004-05	534,983
2005-06	548,806

Notes: Count includes all individuals in our cleaned LEO dataset, for whom we have location data at age 16, and who do not have a location outside England recorded at any age.

[4] Specifically, at least 75% of the area's resident workforce work in the area and at least 75% of people who work in the area also live in the area. (Note these constraints are loosened for some areas with a working population greater than 25,000 people).

An important limitation of this dataset is the potential for mis recording of home addresses. Although these are administrative data, it is likely that a significant number of young adults do not keep their DWP records up-to-date with their actual address, e.g. failing to change their address when moving out of their parents' home – although it is worth noting that three quarters of those we identify to have moved by age 27 do so between ages 22 and 27, i.e. are identified from change of address notifications. Nonetheless, Britton et al. (2021b) estimate that the address data could underestimate the true number of individuals moving away from their parental home by as much as 30%.

This likely underreporting of mobility is a key caveat to bear in mind when interpreting our findings, and those of other papers drawing on LEO data for this purpose. While this will tend to bias down our estimates of overall mobility, unless underreporting is correlated with individual characteristics, it will not affect our relative comparisons of who is likely to move, nor the returns to moving.

2.2 Geographic areas

Levelling up is the government's "*plan to transform the UK by spreading opportunity and prosperity to all parts of it*."[5] We take this to mean a policy focus on places in which there are currently fewer opportunities and lower incomes – those which are 'poorly performing' on these measures. There are several ways to measure the performance of a place, and indeed, the government's White Paper on the issue includes a substantial number of metrics on which Levelling Up will be tracked and assessed.

We consider two such metrics which capture prosperity and opportunity in broad strokes:

1. Index of Multiple Deprivation[6] – aggregated to MSOA level. The index ranks each Lower Layer Super Output Area (LSOA) in England based on seven domains of deprivation – income, employment, education, health, crime, barriers to housing and services, and living environment. The index is a broad-based measure of many of the economic and social domains of concern to the government's Levelling Up plan and is available at granular geographies.

2. Employment Rate (ages 16-64)[7] – available at TTWA level. This provides a narrower focus on labour market conditions.

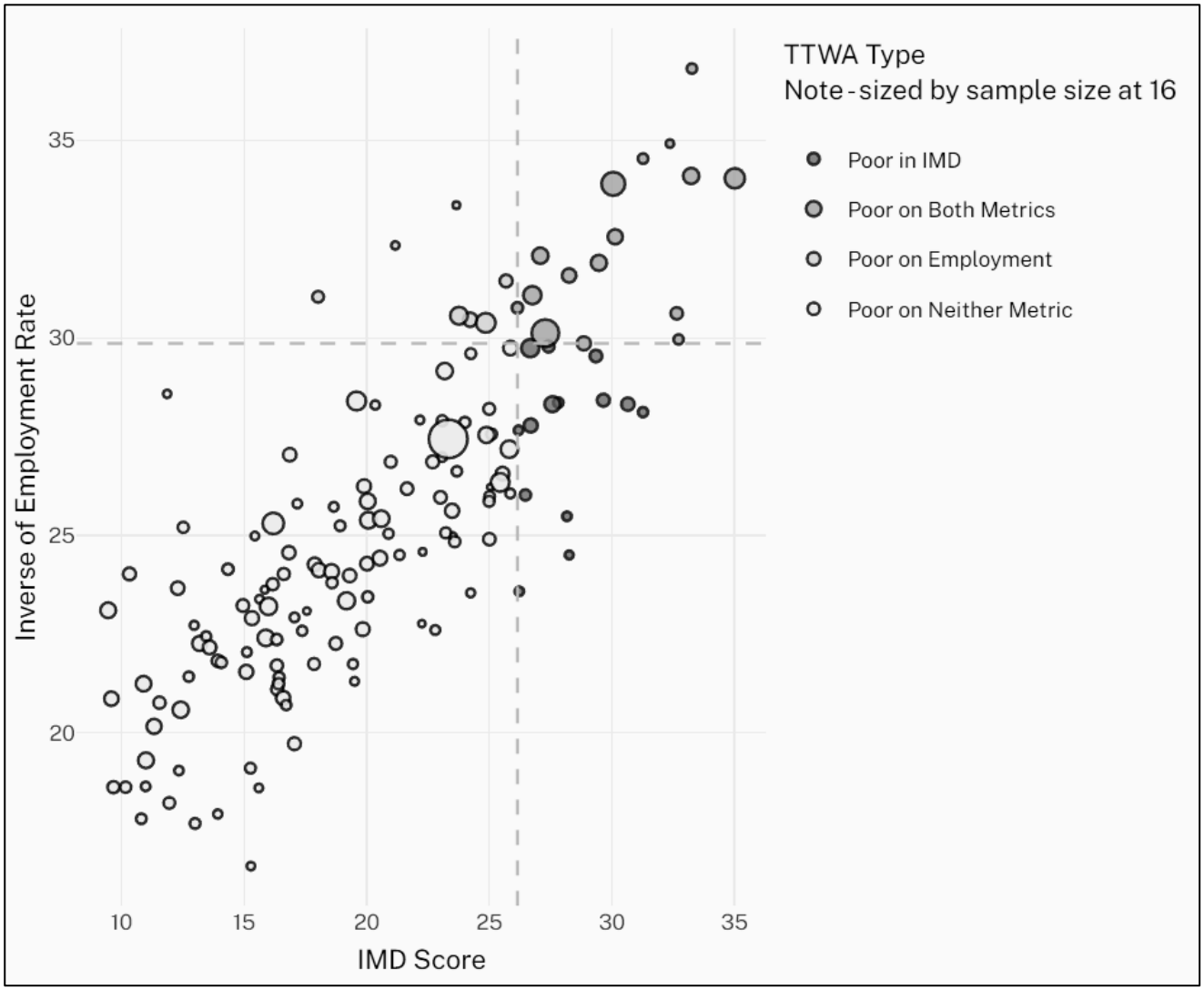
[5] https://www.gov.uk/government/news/government-unveils-levelling-up-plan-that-will-transform-uk

[6] https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015 (note, IMD 2015 are used to align with the time period in which individuals in our dataset were in the labour market)

[7] Based on Annual Population Survey, from September 2012 – September 2017.

Figure 1 shows that, if we aggregate both measures to TTWA level, the two indices are highly correlated. Because of the breadth of the domains covered by the IMD, and the granularity of this data, it is our preferred definition of area-level 'poor performance' for this paper. Results based on using the employment rate as the definition of poor performance are presented in Appendix B.

Figure 1: TTWAs by employment rate and IMD score



Source: LEO data, NOMIS, Indices of Multiple Deprivation.

Notes: Constructed using IMD 2015 data aggregated from LSOA level, and the five-year average of employment rate 2012-2017 at TTWA level from NOMIS. The size of each point is based on each TTWA's sample size in our LEO cohorts at age 16.

The distribution of IMD scores at MSOA level across the country is shown in **Figure 2**. A map showing the distribution of employment rates across TTWAs – which is very similar to the distribution of IMD scores at TTWA level - is shown in **Figure A-1** of Appendix A.

Figure 2: Map of mean IMD score by MSOA

Source: Indices of Multiple Deprivation.

Notes: Includes 6,791 MSOAs which fall within 149 English TTWAs. Darker areas are more deprived.

We define 'poor performance' based on this IMD measure by splitting MSOAs into population weighted quartiles based on their score on the index. The most deprived areas are in quartile 4, and the least deprived in quartile 1. Areas in the most deprived quartile are defined as 'poorly performing'.

2.3 Education levels

The analysis considers the highest level of qualification of any type achieved for each individual by the 2017/18 tax year, derived from KS4/KS5, FE, and HE data. **Table 2** shows the distribution of individuals in the sample across areas on this metric. For

context, Level 2 is equivalent to 5 GCSEs at A*-C, Level 3 is equivalent to 2 A-levels at grades A-E, and Level 6 and above are categorised as graduates. This demonstrates the point made in the introduction: that poorer-performing areas tend to have less educated populations, on average, than other areas. This is particularly clear at graduate level, with a 5-percentage point difference between the proportion of individuals with at least a degree in the poorest performing quartile (27%) compared to other areas (32%).

Table 2: Share of sample by highest qualification level at 27, by area type

Highest Qualification Level	All Areas	Poorest Performing Quartile	Other Areas
1	17.52%	18.82%	17.04%
2	23.46%	25.74%	22.62%
3	24.41%	24.69%	24.31%
4	1.66%	1.75%	1.62%
5	2.34%	2.40%	2.32%
6+	30.61%	26.60%	32.09%

Notes: Count includes all individuals in our cleaned LEO dataset, for whom we have location data at age 16, and who do not have a location outside England recorded at any age.

3. The return to qualifications for the individual

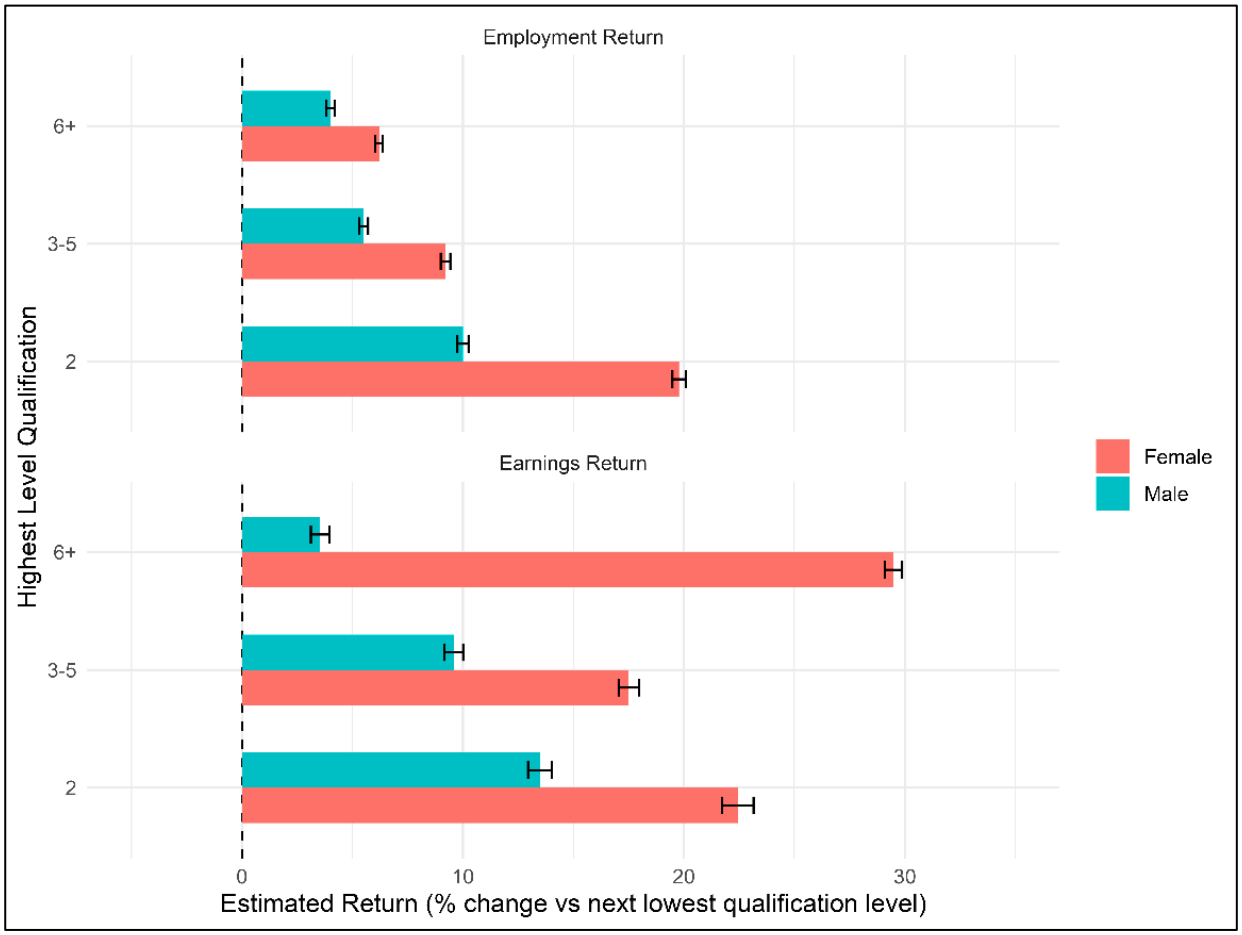
3.1 How much do individuals benefit, on average?

We begin by taking the individual perspective, looking at the benefits to undertaking qualifications for the people receiving those qualifications. We are particularly interested in estimating returns amongst non-graduates (L1-L5). We distinguish individuals whose highest qualification is L1, L2 and L3-5, and compare these to the returns for graduates (L6+). We combine L3-L5 because, as shown in **Table 2**, the proportion of people taking L4 and L5 qualifications is relatively low, and because moving straight from L3 to L6 is a more 'typical' route into HE. We combine L6-L8 as we are studying the returns at age 27, at which point the differential return from different graduate qualification levels is not yet likely to be observable or informative.

We examine the effects of qualifications in terms of both employment and earnings. To do so, we compare employment probabilities and annual earnings between those whose highest qualification is at a given level relative to those at the level immediately below, controlling for a range of other observable differences between individuals with different qualification levels (see Appendix B for the full specification). This means the returns to L2 qualifications are estimated relative to individuals with L1, L3-5 relative to L2, and L6+ relative to L3-5.

Figure 3 shows that higher qualifications are always associated with a higher chance of being in work, and higher earnings, and in each case, these returns are larger for women than men. The payoffs in terms of employment probabilities are particularly large between L1 and L2, with women (men) whose highest qualification is L2 around 19 (10) percentage points more likely to be in work than women (men) whose highest qualification is L1. There are further increases (of 5-10 percentage points each) for those with L3-5 vs. L2, and L6+ vs. L3-5.

Figure 3: Return to qualifications in terms of employment probability and earnings, by sex



Source: LEO data.

Notes: Constructed using data on earnings and employment at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. Excludes individuals with zero earnings. Returns are estimated by comparing earnings/employment status between those with a highest qualification level at level x with those at level x-1 (i.e. individuals with level 3-5 are compared to those with level 2), controlling for (Free School Meals) FSM eligibility at age 16, ethnicity, cohort, and KS2 attainment. Points and confidence intervals show (approximately) the average proportion difference in annual earnings and probability difference of being in employment, by sex, between individuals with the two qualification levels. Qualification levels from 3-5 and 6+ are grouped for sample size reasons.

For women, the earnings returns are substantial at all qualification levels, particularly for graduates, who earn almost 30% more per year than similar women whose highest qualification is L3-L5. Male graduates, by contrast, are not estimated to earn very much more than similar men whose highest qualification is L3-L5. These estimated graduate returns are in line with other estimates using LEO data, e.g. Britton et al. (2021a), who find gross earnings returns to a degree of 27% for state-educated women and 6% for state-educated men at age 30.

3.2 Do the benefits of qualifications vary by area economic performance?

A first step in understanding whether areas are likely to benefit from increasing investment in education or skill development is understanding whether the returns to acquiring education vary by place. **Figure 4** shows the returns to highest educational qualification for employment (in Panel A) and earnings (in Panel B), this time allowing for variation on the basis of where individuals live at age 27 – in particular, whether they live in one of the poorest performing (most deprived) areas (quartile 4) on the basis of IMD score, or the other 80% of areas.

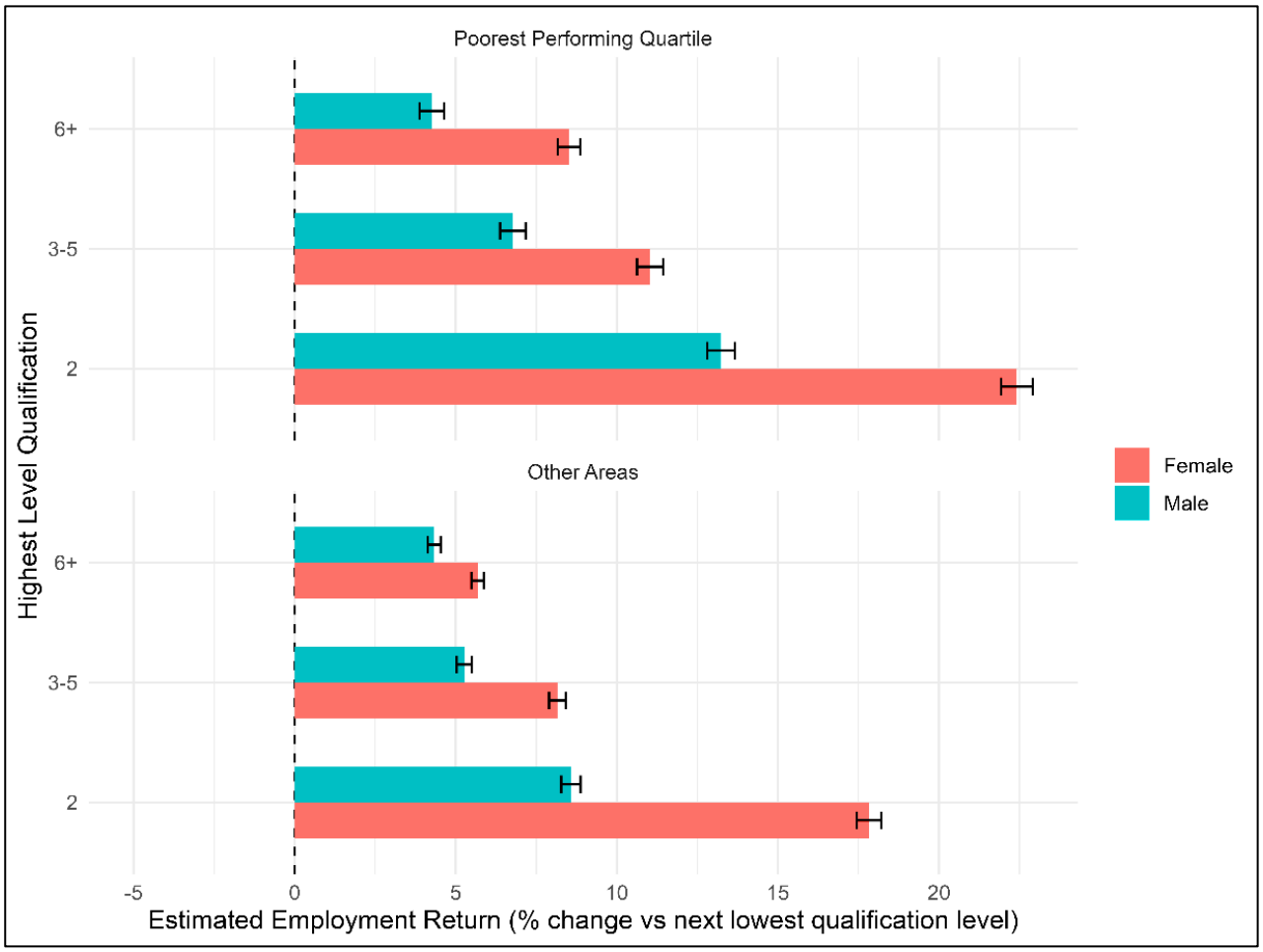
Panel A of **Figure 4** shows that employment returns are somewhat higher for individuals living in the most deprived areas at 27, particularly for non-graduates. For example, women (men) whose highest qualification is L2 have a 25 (13) percentage point higher employment probability than women (men) whose highest qualification is L1 in the most deprived areas, compared to around 17 (8) percentage points in less deprived areas.

Panel B of **Figure 4** shows that there are also some differences between more and less deprived areas in terms of the earnings returns to different qualifications, but these are much smaller and generally not statistically significantly different from each other, i.e. we cannot be sure that any differences in returns across areas are not down to chance.

Figure A-3 in Appendix A shows that these results hold if we additionally account for the area in which individuals grew up; **Figure A-4** shows that the results are very similar if we use childhood rather than adulthood area of residence, and **Figure A-5** shows that it is also true if we were to classify areas based on employment rates rather than IMD scores.

Figure 4: Return at 27 to having a higher level of qualification, by sex and characteristics of current area of residence

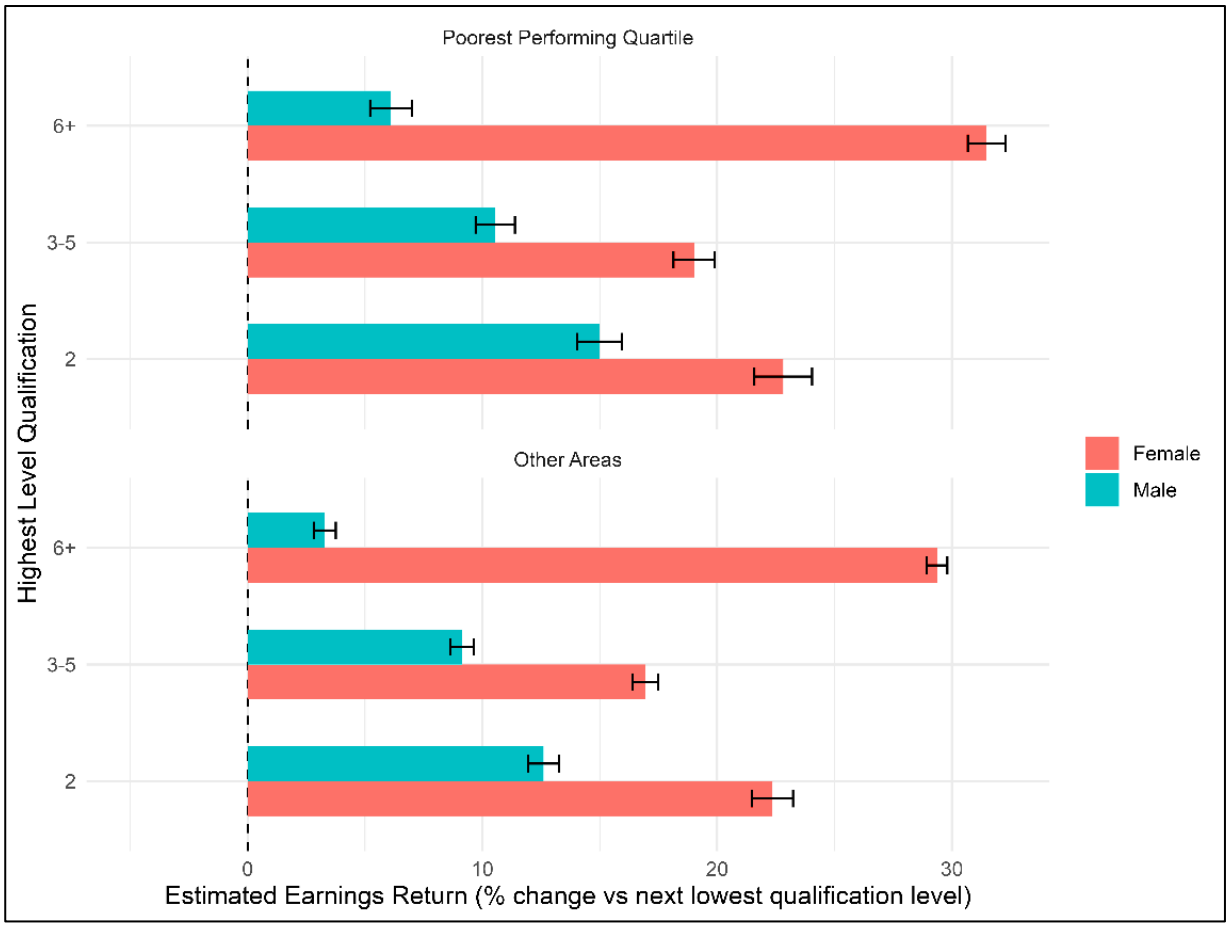
A) Employment Return



Source: LEO data.

Notes: Constructed using data on employment at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. Excludes individuals with zero earnings. Returns are estimated by comparing employment status between those with a highest qualification level at level x with those at level x-1 (i.e. individuals with level 3 are compared to those with level 2), controlling for FSM eligibility at age 16, ethnicity, cohort, and KS2 attainment. Points and confidence intervals show the average probability difference of being in employment, by sex, between individuals with the two qualification levels. The 'poorest performing quartile' results are returns for individuals living in an MSOA in the most deprived IMD quartile at 27, the 'other areas' results are returns for all other individuals.

B) Earnings Return



Source: LEO data.

Notes: Constructed using data on earnings at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. Excludes individuals with zero earnings. Returns are estimated by comparing incomes between those with a highest qualification level at level x with those at level x-1 (i.e. individuals with level 3 are compared to those with level 2), controlling for FSM eligibility at age 16, ethnicity, cohort, and KS2 attainment. Points and confidence intervals show (approximately) the average proportion difference in earnings, by sex, between individuals with the two qualification levels. The 'poorest performing quartile' results are returns for individuals living in an MSOA in the most deprived IMD quartile at 27, the 'other areas' results are returns for all other individuals.

While this analysis has shown that the differences in earnings between qualification levels are roughly similar in different areas, **Figure A-2** in Appendix A shows that the levels of earnings for a given qualification level are slightly lower in more deprived areas than in other areas, particularly for men. In other words, an individual with a particular qualification (e.g. a degree) could potentially earn more – even after adjusting for differences in the cost of living – by moving from a more deprived area to a less deprived area.

These figures suggest that, if all individuals were to stay in the area in which they grew up, those from poorer-performing areas would earn at least as strong returns to the acquisition of qualifications than those from richer areas. At face value, this might

suggest that investments in education or skill development in poorer performing areas would significantly benefit those areas. What we do not know from this analysis alone, however, is whether individuals do indeed stay in the areas in which they grow up, or whether they move to other areas in order to attain higher returns than those they could achieve by staying locally. We turn to this question in the next section.

4. How much do qualifications benefit the areas in which they were obtained?

4.1 How much do people move?

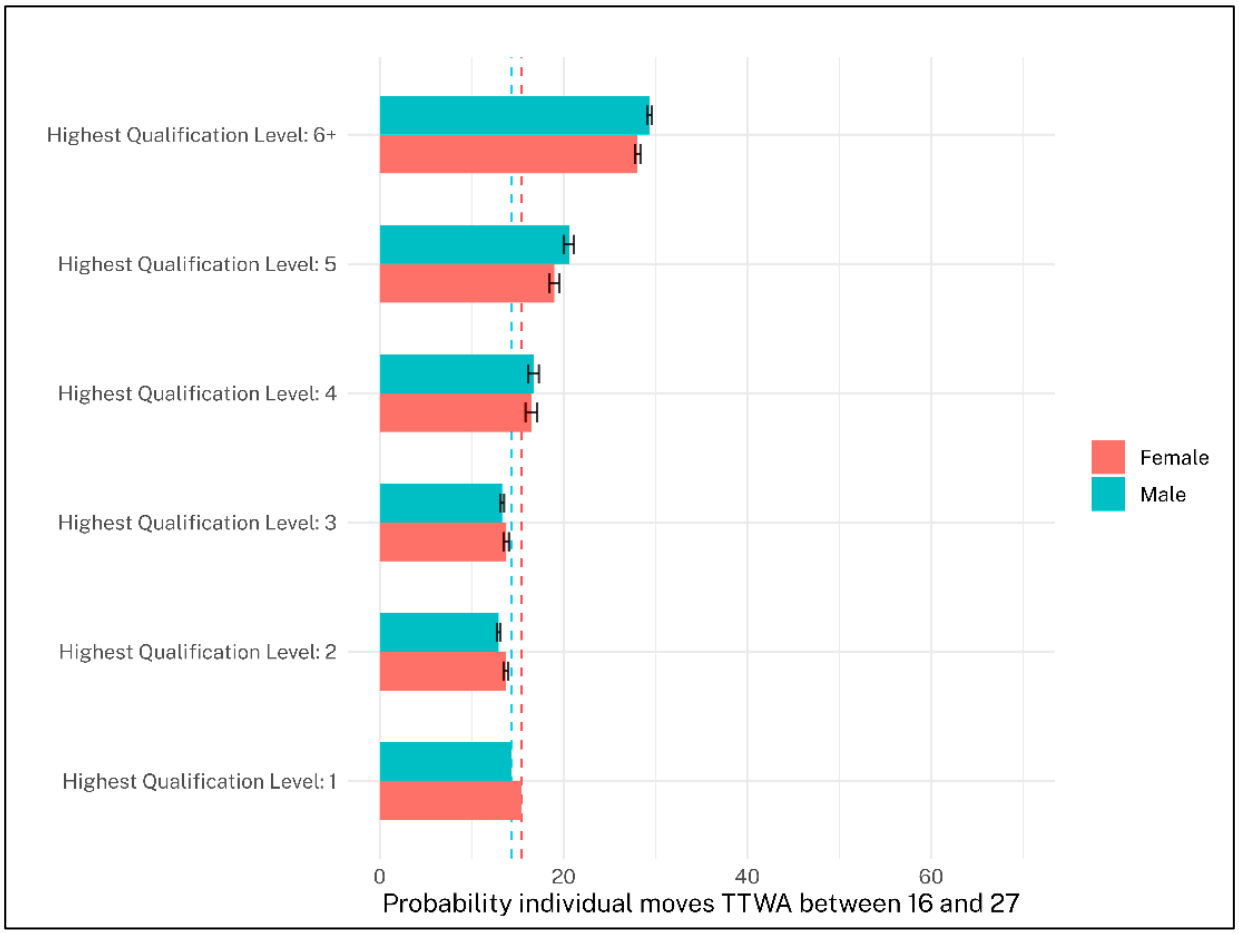
Work by Britton et al. (2021b) shows that graduates are highly mobile across TTWAs and are likely to move to large cities where the returns to degrees are highest, while non-graduates are relatively less mobile. We are interested in exploring whether the acquisition of additional qualifications up to L5 (i.e. amongst the non-graduate group) is associated with a greater propensity to move for the individuals who acquire these qualifications, and additionally to understand how much more mobile individuals appear if we focus on smaller geographic areas (MSOAs rather than TTWAs).

Figure 5 shows the mobility of individuals with different qualification levels (L1-L6+), with Panel A showing mobility across TTWAs and Panel B showing mobility across MSOAs. In both cases, the bars indicate the average probability of moving between the relevant areas between age 16 and age 27 for individuals of a given qualification level *relative* to the average probability of moving for individuals at Level 1, controlling for a set of background characteristics (for full specification see Appendix B). The raw differences are very similar, and can be found in **Figure A-6** in Appendix A. The dashed lines act as references to the proportion of men and women at Level 1 who move: other bars should be interpreted relative to these reference lines.

Panel A reinforces the picture of graduates (those with L6 qualifications and above) as highly mobile across different local labour markets and non-graduates (L1-L5) as relatively less mobile. There is relatively little difference in propensity to move by education level amongst non-graduates, although those whose highest qualification is at L5 (e.g. a foundation degree) are around 5 percentage points more likely to have moved by age 27 than similar individuals whose highest qualification is at Level 1 (not achieving 5 A* to C grades at GCSE). Overall, only around one in six people whose highest qualification is below Level 6 has moved TTWA by the age of 27, compared with around one in three with a Level 6+ qualification.

Figure 5: Probability of moving by age 27 controlling for background characteristics, by level of qualification

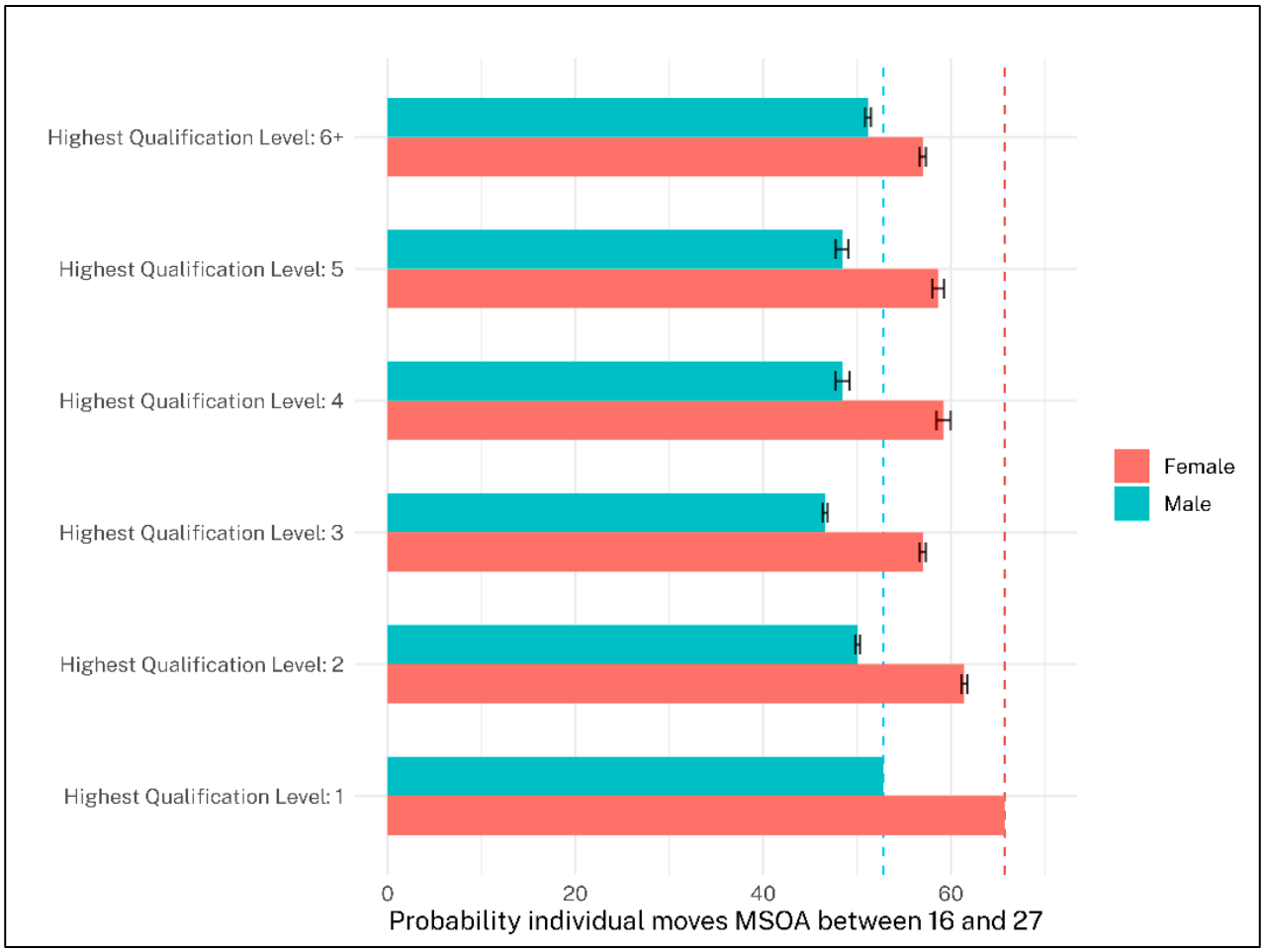
A) Probability of moving across TTWA



Source: LEO data.

Notes: Constructed using data on location at age 16 and 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different TTWA at 27 to their TTWA at 16 Qualification level 1 is the reference group, as indicated by the dashed lines which show the proportion of men and women at qualification level 1 who moved by 27. Differences are estimated by comparing the probability of moving at each other qualification level and the probability of moving at level 1 controlling for cohort, FSM eligibility at age 16, ethnicity, and KS2 attainment. All other estimates are relative to this reference group, calculated by adding the estimated differences between groups to the descriptive proportion of individuals moving at level 1. Estimates are only meaningful in relation to the level 1 reference group (the dashed lines).

B) Probability of moving across MSOA



Source: LEO data.

Notes: Constructed using data on location at age 16 and 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different MSOA at 27 to their MSOA at 16. Qualification level 1 is the reference group, as indicated by the dashed lines which show the proportion of men and women at qualification level 1 who moved by 27. Differences are estimated by comparing the probability of moving at each other qualification level and the probability of moving at level 1 controlling for cohort, FSM eligibility at age 16, ethnicity, and KS2 attainment. All other estimates are relative to this reference group, calculated by adding the estimated differences between groups to the descriptive proportion of individuals moving at level 1. Estimates are only meaningful in relation to the level 1 reference group (the dashed lines).

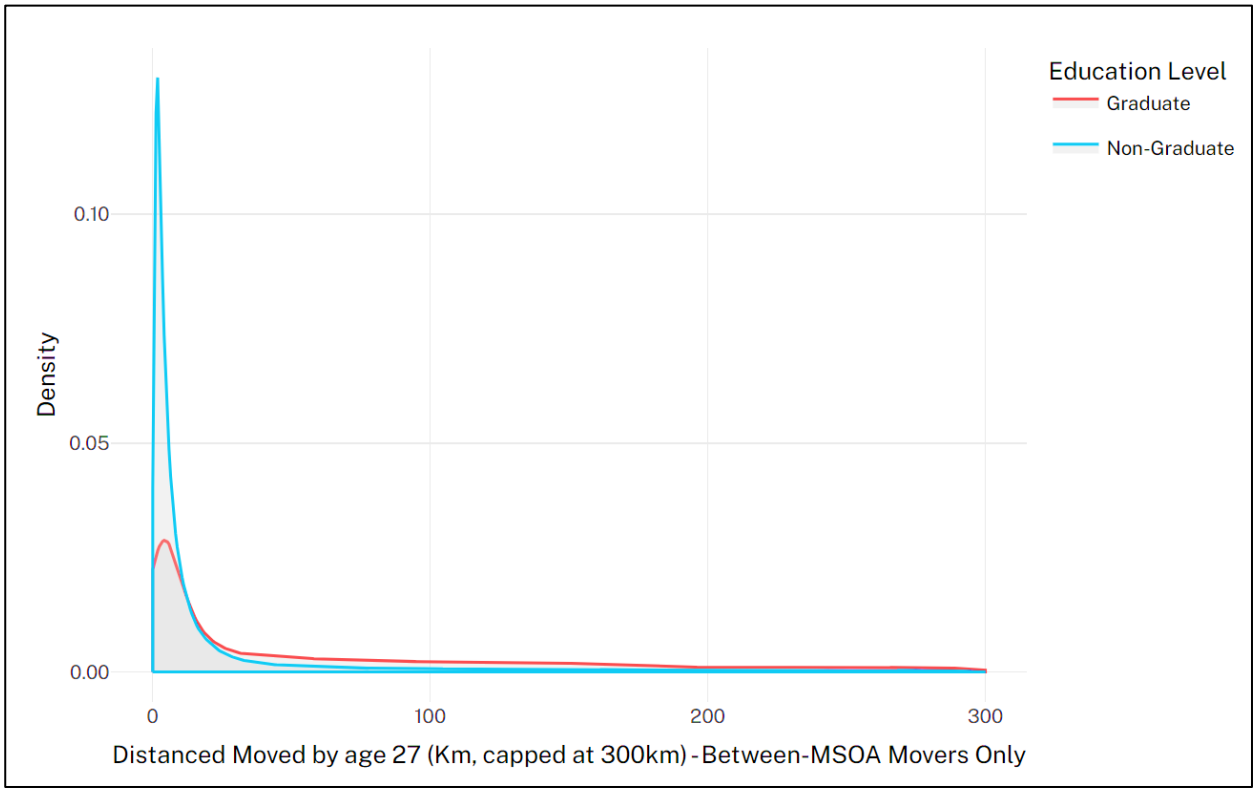
Panel B reproduces the same figure looking instead at more local moves – the likelihood of moving MSOA (rather than TTWA) between age 16 and age 27. Three things stand out: first, the probability of moving across MSOAs is much higher than the probability of moving across TTWAs, particularly for non-graduates, who are more than three times as likely to move MSOAs as TTWAs, with graduates almost twice as likely to move MSOAs as TTWAs. Second, in light of this pattern, the average propensity to move across MSOAs is much more similar for graduates and non-graduates than was the case for moves across TTWAs – and if anything goes in the opposite direction, with those whose

highest qualification is at L1 the most likely to move across MSOAs. Third, at any given education level, women are significantly more likely to move than men.

The distinction between the patterns in Panel A and Panel B suggests that, when graduates move, they are more likely to move longer distances, across TTWAs, potentially for job-related reasons. For non-graduates, however, the moves they are making are relatively shorter distances within TTWAs and are perhaps less likely to be for job-related reasons.

This is borne out by **Figure 6**, which shows the distribution of distance moved for graduates and non-graduates, amongst those who have moved MSOA between age 16 and 27. It shows that the vast majority of moves are very short distances – less than 10km (the median is around 7km) – and that this is especially true for non-graduates (in blue) for whom the median is only around 5km. (For graduates (in red) the median distance moved is around 25km.) Beyond about 20km, graduates are more likely than non-graduates to move across the whole distribution of distances, although the number who are doing so is still dwarfed by the number making moves across MSOAs close to home.

Figure 6: Distribution of distance moved by age 27 for graduates and non-graduates who leave their childhood MSOA



Source: LEO data.

Notes: Constructed using location data at 16 and 27, at the LSOA level. As individual address data is not available, distance moved is calculated as the distance between the centroids of an individual's LSOA of residence at 16 and their LSOA of residence at 27. This chart only includes individuals who have moved from their childhood MSOA by 27.

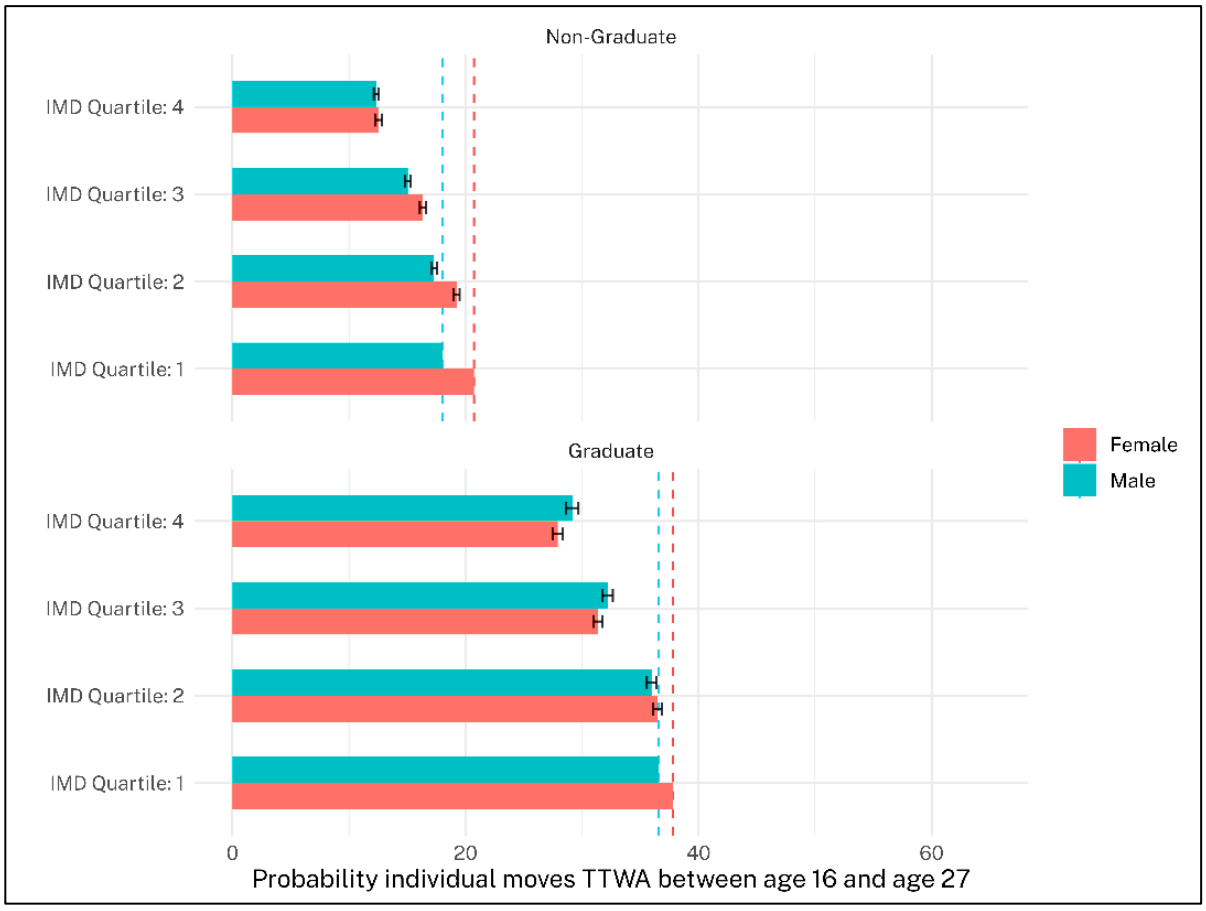
These results suggest that while most individuals move MSOAs between age 16 and age 27, the vast majority amongst non-graduates move relatively small distances, with only around a fifth of movers changing TTWAs. For graduates the split is more even – a little over half of movers change TTWAs. This already suggests that investments in qualifications or skills below graduate level are highly likely to benefit the areas in which the investments are made. To understand whether this is also true for the most deprived areas, where 'levelling up' resources may be concentrated, we consider whether the propensity to move differs depending on the area in which an individual grew up.

Figure 7 shows the propensity to move for non-graduates (upper chart) and graduates (lower chart) depending on the deprivation level of the area in which individuals lived at age 16, after accounting for a set of background characteristics. Panel A considers moves between TTWAs and Panel B considers moves between MSOAs. **Figure A-7** in Appendix A shows that the raw figures are very similar, and also that there is little difference in the propensity to move by qualification level within the non-graduate group.

Panel B shows that the propensity to move across MSOAs is essentially independent of the deprivation level of the area in which individuals grew up, while Panel A suggests a stronger relationship between area characteristics and the propensity to move. In contrast to what we might have expected based on the differences in earnings levels (expected returns) across areas, those from more deprived areas are significantly less likely to move TTWA than similar individuals from less deprived areas, with the relationships slightly stronger for women than men. For example, non-graduate women in the most deprived areas (quartile 4) are around 8 percentage points (40%) less likely to have moved TTWA by age 27 than similar women in the least deprived areas (quartile 1). The difference for female graduates is around 10 percentage points (25%).

Figure 7: Probability of moving by age 27, by graduate status and IMD quartile of childhood MSOA (Quartile 4 is the most deprived)

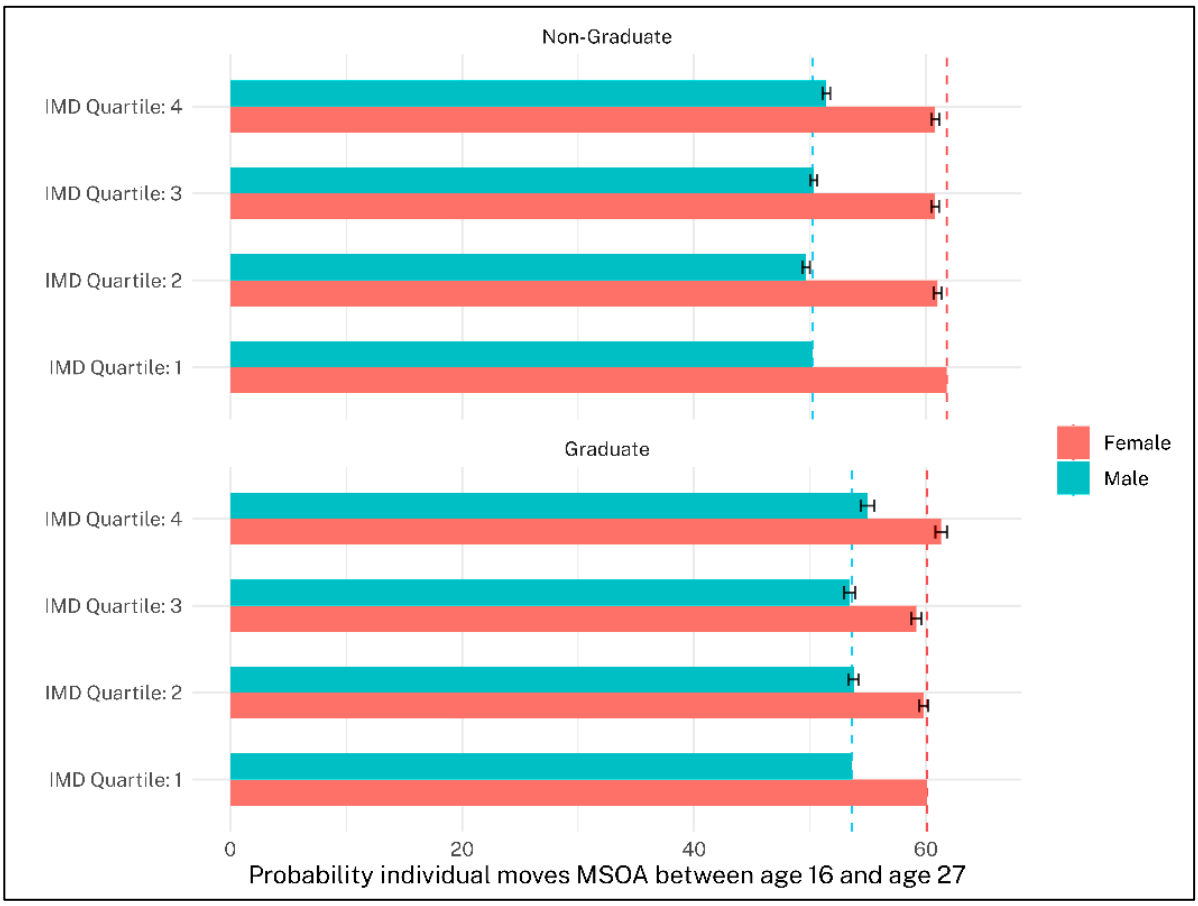
A) Probability of moving across TTWA



Source: LEO data.

Notes: Constructed using data on location at age 16 and 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different MSOA at 27 to their MSOA at 16. The model is estimated separately for non-graduates and graduates. IMD quartile 1 is the reference group in both models, as indicated by the dashed lines which show the proportion of men and women from areas in IMD quartile 1 who moved by 27. Differences are estimated by comparing the probability of moving from each other IMD quartile and the probability of moving from quartile 1 controlling for cohort, FSM eligibility at age 16, ethnicity, and KS2 attainment. All other estimates are relative to this reference group, calculated by adding the estimated differences between groups to the descriptive proportion of individuals moving from quartile 1. Estimates are only meaningful in relation to their reference group (the dashed lines).

B) Probability of moving across MSOA



Source: LEO data.

Notes: Constructed using data on location at age 16 and 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different TTWA at 27 to their TTWA at 16. The model is estimated separately for non-graduates and graduates. IMD quartile 1 is the reference group in both models, as indicated by the dashed lines which show the proportion of men and women from areas in IMD quartile 1 who moved by 27. Differences are estimated by comparing the probability of moving from each other IMD quartile and the probability of moving from quartile 1 controlling for cohort, FSM eligibility at age 16, ethnicity, and KS2 attainment. All other estimates are relative to this reference group, calculated by adding the estimated differences between groups to the descriptive proportion of individuals moving from quartile 1. Estimates are only meaningful in relation to their reference group (the dashed lines).

4.2 Why don't people move more?

We have seen that getting additional levels of further education (particularly qualifications up to and including L3) does not appreciably increase the probability of moving across labour markets (TTWAs). A natural question then, is why don't non-graduates move more?

We noted in the introduction that a caveat of the LEO data is that moves may be under-reported (if individuals do not update their records to show that they have changed addresses). For this to help explain the relatively lower propensity to move of non-

graduates compared to graduates, there would need to be substantially lower reporting of moves by non-graduates than graduates. We do not have any evidence to suggest that this is the case.

It is therefore instructive to consider the reasons why individuals might move in the first place. In terms of the labour market, there may be a 'moving premium': an increased likelihood of employment, or higher earnings for those who are employed, for individuals who move areas. Where there are such benefits to moving location, after accounting for differences in the cost of living, we might expect to see individuals moving across areas.

Figure 8 compares, separately for males and females, the employment probabilities (top bars), earnings (middle bars) and cost-of-living adjusted earnings (bottom bars) of those who move TTWA between age 16 and age 27 with similar individuals who stay in the same TTWA for non-graduates (top chart) and graduates (bottom chart), after accounting for a small number of other differences between individuals who move vs. stay. The raw differences are very similar, and can be found in **Figure A-8** of Appendix A.

In terms of the likelihood of being in work, graduates who move are, on average, more likely to be employed than those who stay, while amongst non-graduates, only male movers see a positive return to moving in terms of their likelihood of being in work; female non-graduates who move are actually slightly less likely to be in work, on average, than those who stay. It is worth noting, however, that these differences are much smaller than the increases in the likelihood of being in work associated with the acquisition of higher qualifications seen in **Figure 5**, with at most a 4 percentage point difference in the likelihood of being in work (after accounting for other observable characteristics) between movers and non-movers.

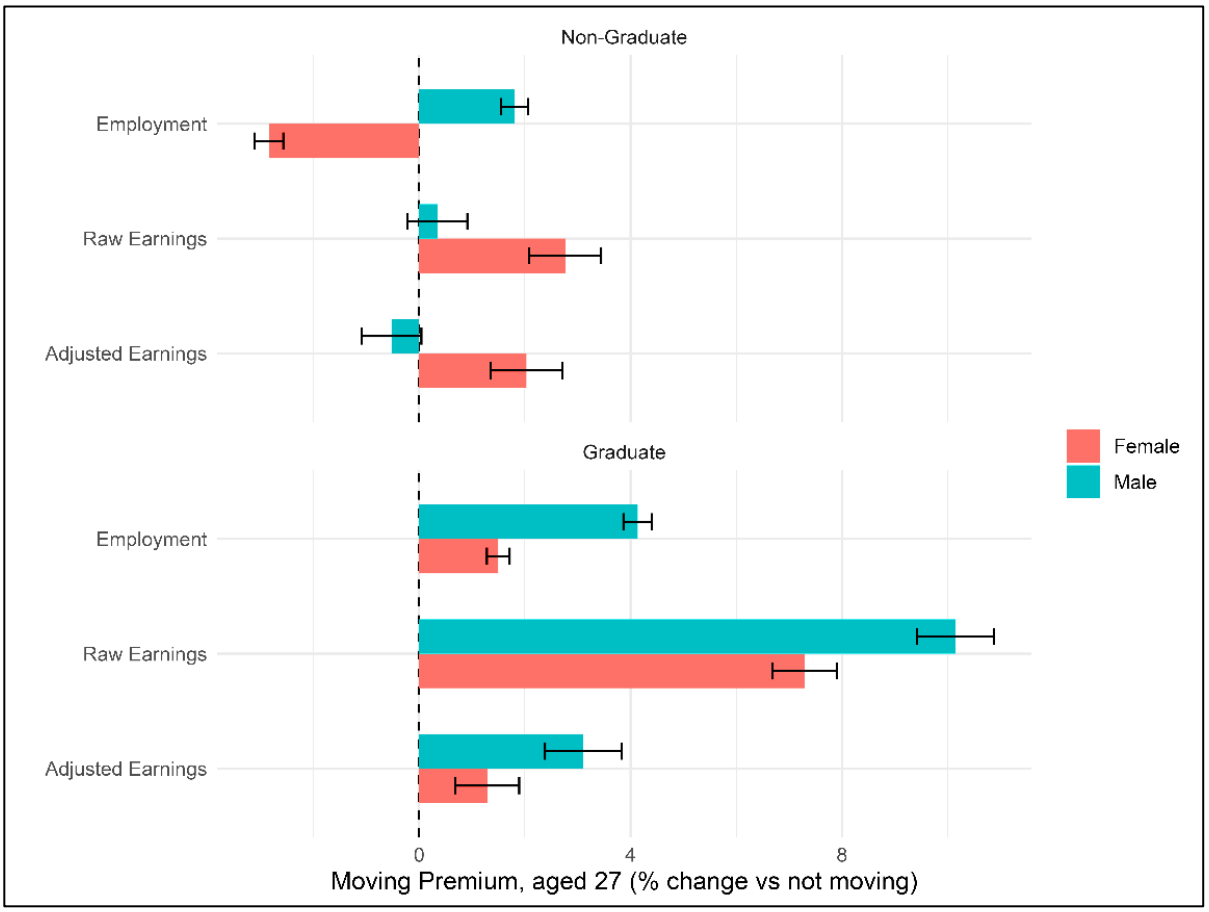
In line with the findings of Britton et al. (2021b), we find a significant nominal earnings premium for graduates who move TTWA by age 27, with male (female) graduates who move earning around 10% (7%) more, on average, than similar individuals who stay. For non-graduates, the estimates are much smaller: only around 2% for women, and not significantly different from zero for men, suggesting that the benefits of moving arise in terms of employment probabilities for men and earnings for women. The returns to moving in terms of employment and nominal earnings do not vary substantially with the deprivation level of the areas in which individuals grew up (see **Figure A-9** in Appendix A).[8]

A substantial proportion of the moving premium for graduates is eroded by differences in the cost of living across areas, however: once account is taken of the fact that graduates tend to move to higher cost areas (including London), the differences between graduate

[8] The cost-of-living adjusted figures show a different pattern, reflecting the types of areas to which individuals growing up in different areas move by age 27, as discussed in the next section.

movers and non-movers look much smaller and more in line with the moving premium for female non-graduates (around 2-3%), which is relatively unaffected by cost-of-living adjustments. These findings are broadly similar if we use our alternative cost-of-living adjustment which does not rely on differences in regional prices – see **Figure A-10** in Appendix A.

Figure 8: Moving premium for employment and earnings by graduate status and sex, at age 27



Source: LEO data.

Notes: Constructed using data on employment, annualised earnings, and cost-of-living adjusted earnings at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different TTWA at 27 to their TTWA at 16. The model is estimated separately for non-graduates and graduates. The premium is estimated by comparing the earnings of movers and stayers, controlling for cohort, FSM eligibility at 16, TTWA at 16, and ethnicity. The estimates on employment can be interpreted as the estimated difference in probability of being in employment at 27 between movers and non-movers, the estimates on earnings as approximately the proportion difference in earnings between movers and non-movers.

Given these findings on the financial returns to moving, it is perhaps not surprising that non-graduate mobility is relatively low, while it is somewhat higher for graduates. There are of course many reasons for moving that go beyond the labour market. Some people may move to a city to benefit from amenities like theatres and restaurants, while others may move to the countryside to have access to more space. These sorts of explanations

can explain why we still see some non-graduates move: they may be moving for reasons other than the financial benefits. They can also explain why most people stay (or only move short distances): friends and family may be located close to where an individual grew up, and these can provide both an important support network and informal economic benefits such as grandparental childcare.

There are two major caveats in interpreting these results. First, observing that cost-of-living adjusted returns are close to zero does not necessarily mean there is no contemporaneous benefit to moving.[9] If, for example, access to theatres and restaurants are an amenity that people value, it will be more expensive to live closer to these amenities. Treating the difference in outlay on housing for a given dwelling size as a pure cost, without accounting for the benefit that individuals receive from this, is therefore an overestimate of the impact on cost of living. The unadjusted and adjusted impacts can therefore be thought of as upper and lower bounds (respectively) on the contemporaneous financial benefits of moving.

Second, and more importantly, they do not rule out longer-term benefits to moving. Given our data we focus here on individuals aged 27, which is the oldest age at which we can observe all of our cohorts.

Figure A-11 in Appendix A shows analogous results to **Figure 8**, but at age 24. The moving premia in terms of both employment and cost-of-living adjusted earnings are actually negative for both non-graduates and graduates at this age. If moving has little immediate benefit but can put individuals on the path for faster earnings growth in future years, then our results will be an underestimate of the financial benefit of moving over the course of a career. However, they still highlight an important constraint. If individuals do not see any short-term benefit to moving, and indeed in the earlier years (age 24) see distinct costs to doing so, then those who are less well-informed about the future benefits, or who cannot finance a move to a more expensive area, will end up not moving despite the long-term returns. There would then be scope for policy to try to tackle this, by improving labour market information (see also Skills and Productivity Board, 2022) and reducing differences in housing costs.

[9] Nor does it mean that individuals who moved in the past did not benefit from doing so, as they may have taken advantage of arbitrage opportunities which have since been exhausted.

4.3 When people move, where do they move to?

We saw above that non-graduates are more likely to move locally – across MSOAs within the same TTWA – than they are to move across TTWAs, whilst graduates are slightly more likely to make cross-TTWA moves than more local ones. Here we focus on where individuals are moving to: do they tend to move to other areas similar to where they lived at age 16, or do they move from poorer-performing areas to better-performing areas (or vice versa)?

Figure 9 is a Sankey chart showing the flow of individuals across different IMD quartiles between age 16 (on the left) and age 27 (on the right), for between-TTWA movers only. Panel A shows results for non-graduates, and Panel B results for graduates. For non-graduates, movers are fairly evenly split across deprivation quartiles at age 16, while for graduates the starting point is very different: the population of graduates who move TTWA by age 27 is highly skewed towards less deprived areas (quartile 1) at age 16, likely reflecting both that individuals who are from less deprived areas are more likely to become graduates, and that – as we saw in **Figure 7** – graduates from less deprived areas are more likely to move.

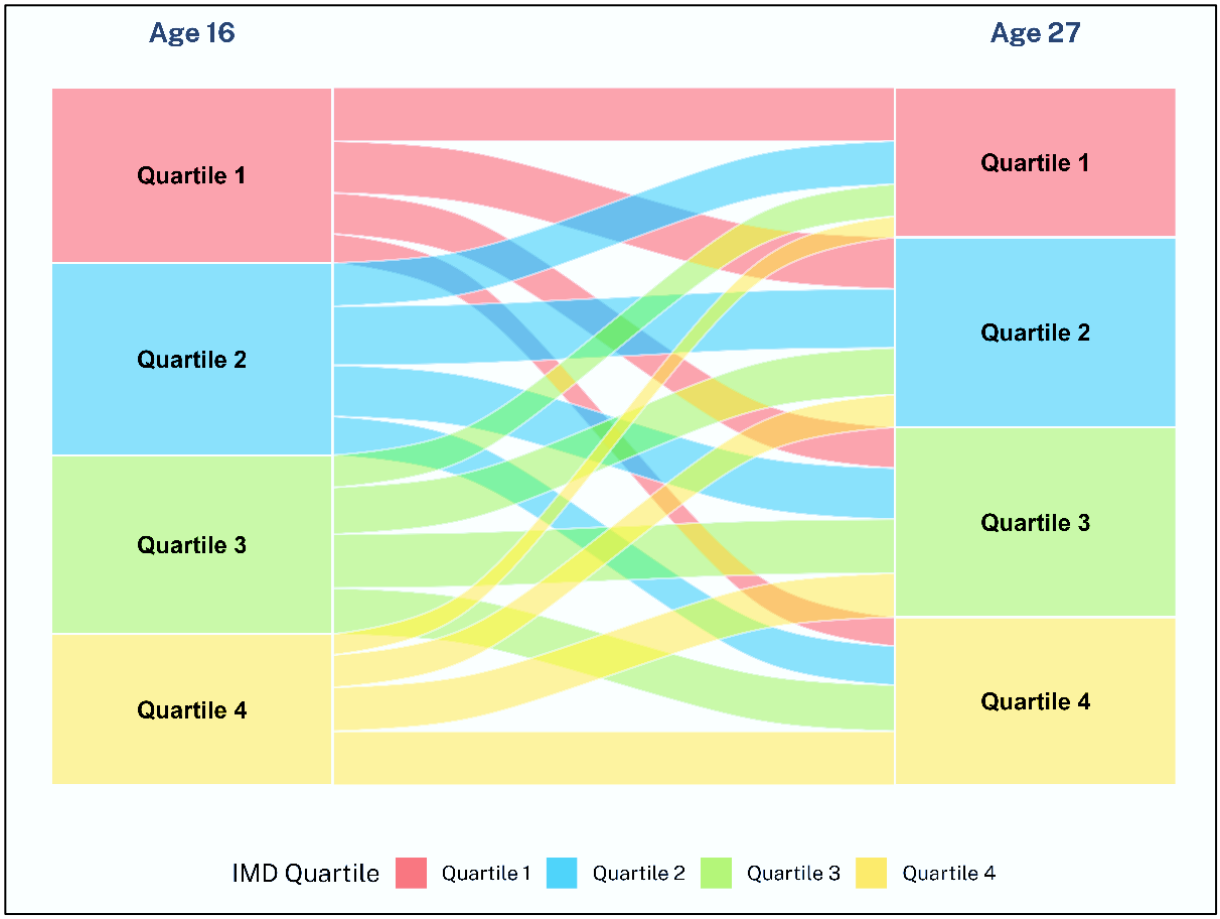
For both graduates and non-graduates, the majority of between-TTWA movers end up in a different deprivation quartile at 27 to their quartile at 16, with the proportion that change quartile similar regardless of the area in which individuals grew up. Moves are also more likely to be made between adjacent quartiles than more disparate quartiles. In other words, individuals are more likely to move to areas that share some common features with where they grew up, rather than to very different areas: the least commonly travelled pairwise paths between quartiles are between the least and most deprived quartiles (and vice versa).

It is also the case that for graduates (and to a much lesser extent non-graduates as well) the least deprived areas are net losers of movers, and the most deprived areas are net gainers. At first glance this may seem surprising, given what we know about the types of areas to which graduates move (e.g. from Britton et al., 2021b). However, it is likely due to a combination of the initial distribution of graduates and within-TTWA sorting. At age 16, future graduates are disproportionately likely to be living with their parents in relatively affluent areas in which few 27-year-olds may be able to afford to live independently. Further, graduates differentially move to more expensive TTWAs, meaning that *within* these TTWAs they may sort into the cheaper (and more deprived) areas.

Figure A-12 in Appendix A shows that these patterns are broadly similar if we categorise areas on the basis of employment rates rather than IMD scores.

Figure 9: Sankey chart showing changes in IMD quartile of area of residence between age 16 and age 27, for between-TTWA movers only (Quartile 4 = most deprived)

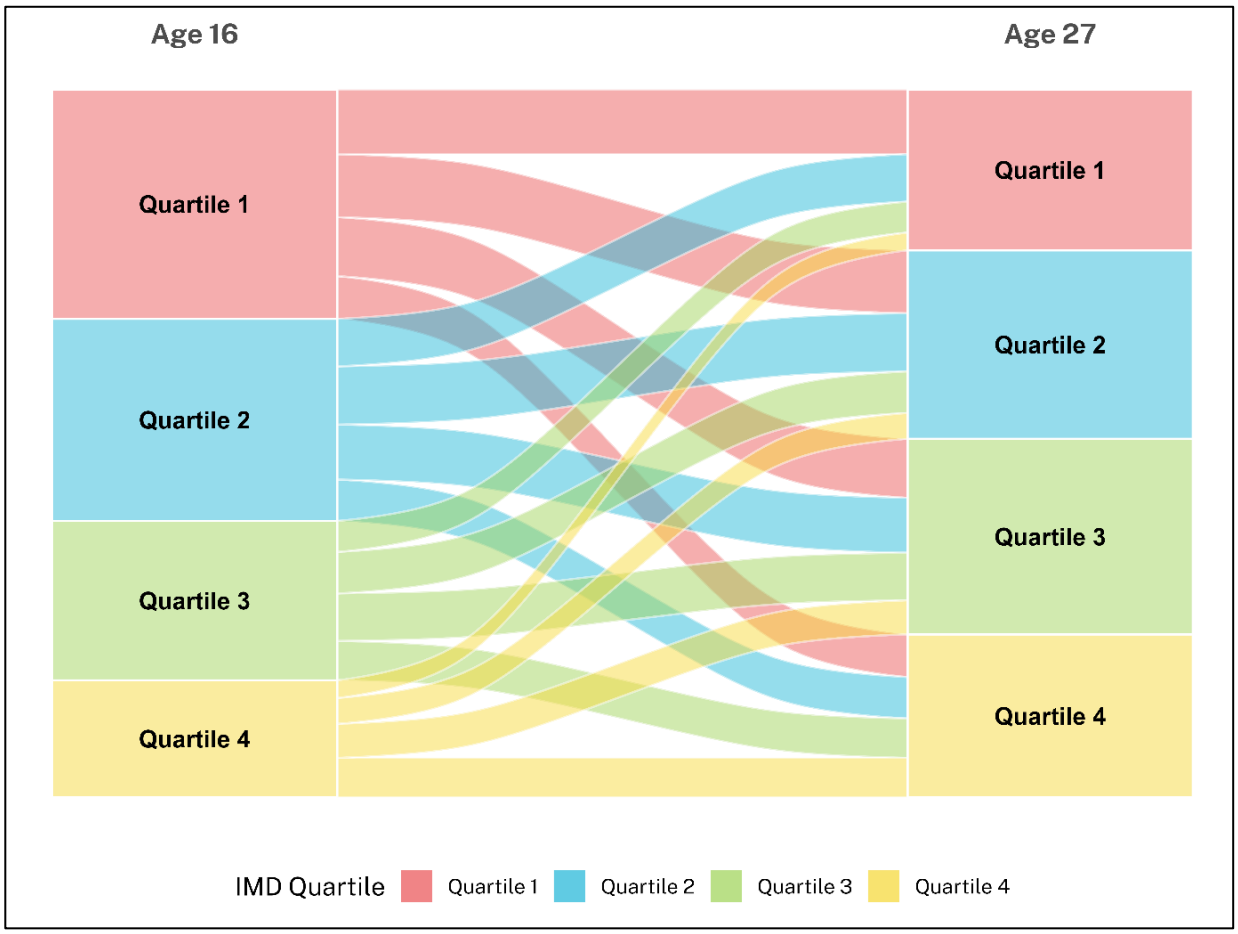
A) Non-graduates



Source: LEO data.

Notes: Constructed using location data at ages 16 and 27, at the MSOA level. Quartiles defined at the MSOA level, where quartile 4 is most deprived, such that individuals can change quartile in this chart with a relatively short move between MSOAs. This chart includes individuals who do not move MSOA by 27.

B) Graduates



Source: LEO data.

Notes: Constructed using location data at ages 16 and 27, at the MSOA level. Quartiles defined at the MSOA level, where quartile 4 is most deprived, such that individuals can change quartile in this chart with a relatively short move between MSOAs. This chart includes only individuals who move MSOA by 27.

5. Are qualifications sufficient to equalise earnings across place?

The results so far show that qualifications are important for earnings, and that there is relatively little contemporaneous premium associated with moving location. It is tempting therefore to think that “place” does not matter and providing skills to individuals is all that is needed to improve both the outcomes of individuals and of the areas where they live.

This idea that almost all the difference in outcomes across areas is driven by differences in the characteristics of the people living in those areas is consistent with earlier research in the UK which focused on TTWAs and used a 1% sample of employees (Gibbons et al., 2014). These findings have become well-established in policy circles. By contrast, recent work using the LEO data comparing the earnings of sons growing up in different local authorities in England highlights large differences in outcomes based on where they grew up (Social Mobility Commission, 2020). Similarly, work in the US using administrative data finds hyper-local differences in outcomes across areas for children from similar families, including across census tracts (small geographic areas with around 2,500-8,000 residents – roughly the size of LSOAs or MSOAs in the UK) (Chetty et al., 2014, 2018).

We explore the extent to which these differences in findings are driven by the size of the geographic areas used. If area effects operate at a relatively local level, then aggregating several smaller areas together, such as with TTWAs, creates the possibility that area effects are underestimated, because doing so averages out these differences. In light of the evidence shown earlier in this report that individuals with lower qualification levels are generally less mobile, and that, as has been shown in previous research, in some cases, decision-making horizons are extremely localised (e.g. Green et al., 2005; Green and White, 2008), this may be important in the UK context.

To understand the extent to which ‘productivity’ (as proxied by individual earnings) can be explained by ‘people’ versus ‘place’ characteristics, we replicate and extend the analysis of Gibbons et al. (2014). The main exercise, described in detail in Appendix A.3, is to first estimate how much of the variation in earnings can be explained by a combination of individual and area fixed effects, and then decompose this explained variation into the part that can be attributed to individual versus area effects (people versus place).

Table 3 shows the findings from this exercise. Column 1 presents the results of Gibbons et al. (2014), and Column 2 shows our replication of these results using our data. Here we maintain the same area definition (TTWA, although we do not merge any TTWAs, as Gibbons et al. are forced to do in some cases because of their smaller sample size), so that the main differences are that we have a larger and younger sample (our oldest individuals are aged 30, while theirs covers the full working age population). Column 3 presents our results using MSOA rather than TTWA as the relevant area level.[10]

Row 1 of **Table 3** presents the variation in earnings across areas, showing that it is relatively small and remarkably similar across specifications, despite the different geographic areas and cohorts used. Row 2 shows how much of that area-level variation remains after we take account of the fact that people with different characteristics – that are likely to be rewarded differently in the labour market – live in different areas. Row 4 presents the ratio between these two figures, with Row 3 showing how much of the variation in earnings can be accounted for by individual characteristics.

The results in Rows 2, 3 and 4 all point towards the fact that we are able to explain less of the variation in earnings across areas using individual characteristics than Gibbons et al. (2014), leaving a greater role for area effects to explain variation in individuals' earnings across areas. Of particular note, is the fact that we are able to explain only around 60% of the variance in earnings using individual fixed effects relative to 86% in Gibbons et al. (2014). This is likely driven by the relative youth of our cohorts: individual characteristics will be able to explain less of the difference in earnings when these differences are small, and as **Figure A-13** in Appendix A shows, the variation in wages across TTWAs rises with age, suggesting that our relatively young sample is likely to have less variation available to be explained than the sample in Gibbons et al. (2014). Similarly, the share of the variation across areas explained by area characteristics (Row 4) is around 2-2.5 times higher in our specifications – regardless of the size of area used – than in Gibbons et al. (2014).

[10] Where an MSOA has an average sample size less than 200 individuals per year, we merge it with the geographically contiguous (neighbouring) MSOA with the largest sample size. This process is performed iteratively until all MSOAs have the necessary average sample size. This leaves around 4,500 distinct areas for use in the analysis.

Table 3: Decomposition of variance in earnings

Metric	Gibbons et al (2014)	Our Replication	
		TTWA Level	MSOA Level
Variation attributable to area	6.00%	4.98%	6.30%
Variation attributable to area, after controlling for individual characteristics	0.70%	1.52%	1.64%
Variation attributable to individual characteristics, after controlling for area	86.40%	59.30%	59.50%
Area share of area disparities	12.60%	30.45%	26.01%
Number of individuals	305,752	1,555,267	1,555,267

Source: LEO data, Gibbons et al (2014): Table 2.

Notes: Constructed using earnings and location data from 2011/12 to 2016/17 tax year (i.e. each year of complete location data) for the 2001/02 to 2005/06 GCSE cohorts. Note that due to a limitation in the data available to the authors, the oldest cohort (2001/02) GCSE year only have four years of data included in this model (up to age 30). Our replication adds controls for individual fixed effects and year dummies in line with specification 4 in gibbons et al 2014. MSOA-level results are based on a set of areas produced by combining geographically contiguous MSOAs until a minimum average sample size of 200 per area per year is achieved. This iteration of the model uses 4,398 combined areas with an average sample of 380 individuals per area per year.

The implications of these findings are two-fold. Firstly, Row 1 of **Table 2** highlights that, in absolute terms, very little of the overall variation in earnings is related to place. This suggests that there is much more variation in wages between individuals within the same areas than there is, on average, across areas. Perhaps surprisingly, this seems to be true even when focusing on very small geographic areas. (**Table A-2** in Appendix A shows that this is even more true if we account for differences in the cost of living across areas, which is likely to be correlated with some of these area-level factors.)

Second, despite the importance of individual characteristics such as qualifications in explaining differences in earnings, between a quarter and a third of the difference in earnings across areas cannot be explained by individual effects (Row 4). Instead, these are driven by other place-specific effects. While this approach cannot itself shed light on the source of those effects, it does highlight the potential importance of investments other than in human capital in reducing disparities in economic performance across areas (Levelling Up White Paper, 2022).

6. Conclusions

Expanded access to education and training can only benefit poorer performing areas of the country under two conditions. First, it must contribute to higher productivity of individuals living in those areas. Second, those individuals must largely remain in their existing locations: there must not be a ‘brain drain’ resulting from improved skills.

Taking these conditions in turn, we see first that there are substantial returns to skill upgrading. Controlling for other characteristics, earnings for women (men) with L2 qualifications – five GCSEs at A*-C – are, on average, more than 20% (10%) above those with only L1 qualifications, who make up 17% of adults in our sample. Marginal returns to other qualifications are lower but still substantial. Moreover, the returns are, if anything, at least as high in poorer performing parts of the country as elsewhere.

Turning to the second condition, we see that, among non-graduates, there is relatively little long-distance mobility. While, by the age of 27, more than half of graduates and non-graduates have moved MSOA, just under 20% of non-graduates live in a different travel to work area, compared with almost 40% of graduates. This difference in mobility is consistent with the difference in contemporaneous ‘moving premium’, which is generally lower for non-graduates than graduates. Perhaps surprisingly, mobility across TTWAs is also lower for those coming from more deprived areas than for those coming from better off areas.

From the perspective of ‘levelling up’, these results highlight the importance of investments in skills and suggest that the benefits of such investments in poorer performing places are highly likely to remain within the targeted areas. They also highlight, however, that education and training alone can only go so far. Decomposing differences in earnings into the parts that can be attributed to individual characteristics and the parts that relate to local area characteristics, we see that between a quarter and a third of the difference in earnings across areas is associated with area characteristics beyond those of the individuals living there. Improving productivity in poorer performing areas will therefore require complementary investments alongside investments in skills.

Appendix A: Additional Results

Data and context

Figure A-1: Map of TTWA rankings on employment rate metric

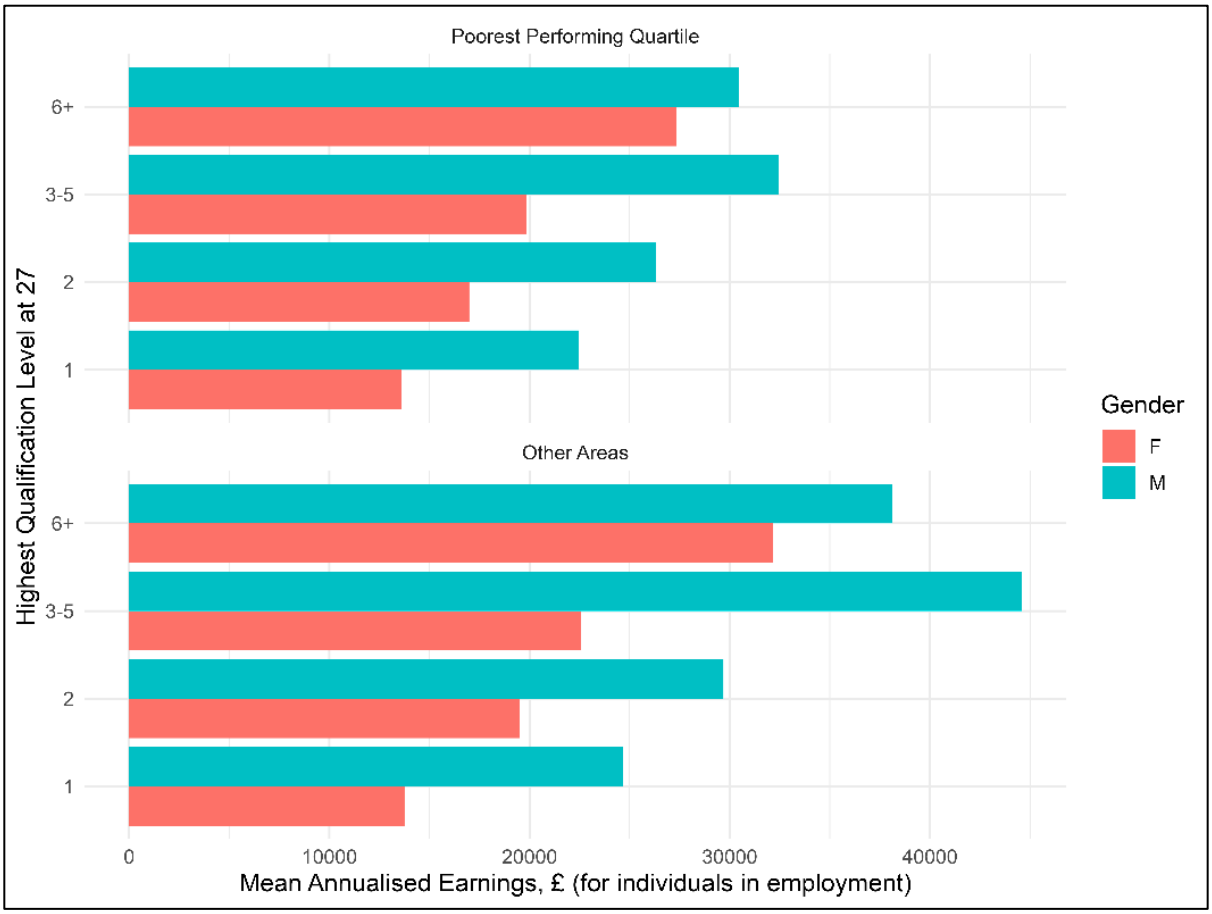
Source: NOMIS.

Notes: Includes 149 English TTWAs. Darker areas are more deprived.

The return to qualifications for the individual

Figure A-2: Descriptive differences in earnings by highest level of qualification at 27, by sex

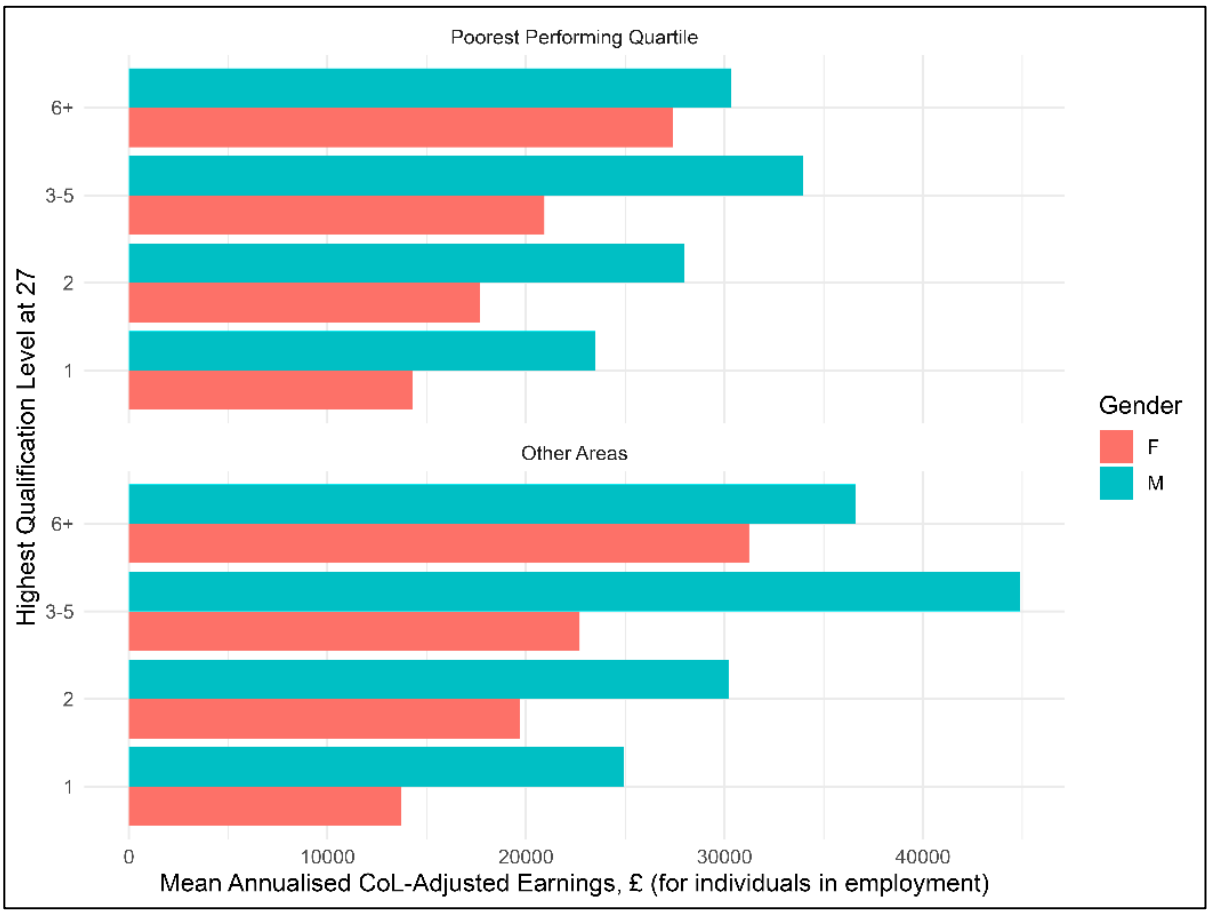
A) Raw Earnings



Source: LEO data.

Notes: Constructed using data on earnings at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. Excludes individuals with zero earnings. The 'poorest performing quartile' results are averages for individuals living in an MSOA in the most deprived IMD quartile at 27, the 'other areas' results are averages for all other individuals.

B) Cost-of-Living adjusted earnings

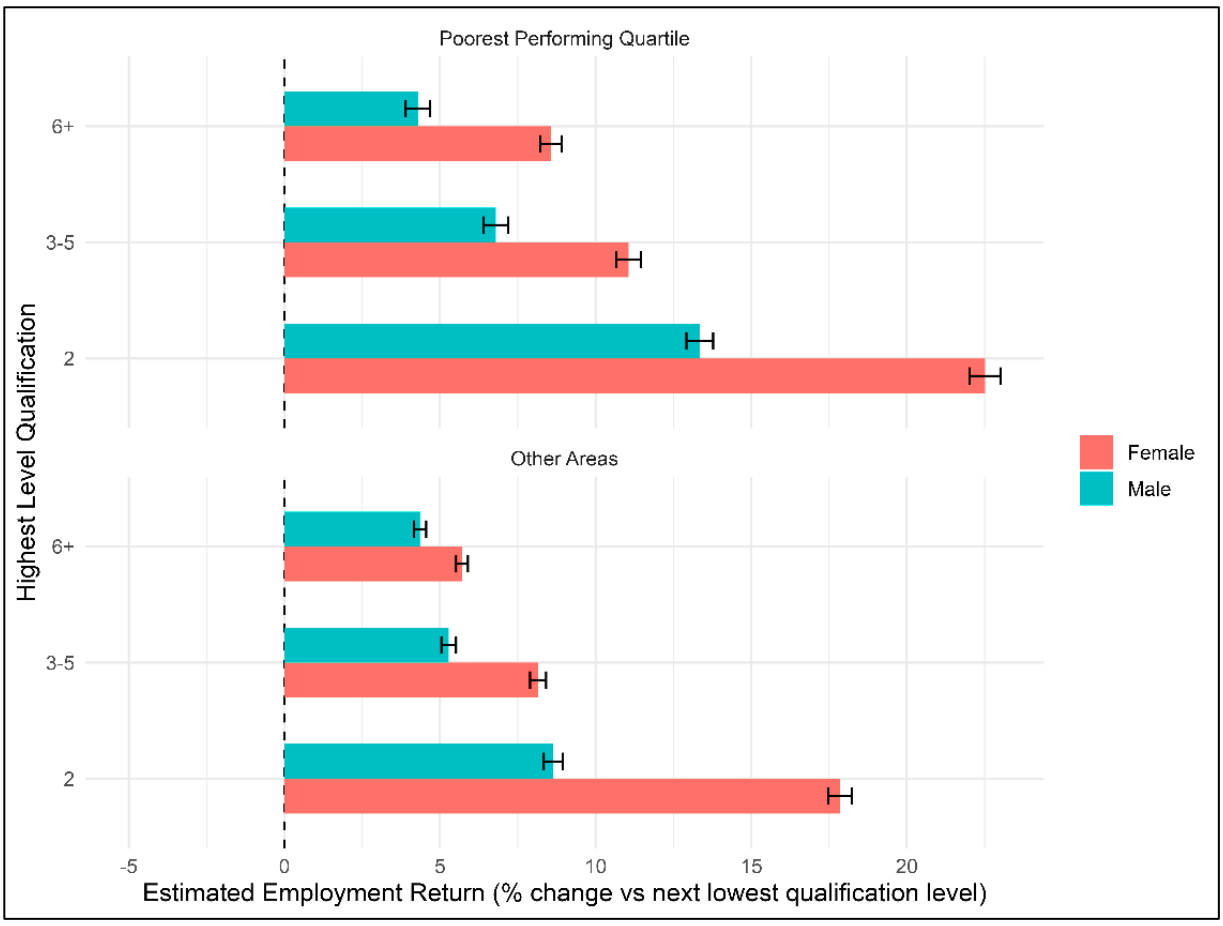


Source: LEO data.

Notes: Constructed using data on earnings at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06, adjusted for cost-of-living variation across the country. Excludes individuals with zero earnings. The 'poorest performing quartile' results are averages for individuals living in an MSOA in the most deprived IMD quartile at 27, the 'other areas' results are averages for all other individuals.

Figure A-3: Return at 27 to having a higher level of qualification, by sex and characteristics of current area of residence, conditional on childhood area of residence.

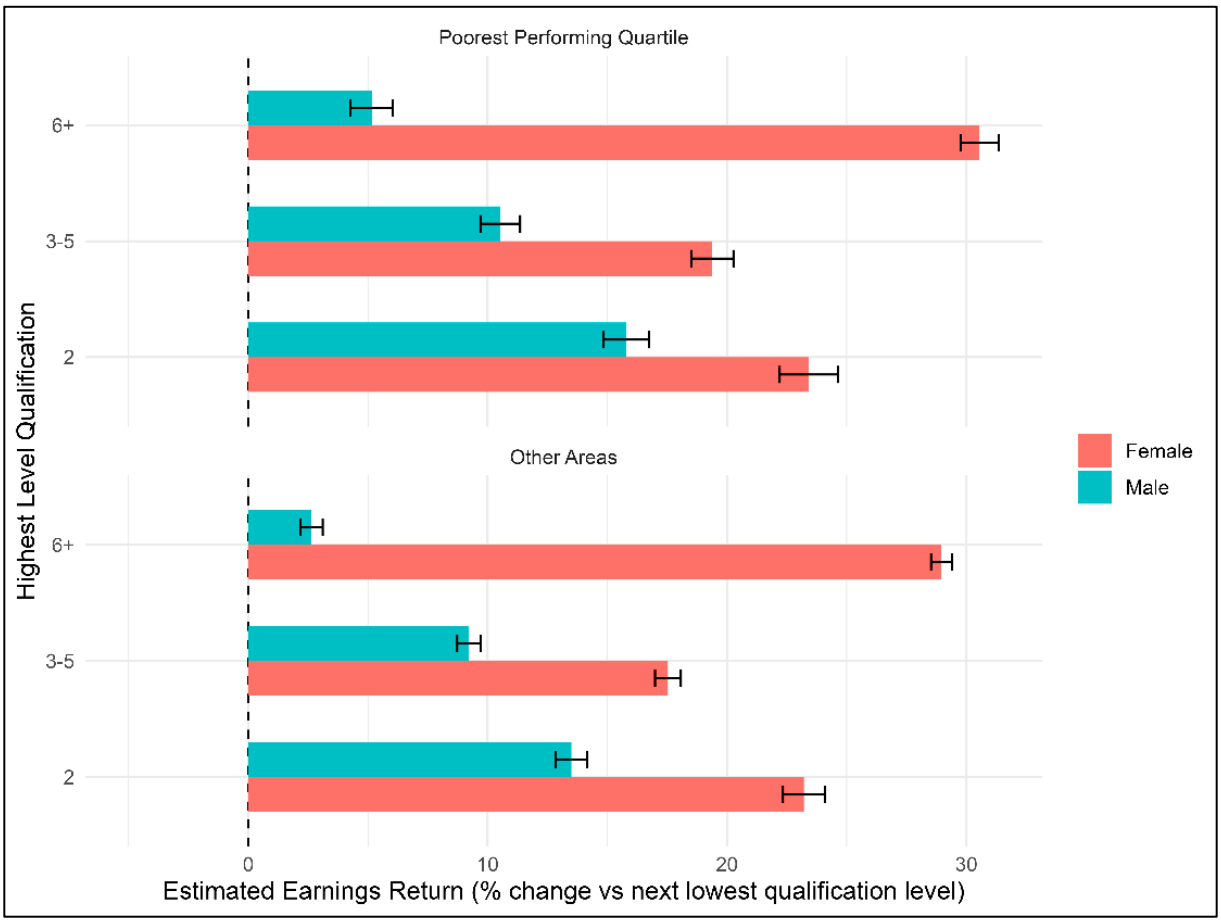
A) Employment Return



Source: LEO data.

Notes: Constructed using data on employment at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. Excludes individuals with zero earnings. Returns are estimated by comparing employment status between those with a highest qualification level at level x with those at level x-1 (i.e. individuals with level 3 are compared to those with level 2), controlling for FSM eligibility at age 16, ethnicity, cohort, KS2 attainment and TTWA of residence at age 16. Points and confidence intervals show the average probability difference of being in employment, by sex, between individuals with the two qualification levels. The 'poorest performing quartile' results are returns for individuals living in an MSOA in the most deprived IMD quartile at 27, the 'other areas' results are returns for all other individuals.

B) Earnings Return

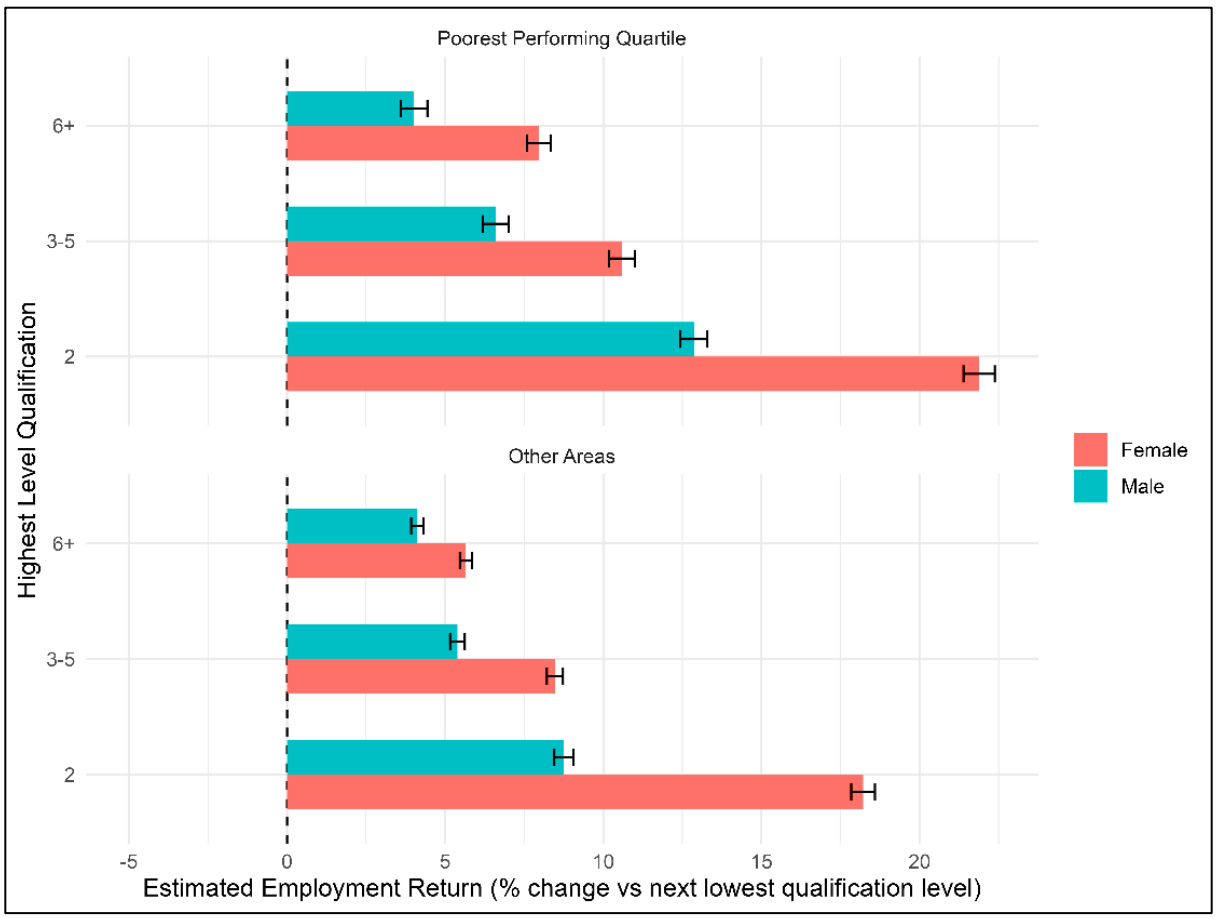


Source: LEO data.

Notes: Constructed using data on earnings at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. Excludes individuals with zero earnings. Returns are estimated by comparing incomes between those with a highest qualification level at level x with those at level x-1 (i.e. individuals with level 3 are compared to those with level 2), controlling for FSM eligibility at age 16, ethnicity, cohort, KS2 attainment and TTWA of residence at age 16. Points and confidence intervals show (approximately) the average proportion difference in earnings, by sex, between individuals with the two qualification levels. The 'poorest performing quartile' results are returns for individuals living in an MSOA in the most deprived IMD quartile at 27, the 'other areas' results are returns for all other individuals.

Figure A-4: Return at 27 to having a higher level of qualification, by sex and childhood area characteristics

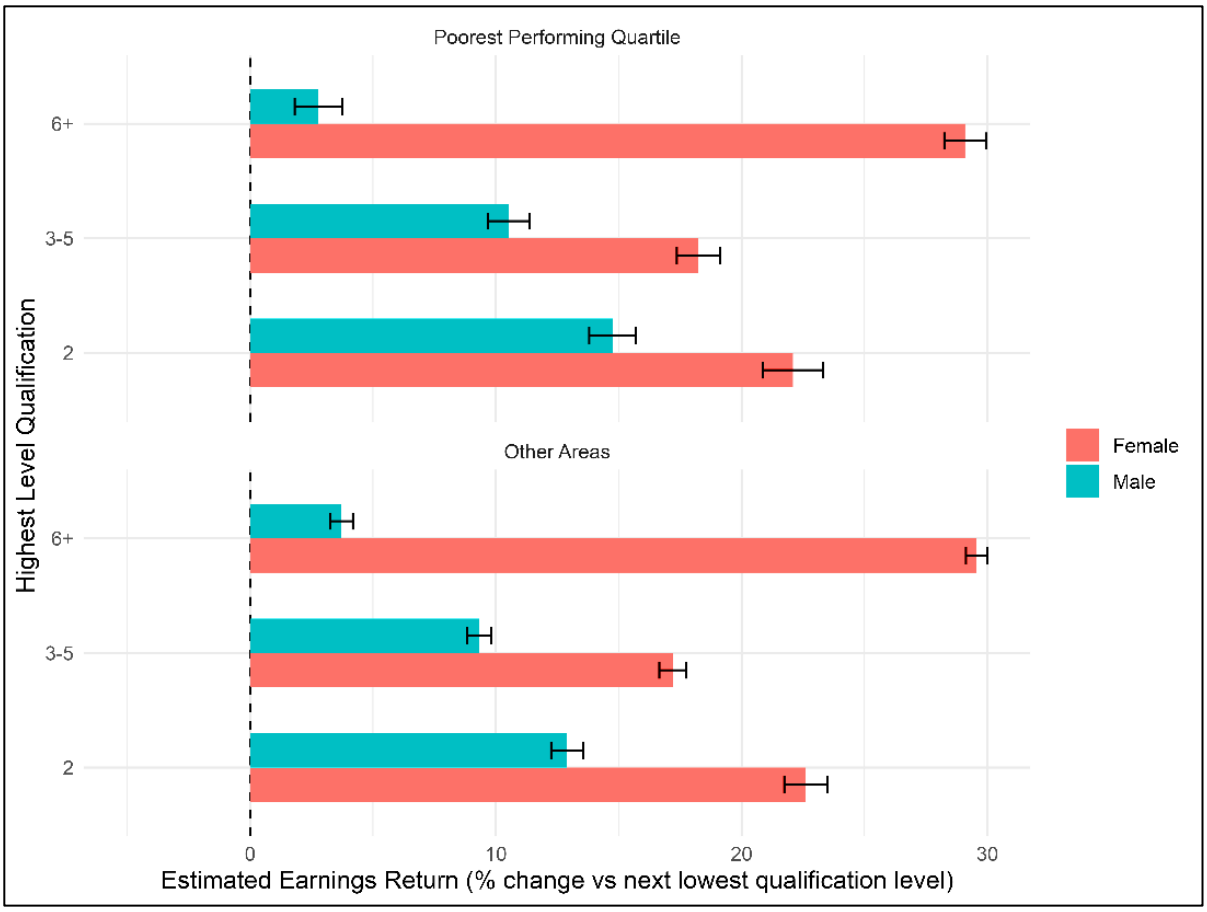
A) Employment Return



Source: LEO data.

Notes: Constructed using data on employment at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. Excludes individuals with zero earnings. Returns are estimated by comparing employment status between those with a highest qualification level at level x with those at level x-1 (i.e. individuals with level 3 are compared to those with level 2), controlling for FSM eligibility at age 16, ethnicity, cohort, and KS2 attainment. Points and confidence intervals show the average probability difference of being in employment, by sex, between individuals with the two qualification levels. The 'poorest performing quartile' results are returns for individuals living in an MSOA in the most deprived IMD quartile at 16, the 'other areas' results are returns for all other individuals.

B) Earnings Return

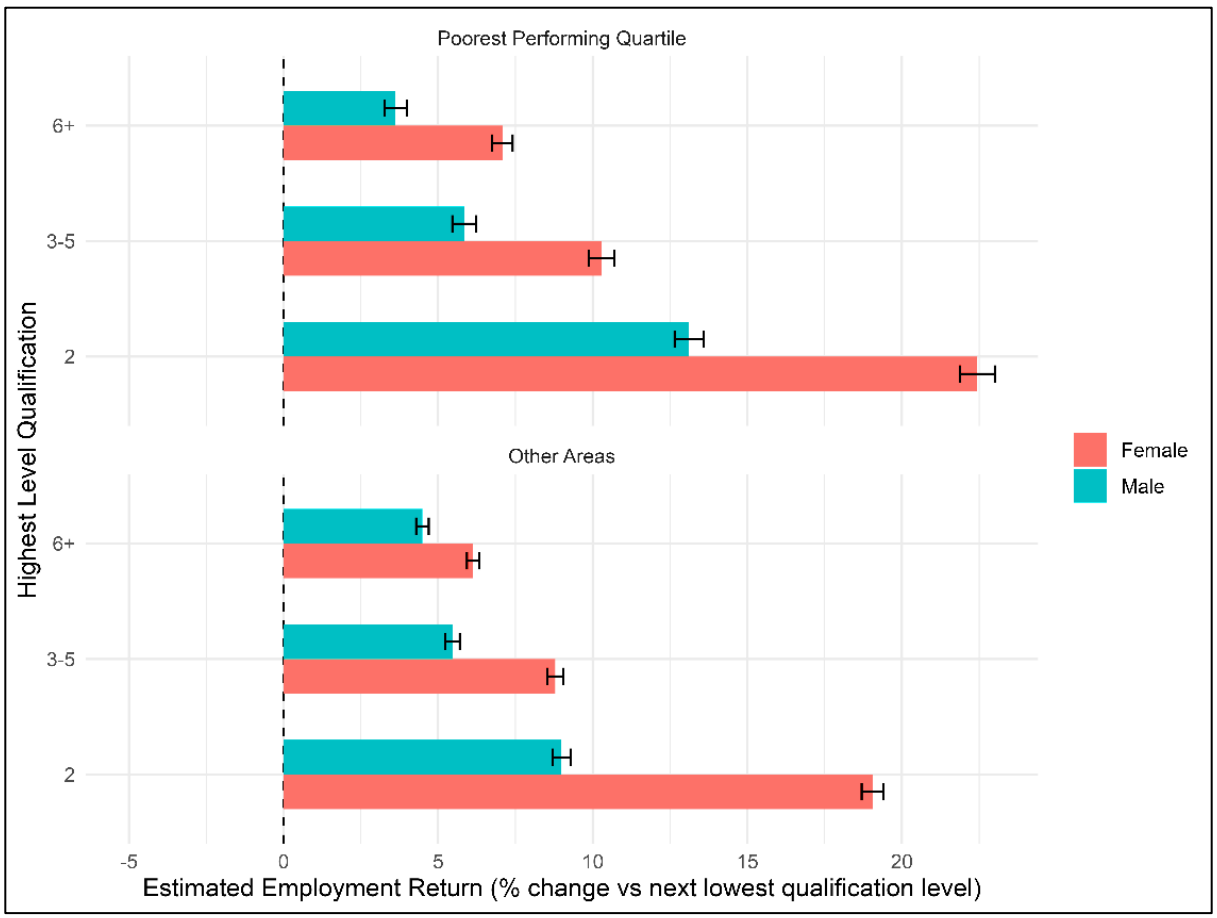


Source: LEO data.

Notes: Constructed using data on earnings at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. Excludes individuals with zero earnings. Returns are estimated by comparing incomes between those with a highest qualification level at level x with those at level x-1 (i.e. individuals with level 3 are compared to those with level 2), controlling for FSM eligibility at age 16, ethnicity, cohort, and KS2 attainment. Points and confidence intervals show (approximately) the average proportion difference in earnings, by sex, between individuals with the two qualification levels. The 'poorest performing quartile' results are returns for individuals living in an MSOA in the most deprived IMD quartile at 16, the 'other areas' results are returns for all other individuals.

Figure A-5: Return at 27 to having a higher level of qualification, by sex and current area characteristics (employment rate)

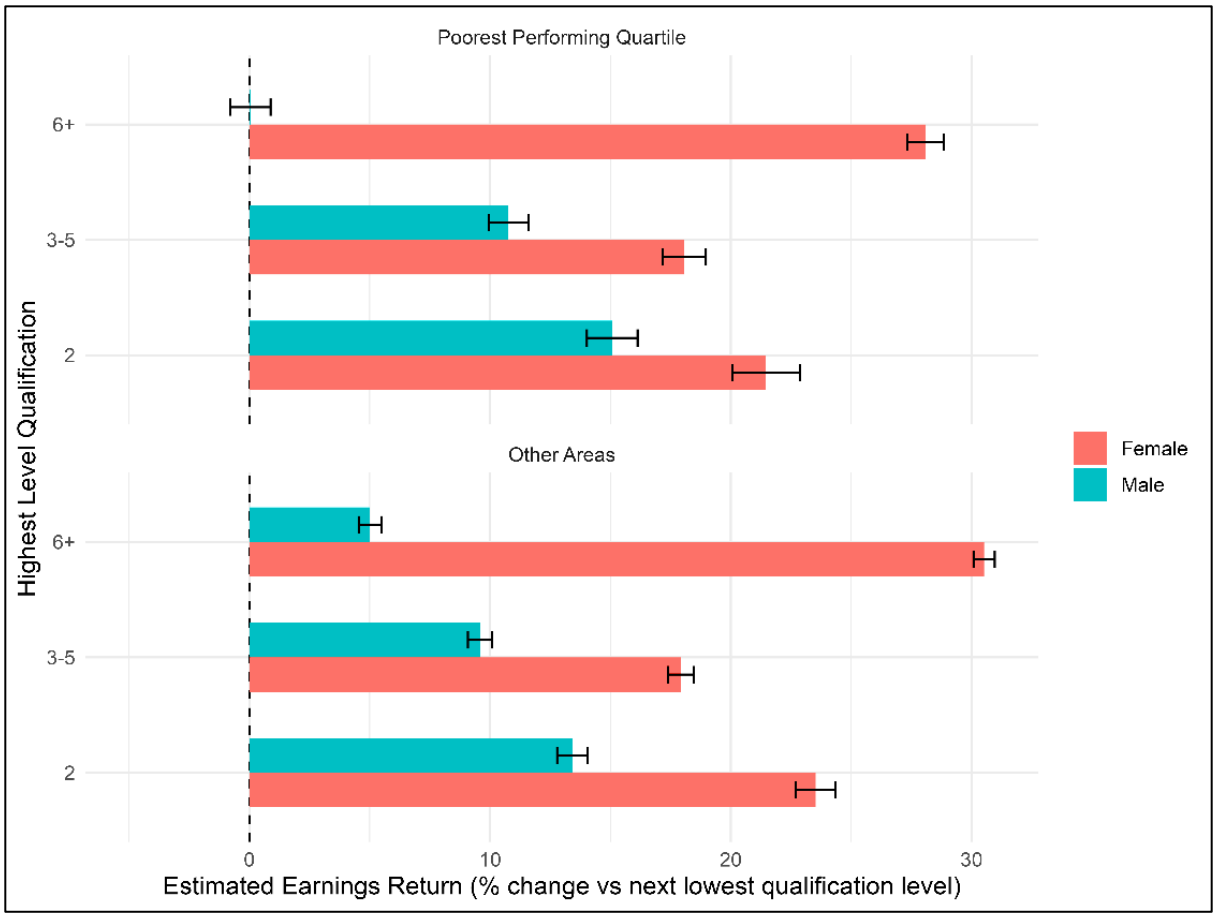
A) Employment Return



Source: LEO data.

Notes: Constructed using data on employment at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. Excludes individuals with zero earnings. Returns are estimated by comparing employment status between those with a highest qualification level at level x with those at level x-1 (i.e. individuals with level 3 are compared to those with level 2), controlling for FSM eligibility at age 16, ethnicity, cohort, and KS2 attainment. Points and confidence intervals show the average probability difference of being in employment, by sex, between individuals with the two qualification levels. The 'poorest performing quartile' results are returns for individuals living in an MSOA in the most deprived IMD quartile at 27, the 'other areas' results are returns for all other individuals.

B) Earnings Return

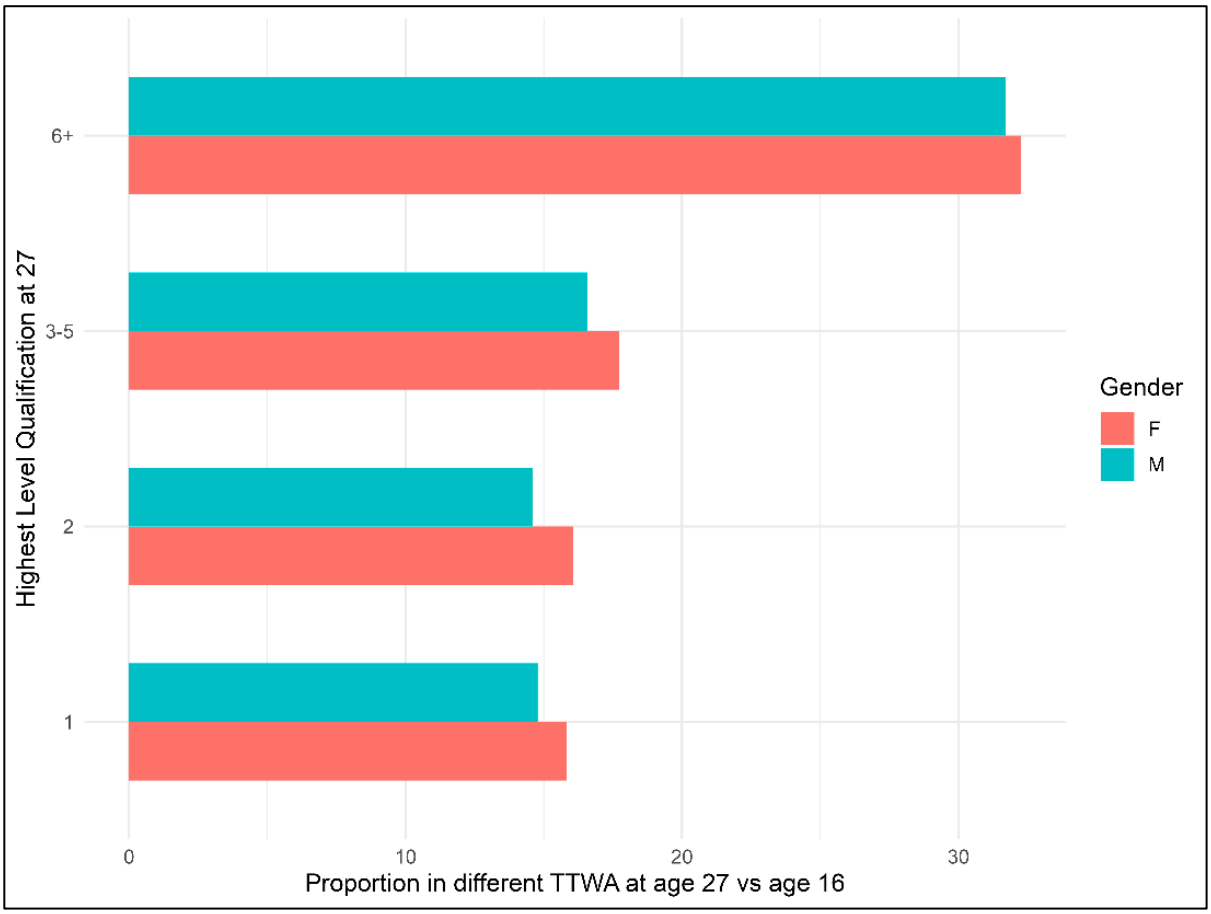


Source: LEO data.

Notes: Constructed using data on earnings at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. Excludes individuals with zero earnings. Returns are estimated by comparing incomes between those with a highest qualification level at level x with those at level x-1 (i.e. individuals with level 3 are compared to those with level 2), controlling for FSM eligibility at age 16, ethnicity, cohort, and KS2 attainment. Points and confidence intervals show (approximately) the average proportion difference in earnings, by sex, between individuals with the two qualification levels. The 'poorest performing quartile' results are returns for individuals living in an MSOA in the most deprived IMD quartile at 27, the 'other areas' results are returns for all other individuals.

Figure A-6: Descriptive proportion of individuals moving by age 27, by sex and highest qualification level

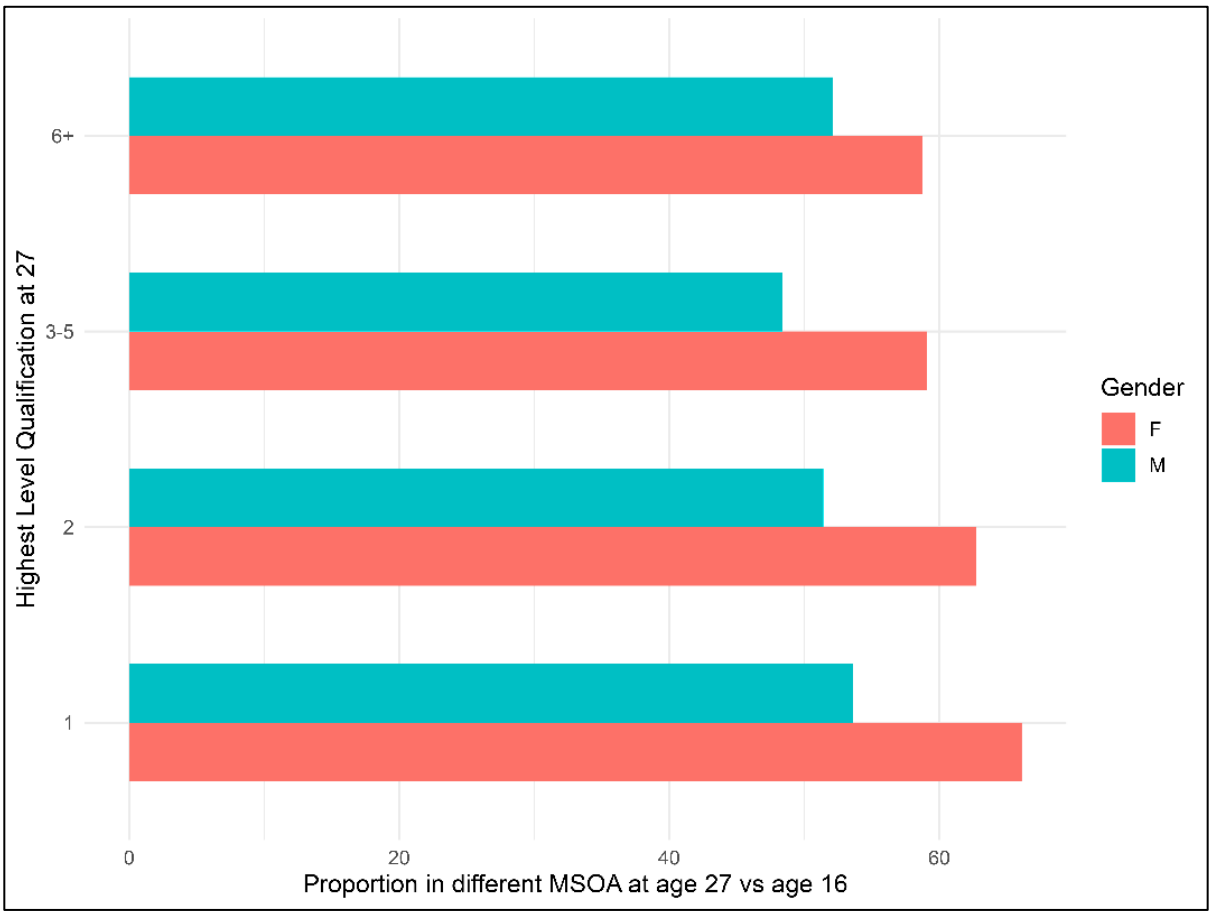
A) Proportion of individuals moving across TTWA



Source: LEO data.

Notes: Constructed using data on location at age 16 and 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different TTWA at 27 to their TTWA at 16. Poor performance quartiles are defined based on IMD at the MSOA level.

B) Proportion of individuals moving across MSOA

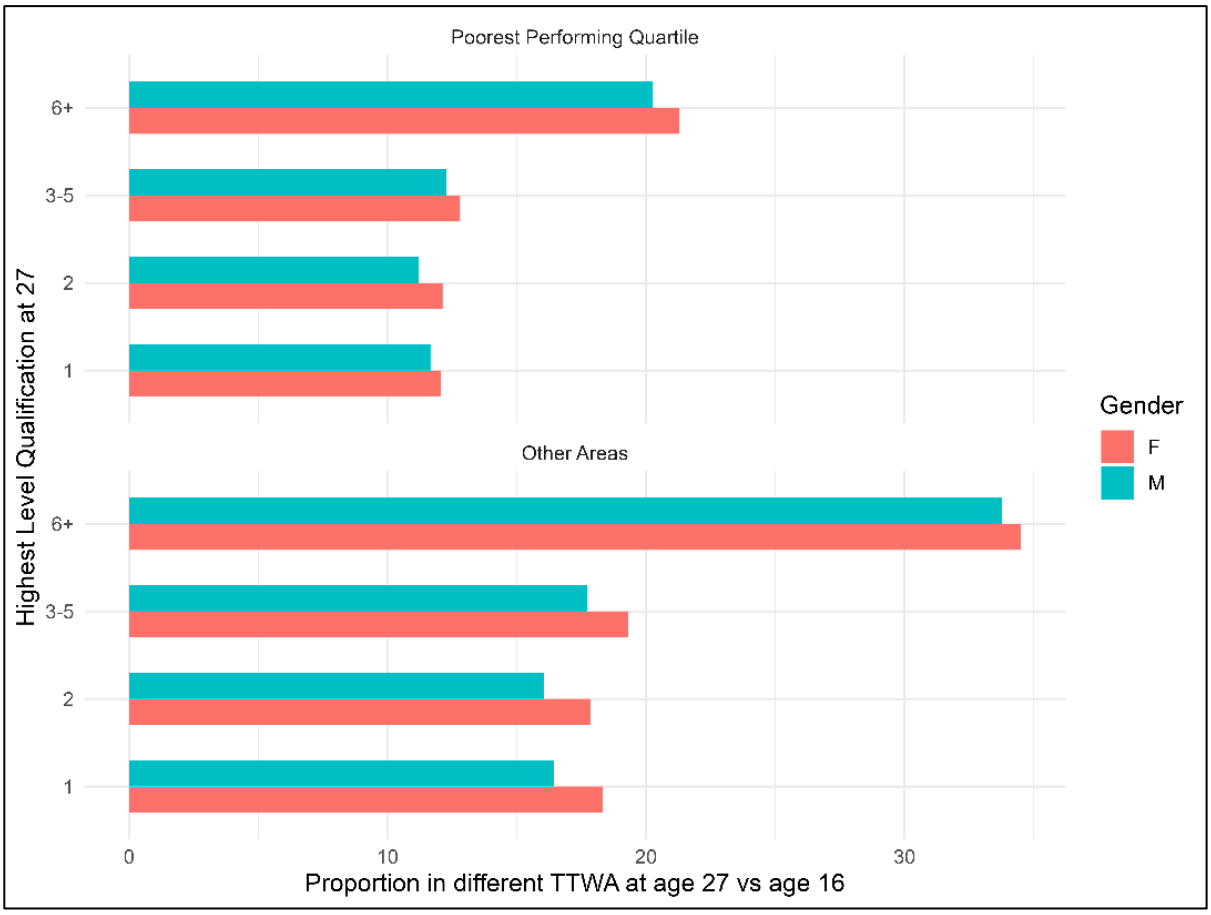


Source: LEO data.

Notes: Constructed using data on location at age 16 and 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different MSOA at 27 to their MSOA at 16. Poor performance quartiles are defined based on IMD at the MSOA level.

Figure A-7: Descriptive proportion of individuals moving by age 27, by sex, highest qualification level and childhood area characteristics

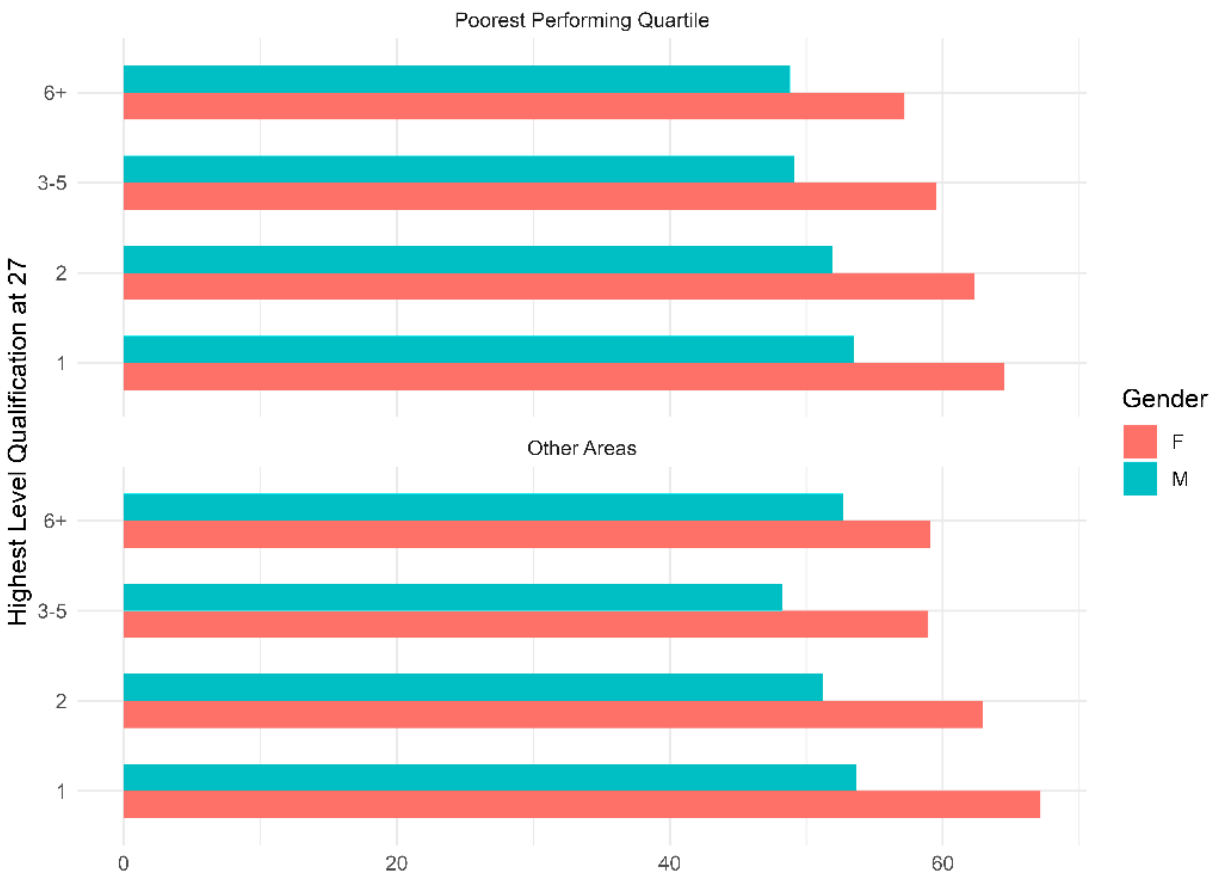
A) Proportion of individuals moving across TTWA



Source: LEO data.

Notes: Constructed using data on location at age 16 and 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different TTWA at 27 to their TTWA at 16. The model is estimated separately for non-graduates and graduates.

B) Proportion of individuals moving across MSOA

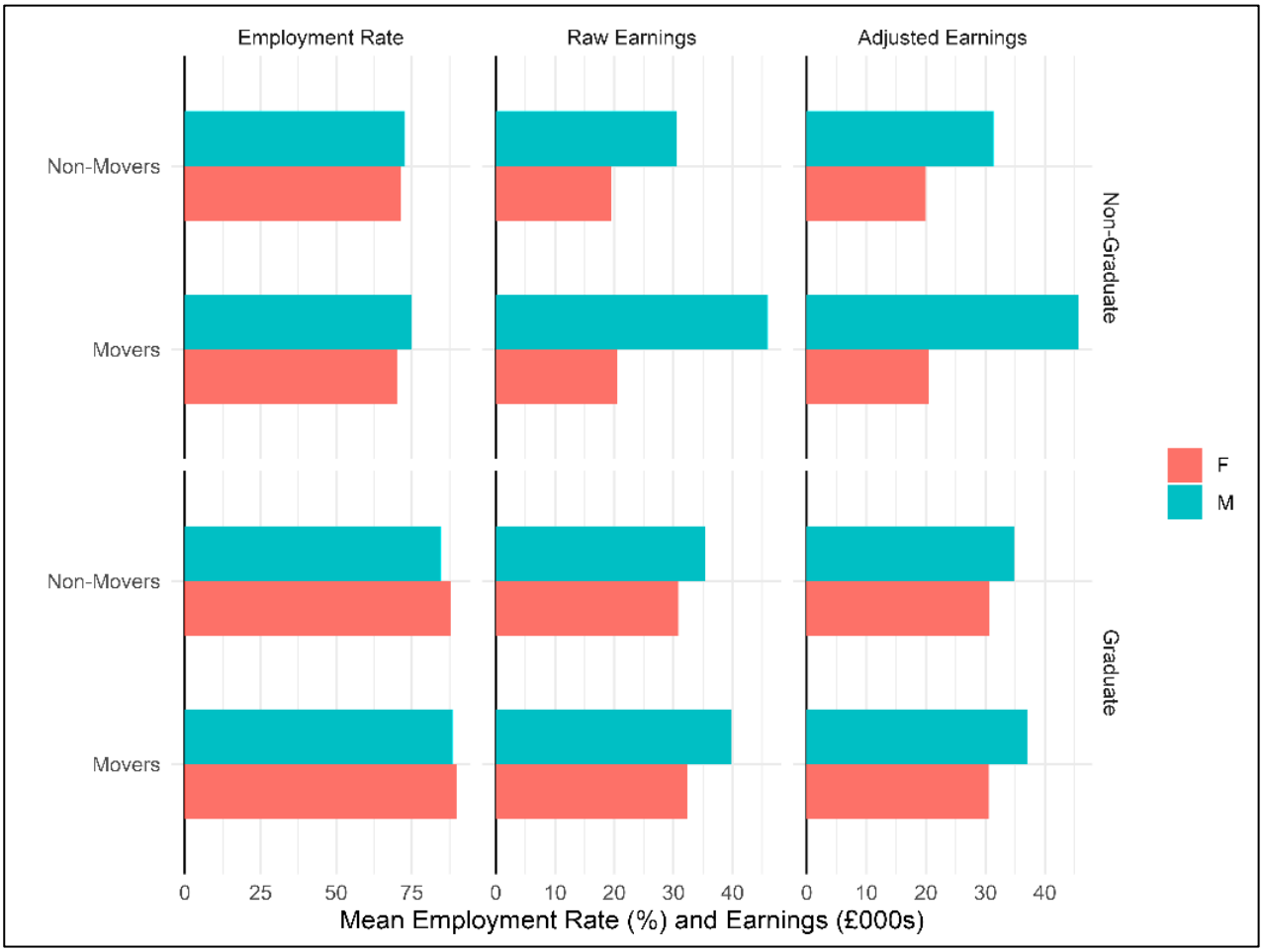


Proportion in different MSOAs at age 27 v age 16

Source: LEO data.

Notes: Constructed using data on location at age 16 and 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different MSOA at 27 to their MSOA at 16. The model is estimated separately for non-graduates and graduates.

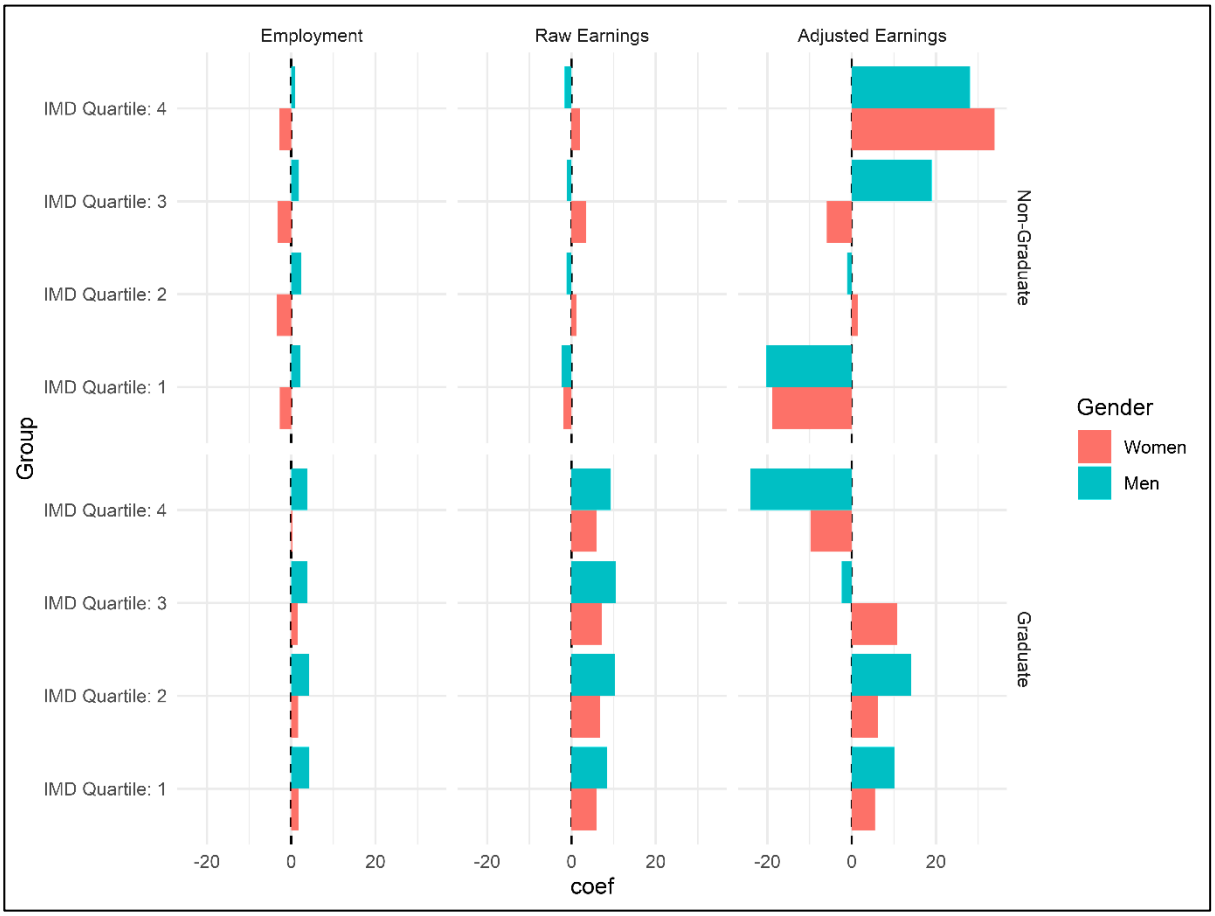
Figure A-8: Descriptive differences in employment and earnings by graduate status, sex, for between-TTWA movers and non-movers at age 27



Source: LEO data.

Notes: Constructed using data on employment, annualised earnings, and cost-of-living adjusted earnings at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different TTWA at 27 to their TTWA at 16. The model is estimated separately for non-graduates and graduates.

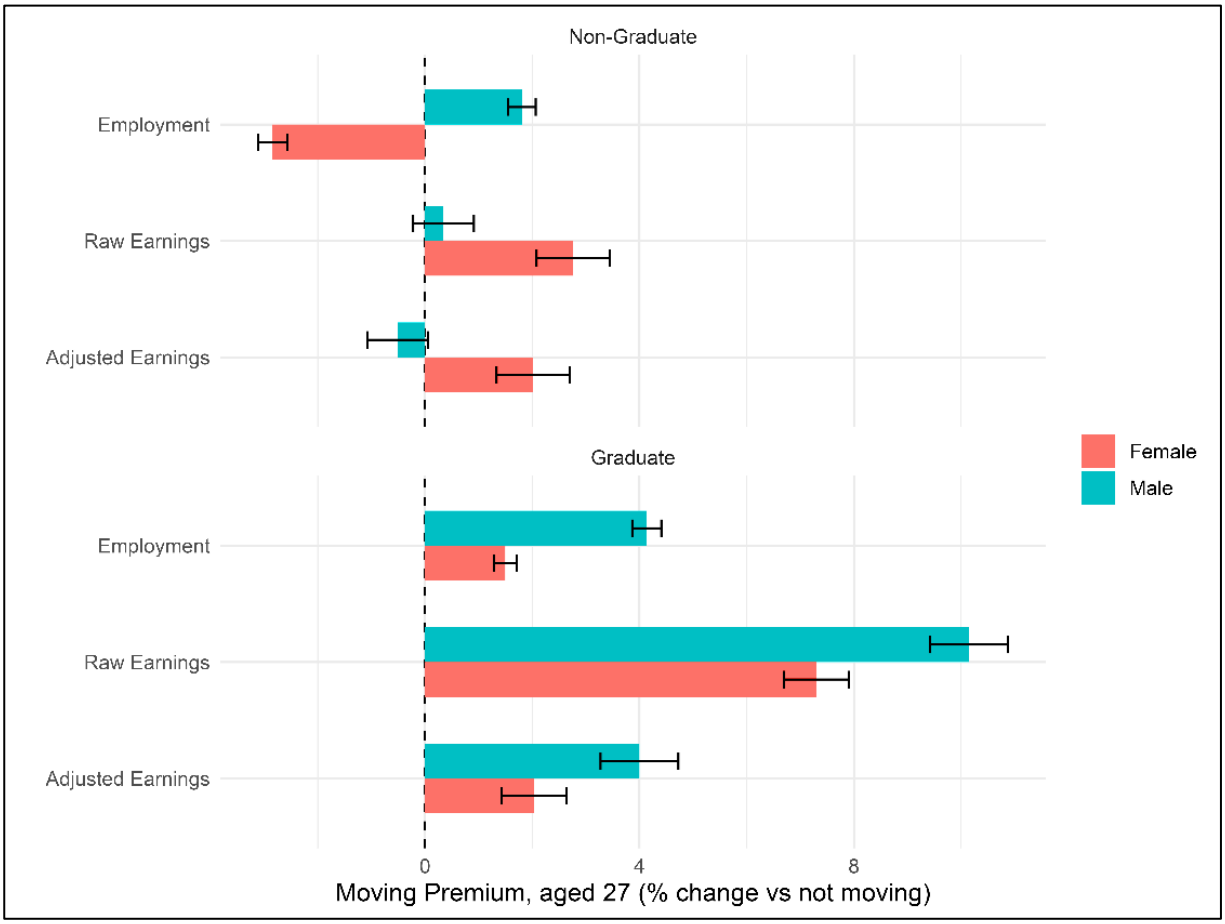
Figure A-9: Moving premium for earnings by graduate status and IMD quartile of childhood MSOA, at age 27



Source: LEO data.

Notes: Constructed using data on employment, annualised earnings, and cost-of-living adjusted earnings at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different TTWA at 27 to their TTWA at 16. The model is estimated separately for non-graduates and graduates. The premium is estimated by comparing the earnings of movers and stayers, interacted with IMD quartile, controlling for cohort, FSM eligibility at 16, TTWA at 16, and ethnicity. The estimates on employment can be interpreted as the estimated difference in probability of being in employment at 27 between movers and non-movers, the estimates on earnings as approximately the proportion difference in earnings between movers and non-movers.

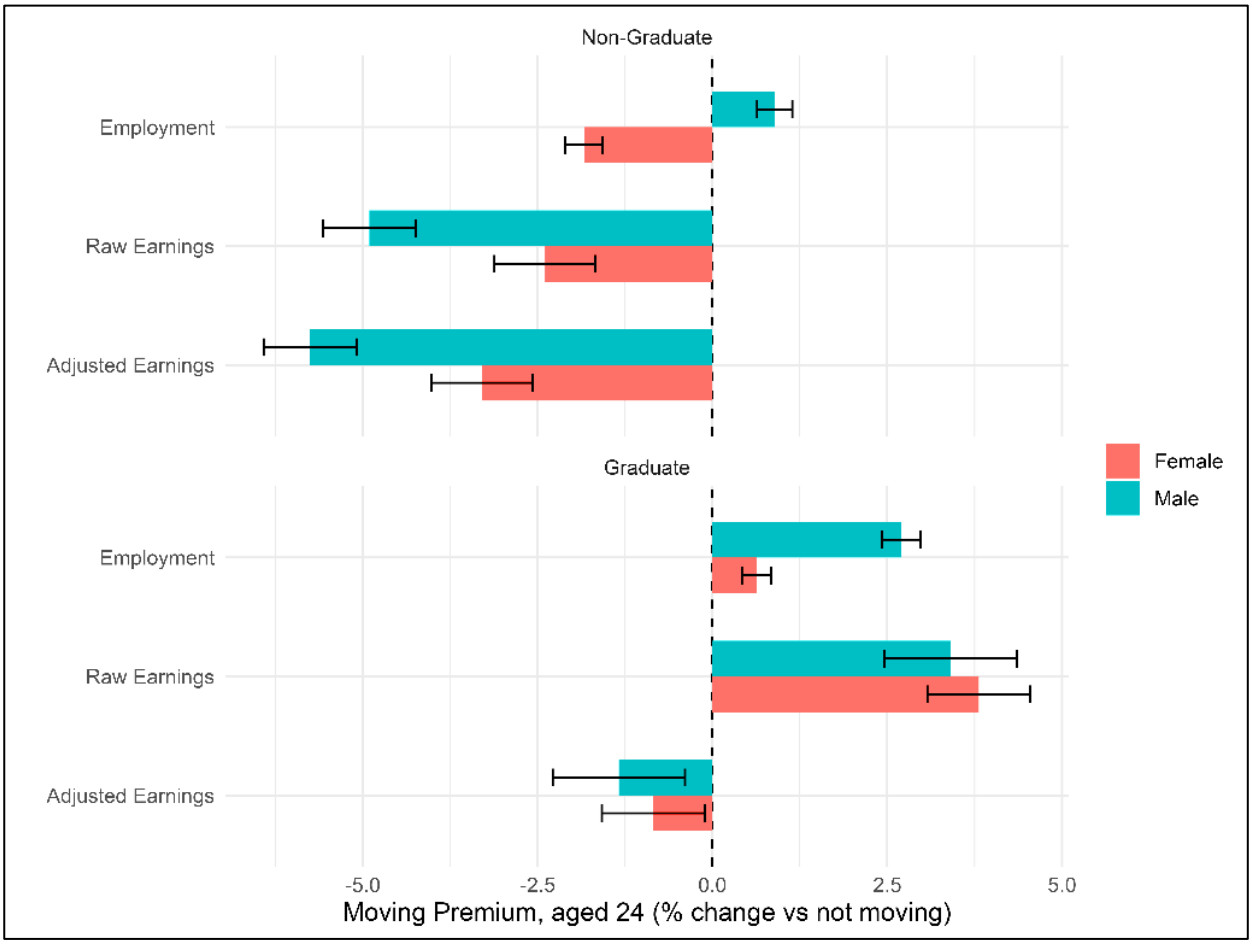
Figure A-10: Moving premium for employment and earnings by graduate status and sex, at age 27 – Alternative cost of living index



Source: LEO data.

Notes: Constructed using data on employment, annualised earnings, and cost-of-living adjusted earnings at age 27 of all individuals who took GCSEs in England between 2001/02 and 2005/06. The cost-of-living index used for this model is more conservative, assuming that price levels are constant across the country (i.e. all variation comes from house prices and rents). 'Movers' are individuals who reside in a different TTWA at 27 to their TTWA at 16. The model is estimated separately for non-graduates and graduates. The premium is estimated by comparing the earnings of movers and stayers, controlling for cohort, FSM eligibility at 16, TTWA at 16, and ethnicity. the estimates on employment can be interpreted as the estimated difference in probability of being in employment at 27 between movers and non-movers, the estimates on earnings as approximately the proportion difference in earnings between movers and non-movers.

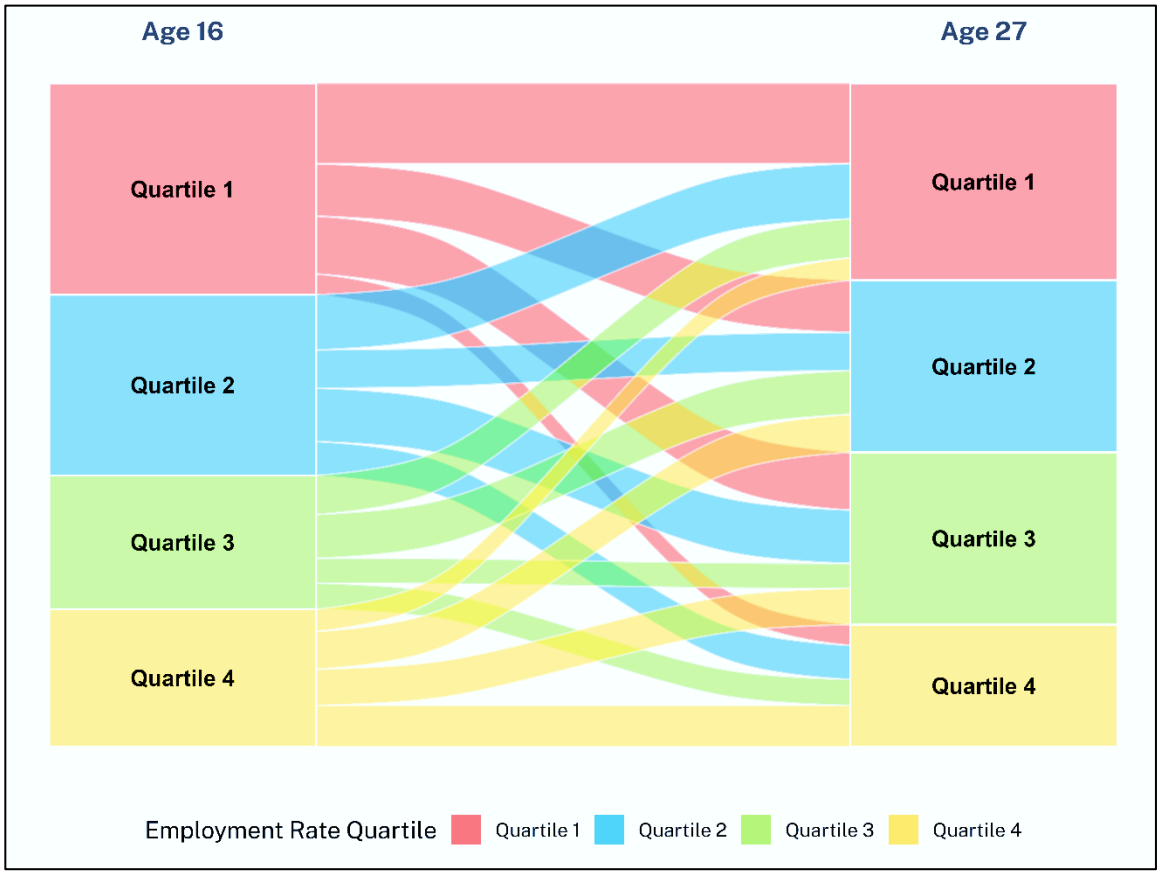
Figure A-11: Moving premium for employment and earnings by graduate status and sex, at age 24



Source: LEO data.

Notes: Constructed using data on employment, annualised earnings, and cost-of-living adjusted earnings at age 24 of all individuals who took GCSEs in England between 2001/02 and 2005/06. 'Movers' are individuals who reside in a different TTWA at 24 to their TTWA at 16. The model is estimated separately for non-graduates and graduates. The premium is estimated by comparing the earnings of movers and stayers, controlling for cohort, FSM eligibility at 16, TTWA at 16, and ethnicity. the estimates on employment can be interpreted as the estimated difference in probability of being in employment at 24 between movers and non-movers, the estimates on earnings as approximately the proportion difference in earnings between movers and non-movers.

Figure A-12: Sankey chart showing movement between areas by employment rate quartile between age 16 and age 27 (Quartile 4 = most deprived)



Source: LEO data.

Notes: Constructed using location data at 16 and 27, at the TTWA level. Quartiles defined at the TTWA level, where quartile 4 is most deprived, such that individuals can change quartile in this chart with a relatively significant move between TTWA. This chart includes only individuals who move TTWA by 27.

Are qualifications sufficient to equalise earnings across place?

Table A-1: Variance decomposition – additional measures

Metric	Gibbons et al (2014)	Our Replication	
		TTWA Level	MSOA Level
Area (RVS)	6.00%	4.98%	6.30%
Area (BVS)	1.10%	1.67%	1.66%
Area (CVS)	0.70%	1.52%	1.64%
Area (UVS)	0.10%	0.10%	0.17%
Individual (BVS)	86.8%	59.45%	59.52%
Individual (CVS)	86.40%	59.30%	59.50%
Individual (UVS)	83.60%	59.02%	57.76%
Area share of area disparities (BVS/RVS)	18.30%	33.49%	26.35%
Area share of area disparities (CVS/RVS)	12.60%	30.45%	26.01%
Area share of area disparities (UVS/RVS)	1.30%	1.98%	2.70%
Number of individuals	305,752	1,555,267	1,555,267

Notes: Constructed using earnings and location data from 2011/12 to 2016/17 tax year (i.e. each year of complete location data) for the 2001/02 to 2005/06 GCSE cohorts. Note that due to a limitation in the data available to the authors, the oldest cohort (2001/02) GCSE year only have four years of data included in this model (up to age 30). Our replication adds controls for individual fixed effects and year dummies in line with specification 4 in gibbons et al 2014. MSOA-level results are based on a set of areas produced by combining geographically contiguous MSOAs until a minimum average sample size of 200 per area per year is achieved. This iteration of the model uses 4,398 combined areas with an average sample of 340 individuals per area per year.

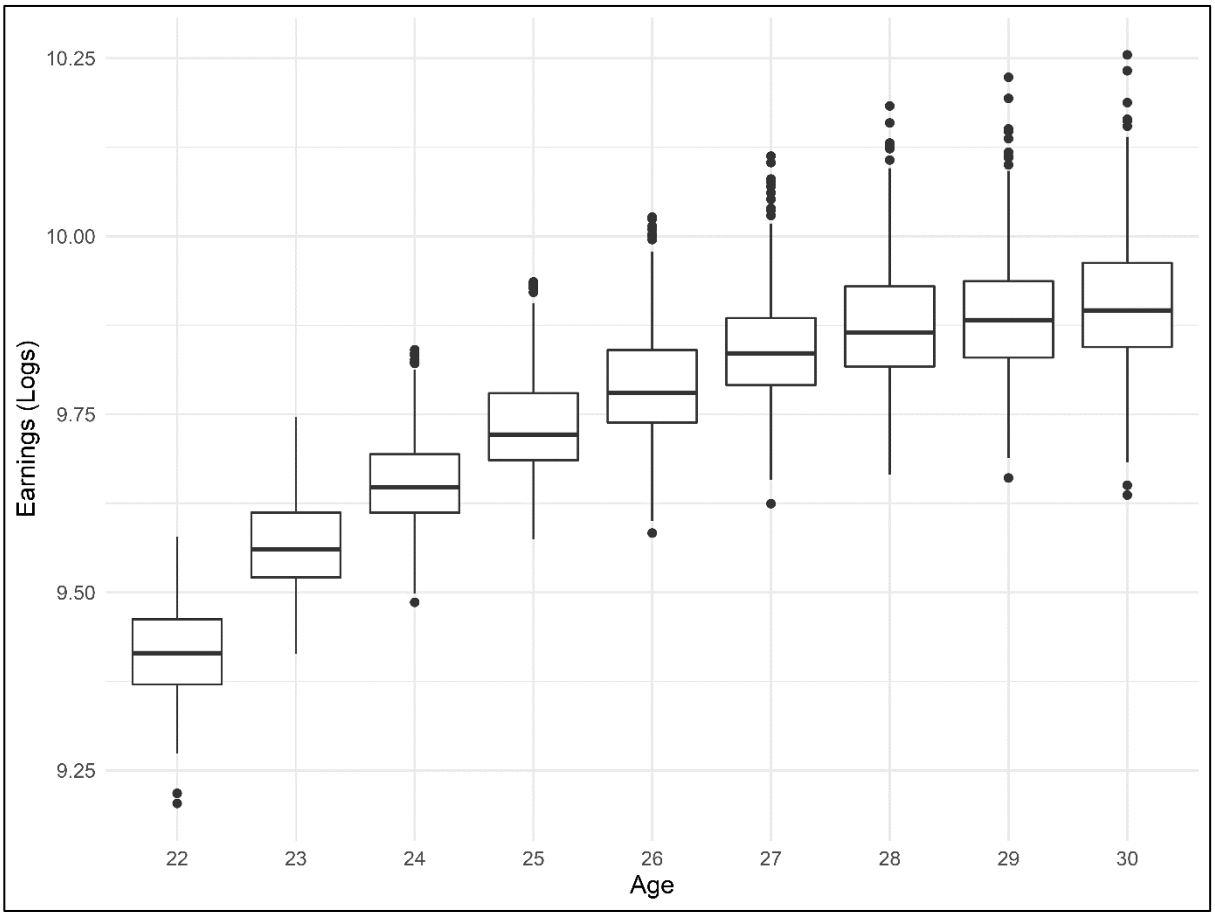
Table A-2: Variance decomposition – cost-of-living adjusted earnings

Metric	Gibbons et al (2014)	Our Replication	
		TTWA Level	MSOA Level
Area (RVS)	6.00%	3.12%	4.58%
Area (BVS)	1.10%	0.22%	0.55%
Area (CVS)	0.70%	0.38%	0.97%
Area (UVS)	0.10%	0.02%	0.15%
Individual (BVS)	86.8%	60.38%	60.18%
Individual (CVS)	86.40%	60.53%	60.60%
Individual (UVS)	83.60%	60.18%	58.85%
Area share of area disparities (BVS/RVS)	18.30%	7.16%	11.94%
Area share of area disparities (CVS/RVS)	12.60%	12.01%	21.11%
Area share of area disparities (UVS/RVS)	1.30%	0.77%	3.31%
Number of individuals	305,752	1,555,216	1,555,216

Notes: Constructed using earnings (cost-of-living adjusted with full index) and location data from 2011/12 to 2016/17 tax year (i.e. each year of complete location data) for the 2001/02 to 2005/06 GCSE cohorts. Note that due to a limitation in the data available to the authors, the oldest cohort (2001/02) GCSE year only have four years of data included in this model (up to age 30). Our replication adds controls for individual fixed effects and year dummies in line with specification 4 in gibbons et al 2014. MSOA-level results are based on a set of areas produced by combining geographically contiguous MSOAs until a minimum average sample size of 200 per area per year is achieved. This iteration of the model uses 4,398 combined areas with an average sample of 340 individuals per area per year.

Figure A-13: Variance decomposition – boxplots of mean earnings by area

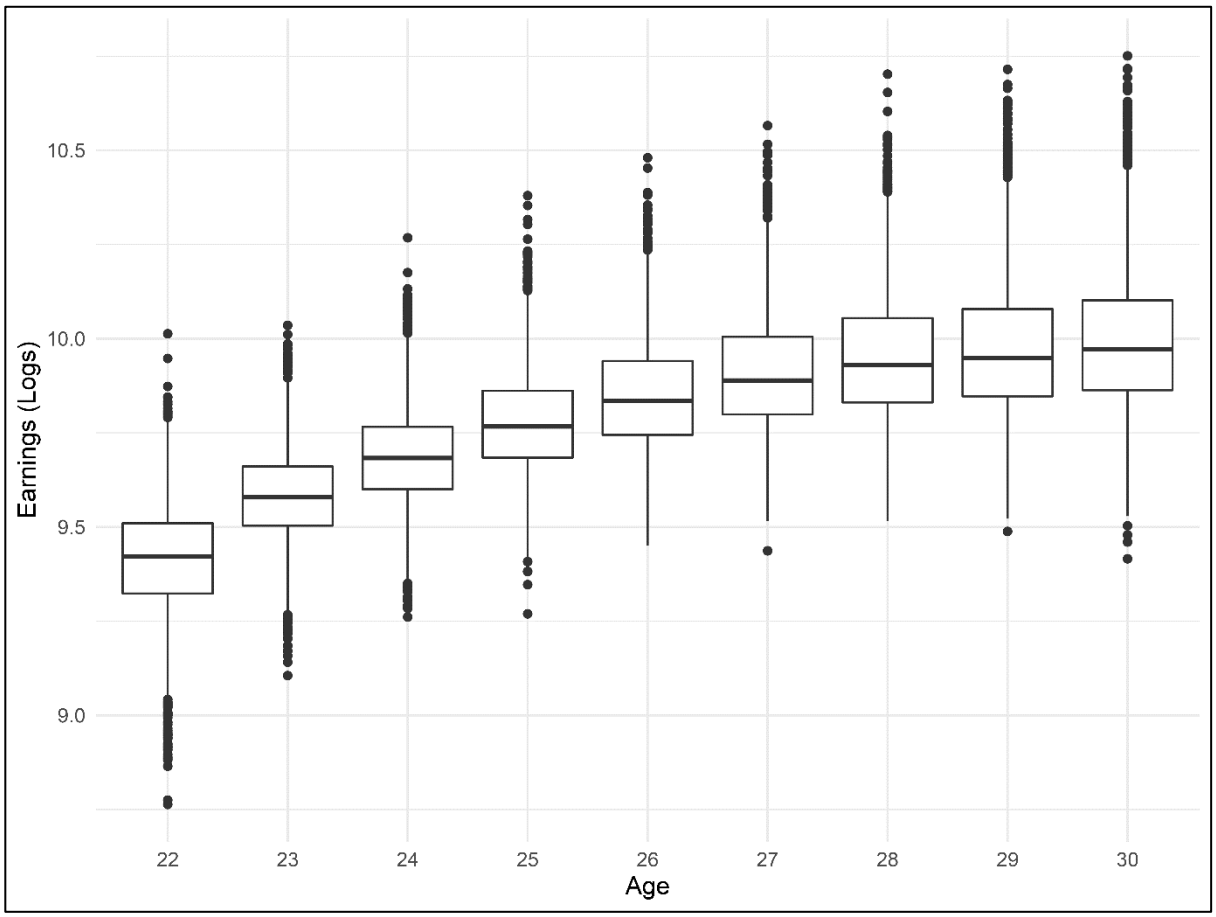
A) By TTWA



Source: LEO data

Notes: Constructed using earnings and location data from 2011/12 to 2016/17 tax year (i.e. each year of complete location data) for the 2001/02 to 2005/06 GCSE cohorts. Note that due to a limitation in the data available to the authors, the oldest cohort (2001/02) GCSE year only have four years of data included in this model (up to age 30). This chart shows the distribution of mean earnings across TTWAs.

B) By combined-MSOAs



Source: LEO data

Notes: Constructed using earnings and location data from 2011/12 to 2016/17 tax year (i.e. each year of complete location data) for the 2001/02 to 2005/06 GCSE cohorts. Note that due to a limitation in the data available to the authors, the oldest cohort (2001/02) GCSE year only have four years of data included in this model (up to age 30). This chart shows the distribution of mean earnings across MSOAs.

Appendix B: Technical Appendix

The return to qualifications for the individual

Return to qualifications in terms of employment probability and earnings, by sex (Figure 3)

Returns to qualification by sex are estimated using Ordinary Least Squares (OLS) estimation of linear regression models, specified as below. These models are run separately for each qualification level, on a limited dataset which contains only individuals at that qualification level or the level one below (e.g. comparing individuals at Level 3 with those at Level 2).

$$Employment27_i = \beta_0 + \beta_1 HighestLevelQual_i + \beta_2 Male_i + \beta_3 HighestLevelQual_i * Male_i + \beta_4 Year27_i + \gamma X_i + \mu_i$$

$$\ln(Earnings27_i) = \beta_0 + \beta_1 HighestLevelQual_i + \beta_2 Male_i + \beta_3 HighestLevelQual_i * Male_i + \beta_4 Year27_i + \gamma X_i + \mu_i$$

Where:

- i indexes individuals;
- $Employment27_i$ is a dummy for employment at age 27, and $Earnings27_i$ are annualised earnings at age 27;
- $HighestLevelQual_i$ is a dummy equal to 1 if an individual has a qualification at level X, and 0 if their highest qualification is at level X-1 (X varies depending on the model iteration);
- $Male_i$ is a dummy equal to 1 if an individual is male, and 0 if they are female;
- $Year27_i$ are a set of dummies for the year in which an individual is 27 years old;
- X_i is a vector of characteristic controls, including a dummy variable for FSM eligibility at age 16, a set of dummies for ethnicity, and KS2 attainment in English and Maths; and
- μ_i are residuals.

Return at 27 to having a higher level of qualification, by sex and current area characteristics (Figure 4)

Returns to qualification by sex and current area characteristics are estimated in much the same way as the previous result, adding an additional set of interactions between current area characteristics and sex:

$$Employment27_i = \beta_0 + \beta_1 HighestLevelQual_i + \beta_2 PP_i + \beta_3 HighestLevelQual_i * PP_i + \beta_4 Male_i + \beta_5 Male_i * (\beta_6 HighestLevelQual_i + \beta_7 PP_i + \beta_8 HighestLevelQual_i * PP_i) + \beta_9 Year27_i + \gamma X_i + \mu_i$$

$$\ln(Earnings27_i) = \beta_0 + \beta_1 HighestLevelQual_i + \beta_2 PP_i + \beta_3 HighestLevelQual_i * PP_i + \beta_4 Male_i + \beta_5 Male_i * (\beta_6 HighestLevelQual_i + \beta_7 PP_i + \beta_8 HighestLevelQual_i * PP_i) + \beta_9 Year27_i + \gamma X_i + \mu_i$$

Where definitions are as in the previous specification, and:

- PP_i is a dummy equal to 1 if an individual lives in an area (MSOA or TTWA depending on the specification) in the poorest performing quartile on the IMD metric (or employment rate metric in alternate specifications) at age 27.

Return at 27 to having a higher level of qualification, by sex and current area characteristics, conditional on childhood area of residence (Figure A-3)

This result uses models almost identical to the previous specification, adding fixed effects for current area of residence:

$$Employment27_i = \beta_0 + \beta_1 HighestLevelQual_i + \beta_2 PP_i + \beta_3 HighestLevelQual_i * PP_i + \beta_4 Male_i + \beta_5 Male_i * (\beta_6 HighestLevelQual_i + \beta_7 PP_i + \beta_8 HighestLevelQual_i * PP_i) + \beta_9 Year27_i + \gamma X_i + \delta A_i + \mu_i$$

$$\ln(Earnings27_i) = \beta_0 + \beta_1 HighestLevelQual_i + \beta_2 PP_i + \beta_3 HighestLevelQual_i * PP_i + \beta_4 Male_i + \beta_5 Male_i * (\beta_6 HighestLevelQual_i + \beta_7 PP_i + \beta_8 HighestLevelQual_i * PP_i) + \beta_9 Year27_i + \gamma X_i + \delta A_i + \mu_i$$

Where definitions are as in the previous specification, and:

- A_i are fixed effects for TTWA of residence at age 16.

The return to qualifications for the individual

Probability of moving by age 27 controlling for background characteristics, by level of qualification (Figure 5)

Estimates of differences in moving probability by level of qualifications are produced by OLS estimation of a linear probability regression model, specified as below, and run separately by sex to produce different estimates for men and women:

$$Moved27_i = \beta_0 + \beta_1 HighestLevelQual_i + \gamma X_i + \mu_i$$

Where:

- i indexes individuals;
- $Moved27_i$ is a dummy equal to 1 if an individual has moved area (either TTWA/MSOA depending on the specification) between 16 and 27, and 0 if not;
- $HighestLevelQual_i$ is a dummy equal to 1 if an individual has a qualification at level X, and 0 if their highest qualification is at level X-1 (X varies depending on the model iteration); and
- X_i is a vector of characteristic controls, including a dummy variable for FSM eligibility at age 16, a set of dummies for ethnicity, and KS2 attainment in English and Maths.

Probability of moving by age 27, by graduate status and IMD quartile of childhood MSOA (Figure 7).

This result uses a similar linear probability model as the above, this time run separately for each combination of graduate status and sex:

$$Moved27_i = \beta_0 + \beta_1 PPQuartile_i + \gamma X_i + \mu_i$$

Where definitions are as in the previous specification, and:

- $PPQuartile_i$ is a vector of dummies for the quartile of the individual's childhood MSOA on our IMD metric.

Moving premium for employment and earnings by graduate status and sex, at age 27 (Figure 8)

The 'moving premium' for employment and earnings is calculated using OLS estimation of a linear regression model, with the specification as below. The specification is very similar to the definition of the moving premium in Britton et al. (2021b). This model is run separately for each combination of graduate status and sex to produce separate estimates for these subgroups:

$$Employment27_i = \beta_0 + \beta_1 Moved_i + \gamma X_i + \delta A_i + \mu_i$$

$$\ln(Earnings_i) = \beta_0 + \beta_1 Moved_i + \gamma X_i + \delta A_i + \mu_i$$

Where:

- i indexes individuals;
- $Employment_i$ is a dummy for employment status, and $Earnings_i$ are annualised earnings – these variables are defined either at age 24 or 27 depending on the specification;
- $Moved_i$ is a dummy equal to 1 if an individual has moved area (either TTWA or MSOA depending on the specification) and 0 otherwise – this variable is defined either at age 24 or 27 depending on the specification;
- X_i is a vector of characteristic controls, including a dummy variable for FSM eligibility at age 16, a set of dummies for ethnicity, and KS2 attainment in English and Maths; and
- A_i are fixed effects for TTWA of residence at 16.

Estimates of the moving premium by highest qualification level and sex use very similar models to the previous specification, using an interaction term instead of running the models separately by subgroup:

$$Employment27_i = \beta_0 + \beta_1 Moved_i + \beta_2 HighestLevelQual_i + \beta_3 Moved_i * HighestLevelQual_i + \gamma X_i + \delta A_i + \mu_i$$

$$\ln(Earnings_i) = \beta_0 + \beta_1 Moved_i + \beta_2 HighestLevelQual_i + \beta_3 Moved_i * HighestLevelQual_i + \gamma X_i + \delta A_i + \mu_i$$

Where:

- $HighestLevelQual_i$ is a vector of dummies for an individual's highest level of qualification.

Are qualifications sufficient to equalise earnings across place?

Decomposition of variance in earnings (Table 3)

This result follows methodology outlined in Gibbons et al (2014) to produce different measures of the proportion of variation in wages in our LEO dataset which are explained by area effects, as opposed to individual characteristics. Panel data on location and earnings each year from 2012-2017 allows us to use movers between places to identify these area effects, controlling for individual characteristics.

The core specification is as follows (the equivalent of specification 4 in Gibbons et al (2014)):

$$\ln(Earnings_{it}) = \lambda year_t + \gamma PMR_i + \delta A_{it} + \mu_i$$

Where:

- i indexes individuals, t indexes time;
- $\ln(Earnings_{it})$ is the log of earnings;
- $year_t$ is a vector of dummies for the year;
- PMR_i is a vector of unique dummies for each individual i (individual fixed effects);
- A_{it} is a vector of dummies for an individual's area of residence at a given time (defined at either TTWA or combined-MSOA level depending on the specification); and
- μ_i are residuals.

There are a number of possible estimators for the share of variation in wages that is attributable to the area fixed effects in this specification. The measures proposed in Gibbons et al. (2014) are as follows.

As a clear upper bound, there is the Raw Variance Share:

$$RVS = \frac{var(\delta A_{it})}{var(\ln(Earnings_{it}))} = R^2(\ln(Earnings_{it})|A_{it})$$

Where $R^2(\ln(Earnings_{it})|A_{it})$ is the R-squared from regressing earnings on area dummies with no control for individual fixed-effects. If (as we would expect) average individual characteristics vary by area due to sorting (and are not *caused* by area differences) then RVS over-estimates the share of earnings variation attributable to area effects.

Correlated Variance Share (CVS) is an alternative drawn from the full specification above, i.e. controlling for individual fixed effects when estimating area fixed effects:

$$CVS = \frac{var(\delta A_{it})}{var(ln(Earnings_{it}))}$$

This measure is the simple ratio of the variance of area effects to the variance of log wage – however while it excludes the direct contribution of sorting across areas, it includes any indirect effect of sorting on the variance of area effects themselves.

Uncorrelated Variance Share (UVS) isolates the contribution to wage variance of the parts of area effects which are uncorrelated with area characteristics:

$$UVS = R^2(ln(Earnings_{it})|PMR_{it}, A_{it}) - R^2(ln(Earnings_{it})|PMR_{it})$$

Note this measure uses both the core specification above, and a model which regresses earnings on individual fixed effects alone to calculate $R^2(ln(Earnings_{it})|PMR)$.

Balanced Variance Share is similar to CVS in that it considers the ratio of the variance of area effects to the variance of wages – except rather than excluding the direct effect of sorting between areas, it apportions the covariance component of total earnings variance equally between area and individual effects.

$$BVS = \frac{var(\delta A_{it}) + cov(\gamma PMR_i, \delta A_{it})}{var\ (ln(Earnings_{it}))}$$

Gibbons et al (2014) show that these estimators all give different answers when there is associative sorting across areas – and that which variance share is most appropriate depends on the underlying structure of area and individual effects on earnings. These four estimators (RVS, CVS, UVS, BVS) are provide reasonable boundaries for the contributions of area effects to earnings variance.

Bibliography

Britton, J., L. Dearden and B. Waltmann (2021a), *The returns to undergraduate degrees by socio-economic group and ethnicity*. IFS Research Report. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/973239/The_returns_to_undergraduate_degrees_by_socio-economic_group_and_ethnicity.pdf

Britton, J., L. Van der Erve, B. Waltmann and X. Xu (2021b), *London calling? Higher education, geographical mobility and early-career earnings.* IFS Research Report. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1022365/London_calling_final_for_dfe_with_accessibility.pdf

Chetty, R., Hendren, N., Kline, P. and Saez, E. (2014), Where is the land of opportunity? The geography of intergenerational mobility in the United States, *The Quarterly Journal of Economics*, Vol. 129, pp.1553-1623.

Chetty, R. and Hendren, N. (2018), The impacts of neighbourhoods on intergenerational mobility II: county-level estimates, *The Quarterly Journal of Economics*, Vol. 133, pp.1163-1228.

HM Government (2022), *Levelling Up the United Kingdom,* CP 604. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1052706/Levelling_Up_WP_HRES.pdf.

Gibbons, S., Overman, H.G. and Pelkonen, P. (2014), Area disparities in Britain: Understanding the contribution of people vs. place through variance decompositions, *Oxford Bulletin of Economics and Statistics*, Vol. 76, pp.745-763.

Green, A., Shuttleworth, I. and Lavery, S. (2005), Young people, job search and local labour markets: the example of Belfast, *Urban Studies*, Vol. 42, pp.301-324.

Green, A.E. and White, R. (2008), Shaped by place: young people's decisions about education, training and work, *Benefits*, Vol. 16, pp.213-224.

Skills and Productivity Board. (2022) *Opportunities and challenges for improving labour market information on skills.* Available at: https://www.gov.uk/government/groups/skills-and-productivity-board

Social Mobility Commission (2020), *The long shadow of deprivation: Differences in opportunities across England*. Research report. Available at: HTTPS://WWW.GOV.UK/GOVERNMENT/NEWS/THE-LONG-SHADOW-OF-DEPRIVATION

About this publication:
enquiries www.education.gov.uk/contactus
download www.gov.uk/government/publications