Department for Education

# Labour Market Outcomes: Chapter 2

## Socioeconomic Status

### November 2024


### Moira Nelson and Oliver Anderson, Department for Education

# Contents

# List of figures

# Chapter summary

The second chapter in this three-part series focuses on the impact of **socioeconomic status** on labour market outcomes. Here, we use a measure of socioeconomic status derived from free school meals (FSM) eligibility and some local area deprivation measures. The variables are combined into a continuous measure of SES and split into quintiles of deprivation. The lowest SES group (1st quintile) is the most disadvantaged fifth of the population and the highest SES group (5th quintile) is the least disadvantaged. The other chapters look at ethnicity and special educational needs (SEN).

# Key findings

1. **Individuals from lower SES quintiles are less likely to be in a good outcome and more likely to be in a poor outcome than those from higher SES quintiles**

Those from the highest SES quintile are over three times more likely to be in a good labour market outcome than those from the lowest SES quintile. There is gradual progression in the likelihood of good outcome with each higher quintile.

Those from the lowest SES quintile are around five times more likely to be in a poor outcome than those from the highest SES quintile. The biggest gap between quintiles is between the two lowest SES quintiles, with smaller gaps between the middle and top two quintiles.

These gaps in outcomes between socioeconomic backgrounds persist across demographics and educational achievement breakdowns.

2. **Females show a bigger disparity between the highest and lowest SES quintiles than males for both good and poor outcome**

Females from the lowest SES quintile are **less** likely to be in a good outcome than males from the same quintile (6 per cent compared to 8 per cent, respectively) while females from the highest SES quintile are **more** likely to be in a good outcome than males from the same SES quintile (26 per cent compared to 22 per cent, respectively). Consequently, the disparity in good outcome between the highest and lowest SES quintiles is bigger for females (19 percentage points) than it is for males (15 percentage points).

Females from the lowest SES quintile are both **more** likely to be in a poor outcome than males from the same quintile (23 per cent compared to 14 per cent) and **more** likely to be in a poor outcome than males from the same quintile (5 per cent compared to 3 per cent). In addition, the gap in poor outcome between the lowest and highest SES quintiles are much higher for females than males (19 percentage points for females compared to 11 percentage points for males).

3. **The important factors in explaining differences in outcomes between the most and least deprived groups are similar for males and females, and for good and poor outcome**

For females, having a degree is the most substantial factor in explaining the gap in good outcome between the highest and lowest SES quintiles, followed by pre-16 attainment, then pre-16 school and peer factors. Demographic factors are less important, but still explain some of the gap, more so region of school than the others.

For males, the same factors are important in explaining the gap in good outcome between the highest and lowest SES quintiles as for females, although pre-16 attainment for males is more important in explaining the difference in good outcome between the highest and lowest SES quintiles than achieving a degree.

When looking at poor outcome for males and females, again we see that level 6 and above achievement and pre-16 attainment are the most important factors in explaining the gap between those from the highest and lowest SES groups, but for poor outcome very little of the gap is explained by pre-16 school and peer factors.

4. **Demographic and education factors only partly explain the differences in outcomes between SES quintiles**

More of the difference in good outcome between SES quintiles for females is explained by demographic and education factors than it is for males: around 80 per cent of the gap between the lowest and second lowest quintile for females is explained by these factors, compared to around three quarters of the gap between males from the same quintiles.

For both males and females, we have less of an understanding of what is driving the differences in good outcome for those in the lower socioeconomic groups, compared to those in the higher groups. As such, more of the gap between the lowest and highest SES quintiles is explained by demographic and education factors, whereas we understand less of the gap between the lowest and the second, third and fourth quintiles.

For poor outcome, the amount of the gap between SES quintiles which is explained by demographic and education factors is similar for males and females, with around 80 per cent of the gap between the lowest and second lowest quintile for females explained by these factors. More of the gap between the first and second (the two lowest SES groups) quintiles is explained by these factors than the gaps between the first and the third, fourth and highest quintiles.

For males, similar proportions of the gaps between the lowest and all higher SES quintiles are explained by the demographic and education factors (76 to 78 per cent).

## Methodology

The differences in the composition of the socioeconomic quintiles regarding demographic and education variables mean that it is difficult to compare fairly the labour market outcomes of those from different socioeconomic backgrounds. This chapter uses probit regression methods to hold the demographic and education factors associated with labour market outcomes constant so we can compare socioeconomic groups on a like for like basis to determine how much of the differences in labour market outcomes can be explained by these factors. The regression results show the chance of good or poor labour market outcome for each socioeconomic group compared to the lowest i.e. the gap, or difference, in outcomes between the most disadvantaged and the higher SES groups.

Decomposition analysis is then used to apportion the explanatory power of each demographic and education factor in explaining the gaps in good (and poor) labour market outcomes between males and females from the highest and lowest SES quintiles to shed light on the most important factor or factors.

## Conclusions

The descriptive analysis in this chapter tells us that the lower an individual's socioeconomic status during GCSEs, the less likely they are to be an upper quartile earner in their early career (good outcome), and the more likely they are to be claiming out of work benefits for 6 months in this same period (poor outcome). This persists across all demographic and education breakdowns. The regression analysis then shows that although a large proportion of the difference in good (and poor) outcome is explained by differences in demographic, school and education factors between the socioeconomic groups (for example, those from the lowest SES group are more likely to be identified with SEN, less likely to have a degree and more likely to have gone to school in the north of England than those from each higher SES group), differences still remain. In other words, once the differences in the demographics, pre- and post-16 education and achievements are taken into account, those from the highest SES quintile are still more likely to be in a good outcome, and less likely to be in a poor outcome than those from the lower SES quintiles.

Whilst this analysis offers some insight into the drivers of labour market outcomes between different socioeconomic groups, the demographic and education factors in the administrative data do not fully explain these. Further work would be required to shed light on what these factors are and their relative importance, by linking to other datasets or including labour market data as controls.

# Introduction

This chapter is one of a series of three exploring the early labour market outcomes of individuals from different backgrounds and how these outcomes differ across these groups. This chapter focuses on socioeconomic status (SES); the others look at ethnicity and special educational needs (SEN).

The Department for Education (DfE) previously published analysis[1] documenting the differences in post-16 education and labour market outcomes for a number of different socioeconomic, demographic and education sub-groups. This new series uses more sophisticated methods to build on the analysis for some of these groups in combination with other factors. Specifically, the analysis in this chapter aims to answer the following questions:

- How is socioeconomic status linked with different early career labour market outcomes when demographic and education factors are taken into account?
- Which of these demographic and education factors are most important for explaining differences in early career labour market outcomes between those from different socioeconomic backgrounds?

The research uses the Longitudinal Education Outcomes (LEO) administrative data set which links information about individuals, including:

- personal characteristics such as gender, ethnic group, special educational needs, free school meals eligibility
- education: including schools, colleges and higher education institutions attended, courses taken and qualifications achieved
- income and employment status
- claims for out of work benefits

By combining these sources, we can look at the progress of individuals doing their GCSEs into post-compulsory education and the labour market. Further information on the LEO dataset can be found in the accompanying technical report, which includes information on the data quality and match rates.

This chapter is split into four sections:

1. discussion on socioeconomic status and how it is used in the literature and in this report (Section 1)

[1] Post-16 education and labour market activities, pathways and outcomes (LEO) - GOV.UK (www.gov.uk). The accompanying dashboard is available here Longitudinal Education Outcomes (LEO): post-16 education and labour market activities and outcomes (shinyapps.io)

2. context on composition of SES groups with regard to demographic and education factors, and descriptive analysis on labour market outcomes by socioeconomic status showing the association with demographic and education factors (Section 2)
3. regression analysis to control for demographic and education factors to determine how much these account for differences in labour market outcomes for those from different socioeconomic backgrounds (Section 3)
4. analysis to shed light on the relative importance of these demographic and education factors when comparing labour market outcomes of those from the highest and lowest socioeconomic groups (Section 4)

The results in all chapters of this report are concerned with statistical measurement of the relationship between socioeconomic, demographic and education factors and measures of disadvantage and do not imply any causality.

## Coverage

The analysis in this chapter looks at 4.5 million individuals who finished key stage 4 (KS4), i.e. took their GCSEs, in a state-funded[2] school in England between the 2001/02 to 2008/09 academic years. Their labour market outcomes are measured in the 2017-18 tax year[3]. For the oldest cohort, we therefore look at outcomes 15 years after GCSEs (age 31), and for the youngest, 8 years after GCSEs (age 24). See the accompanying technical report for further details.

All labour market figures are based on UK tax and out of work benefits[4] records only, further education data is from English institutions only, and higher education figures are from UK institutions.

The eight cohorts of individuals who completed KS4 in England between 2002 and 2009 have been combined to produce a more representative and robust picture of people's labour market outcomes. Combining several cohorts of individuals completing their GCSEs at the same age means any changes or patterns are more likely to be real differences and not reflective of variations between year groups. Although this means

[2] State-funded schools in those academic years were: sponsor-led academies; city technology colleges; LA maintained mainstream; and LA maintained special schools.
[3] Although more recent employment, earnings and benefits data was available, 2017-18 tax year data was used for consistency with Post-16 education and labour market activities, pathways and outcomes (LEO) - GOV.UK (www.gov.uk).
[4] Northern Ireland benefits system is not covered by DWP. Although the benefits have the same criteria and payments, we do not have benefits claims for Northern Ireland.

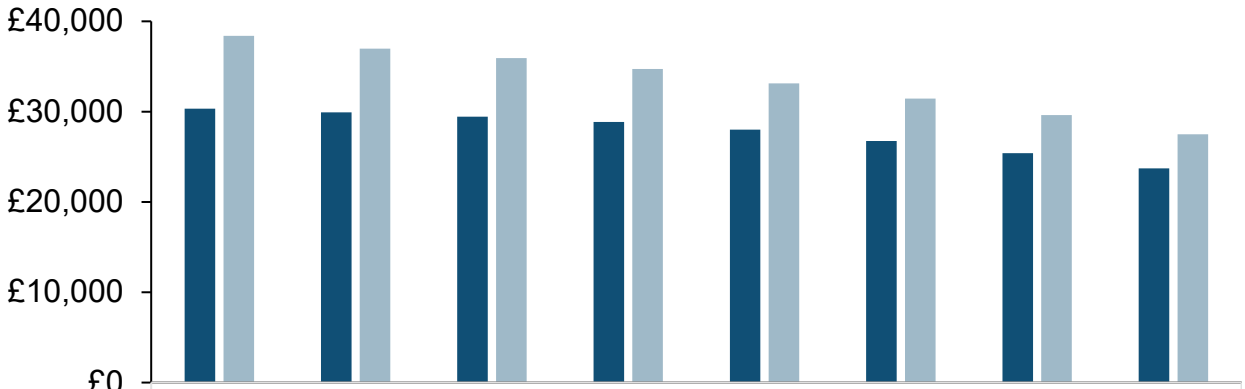
they are different ages when their outcomes are measured[5], this maximises the number of years in the labour market available for each cohort.

## Definitions

Two different measures of labour market outcomes have been defined for the analysis in this report:

**Good labour market outcome** – the individual was in paid[6] employment for at least one day in each of the 12 months of the 2017-18 tax year **and** had upper quartile earnings. Earnings quartiles have been calculated separately for males and females[7] and for each GCSE cohort (to allow for different earnings profiles at different ages). The upper quartile earnings thresholds are shown in Figure 1. Around 60 per cent of males and females in each cohort meet the employment threshold above. Taking the top quartile earners means that those in good outcome represent around 15 per cent of each cohort and gender.

**Figure 1: Upper quartile earnings thresholds in 2017-18 tax year: females and males by KS4 cohort**




| | 2001/02 | 2002/03 | 2003/04 | 2004/05 | 2005/06 | 2006/07 | 2007/08 | 2008/09 |
|---|---|---|---|---|---|---|---|---|
| Females | £30,333 | £29,915 | £29,445 | £28,866 | £27,999 | £26,744 | £25,399 | £23,722 |
| Males | £38,400 | £36,984 | £35,928 | £34,731 | £33,138 | £31,444 | £29,627 | £27,494 |

Source: Authors' analysis using Longitudinal Education Outcomes data

**Poor labour market outcome** – the individual was claiming out-of-work benefits for at least one day in each of six or more consecutive months of the 2017-18 tax year. Details

[5] Robustness checks performed on all eight cohorts at age 24, and three cohorts at age 29 produce very similar results to those produced using the 2017-18 tax year. See technical report for further information.
[6] Excludes earnings from self-employment.
[7] Information on hours worked is not available. Female earnings are likely to be affected by lower earnings due to part-time working.

on the types of benefits included can be found in the technical report. This represents around 8 per cent of the males and around 12 per cent of the females in each cohort.

The good outcome and poor outcome definitions are such that each outcome is mutually exclusive, where an individual cannot appear in multiple categories. However, the majority of individuals (around 75 per cent) sit within neither category.

These measures are used to categorise groups of individuals into those who are observed to have labour market outcomes which could be considered good in an **economic** sense (the individual has steady employment, is well paid and contributing to the exchequer) and those who may be reliant on the state with a low income, and therefore in a poor economic outcome.

This does not intend to negate any individual's labour market choices, contribution to society or personal well-being, or whose economic contribution may be lower due to no fault of their own.

**Socioeconomic status quintile** - socioeconomic status during GCSE year. This is an indicator derived from an individual's free school meals (FSM) eligibility, combined with local area statistics (deprivation, occupation, education and housing tenure) to give a combined household income and place-based measure. The index is then divided into quintiles. The lowest SES group (1st quintile) is the most disadvantaged fifth of the population and the highest SES group (5th quintile) is the least disadvantaged.

The demographic and education variables used in the descriptive analysis and as controls in the regression analysis in this chapter include:

- pupil characteristics as collected in the GCSE year[8]: ethnicity, first language, special educational needs (SEN) status
- school measures: type of school attended in the GCSE year, demographics (proportion of pupils eligible for FSM and with SEN), cohort attainment and school effectiveness (progress measure)
- school attainment: maths and English at key stage 2, key stage 4 performance points, A level total points and subjects studied
- higher education: classification of degree, subject studied, type of institution
- further education: subject of apprenticeship or classroom learning

[8] Some missing characteristics have been backfilled using earlier School Censuses, the Individualised Learning Record (ILR) and Higher Education Statistics Agency (HESA) data. See technical report for more details.

- highest level of achievement[9] by tax year 2017-18
- local authority of residence in GCSE year and during tax year 2017-18

For further information on the sources and derivation of these variables, please see the technical report.

[9] What qualification levels mean: England, Wales and Northern Ireland - GOV.UK (www.gov.uk)

# Section 1: Review of socioeconomic status in the literature

The analysis in this chapter looks at differences in a number of education factors in combination with demographics to explore how these relate to different labour market outcomes across the range of socioeconomic backgrounds. There are large differences in these factors between socioeconomic groups and there is a wealth of statistics and literature available which describe and explore them[10]. However, there is a variety of ways of describing socioeconomic status depending on the data being used and at what stage in life status is being measured. This section explores what socioeconomic status is and different proxies used to define it before outlining the measure used in this chapter and the reasoning behind it.

Socioeconomic status is a combination of economic and sociological measures regarding occupation, access to economic resources and relative social position. This can be measured for either the individual or a family depending on who or what is being examined. In education research, it is usually the socioeconomic status of the parents or of the household in which the individual grew up.

The main factors in determining socioeconomic status for education research and statistics would ideally be parental income, parental level of education and parental occupation, which together can give an indication of the economic circumstances of an individual's family during their childhood, as well as a proxy for the social class of the family. Unfortunately, some or all of these variables are seldom available, particularly in administrative data, as this is often limited to information on the student themselves. Other variables must be used as proxies for these, some of which are at the local area level and some are household or individual measures.

The Office for National Statistics (ONS) has published research into the factors involved in earnings for those from poor backgrounds[11] using LEO data. FSM eligibility is used as a proxy for household socioeconomic status during childhood. This is a binary measure which captures those families in the most economic need, but arguably many of those who could be classed as lower socioeconomic status miss out due to the eligibility criteria.

Higher education institutions use local area measures as proxies for socioeconomic status which track participation in higher education in order to attempt to target access initiatives, such as POLAR or TUNDRA[12]. These identify potential students who live in

[10] For example, the Social Mobility Commission has commissioned a large body of research and reports relating to socioeconomic status and social mobility. The commission has recently added more measures to its Social Mobility Index which can now be explored using its Data Explorer Tool
[11] Why free school meal recipients earn less than their peers - Office for National Statistics (ons.gov.uk)
[12] About the TUNDRA area-based measures data - Office for Students

areas of low participation but the lack of household measures can result in mistargeting of, for example, parents with higher education levels living in more deprived areas.

ONS publish data on young people in the labour market by socioeconomic background of young people[13] using the National Statistics Socioeconomic classification (NS-SEC), which is based on the occupation of their highest earning parent (or family member living in the household) when they were 14 years old. While this is a good measure of socioeconomic status, it is taken from survey data and data items of this nature are not generally available in combination with administrative education data.

In order to address some of these issues, a measure of socioeconomic status combining local area and household factors is being increasingly used in education research using administrative data[14][15][16] and is the one used in this report. This combines FSM eligibility and IMD[17] (as proxies for household income) and local areas statistics on occupation, education level and housing tenure (as proxies for parental social status).

## Contribution of this series

The analysis presented in the three chapters of this report uses LEO administrative data. The completeness and size of this dataset allows for robust analysis, particularly for smaller characteristics groups that cannot be obtained using survey data. The Institute for Fiscal Studies (IFS) has previously used LEO data to show returns for graduates, but no studies cover the whole range of post-16 education and the relationship with earnings and employment for ethnic groups. In this analysis, labour market outcomes data has been linked to DfE's administrative data to utilise an unprecedented range of socioeconomic and demographic factors during GCSEs as well as prior attainment, achievements at age 16 and post-16. Individuals are tracked through from the end of compulsory education rather than examining the population as a whole. In addition, novel analysis of the relationship between education, socioeconomic and demographic factors and out-of-work benefits is presented. Together the three chapters (ethnicity, socioeconomic status and special educational needs) provide extensive insight into the outcomes across different aspects of disadvantage.

[13] Young people in the labour market by socio-economic background, UK - Office for National Statistics (ons.gov.uk)
[14] Widening participation in higher education: analysis using linked administrative data - Chowdry - 2013 - Journal of the Royal Statistical Society: Series A (Statistics in Society) - Wiley Online Library
[15] The long shadow of deprivation: differences in opportunities - GOV.UK (www.gov.uk)
[16] The returns to undergraduate degrees by socio-economic group and ethnicity | Institute for Fiscal Studies (ifs.org.uk)
[17] English indices of deprivation - GOV.UK (www.gov.uk)

# Section 2: Descriptive analysis of socioeconomic status

Section 1 discusses the nature of socioeconomic status and how it is used in the literature and that the variety of different socioeconomic measures makes it difficult to consolidate information from multiple sources. To address this, we first present descriptive statistics on the demographic and educational achievements and participation of the socioeconomic status quintiles used in this chapter to provide contextual information on the cohorts in the analysis. We then describe how different labour market outcomes are related to some of these attributes for the socioeconomic status measure used. See the Coverage section for more information on the individuals included and the Definitions section for details of the definitions and derivation of the data provided here.

As a reminder, the socioeconomic status measure is an index split into quintiles (SES quintile) derived from an individual's free school meals (FSM) eligibility, combined with local area statistics (deprivation, occupation, education and housing tenure). The 1st quintile is the lowest SES (or most deprived) group, and the 5th quintile is the highest SES (or least deprived) group.

## A: Socioeconomic quintiles in this report

The demographic and education composition of the socioeconomic groups used in this chapter are not readily available from published sources, so this section includes some breakdowns to give a picture of the different makeup of the quintiles of socioeconomic status, to give context to the differences observed in the labour market.

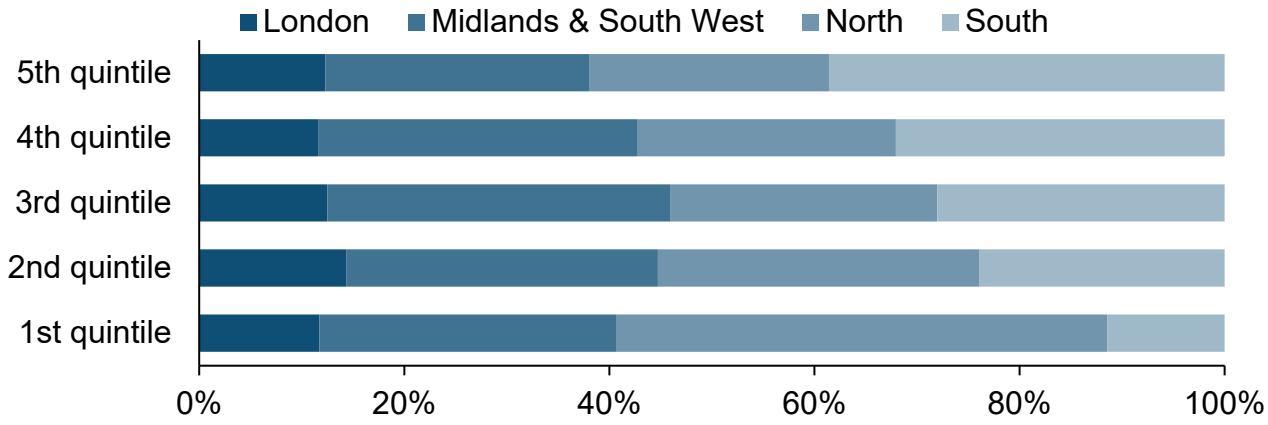
### Methodology

The percentage of individuals in each SES quintile in each of the factors listed in the Definitions section has been calculated. A selection of these is shown below.

### Location

The administrative data used in this report allows us to see which region of England the school attended during GCSEs was located in. Regions have been grouped as follows: London, Midlands & South West (East Midlands, West Midlands and South West), North (Yorkshire and the Humber, North East and North West), and South (South East and East of England).

**Figure 2: Proportion of each SES quintile by region of school**



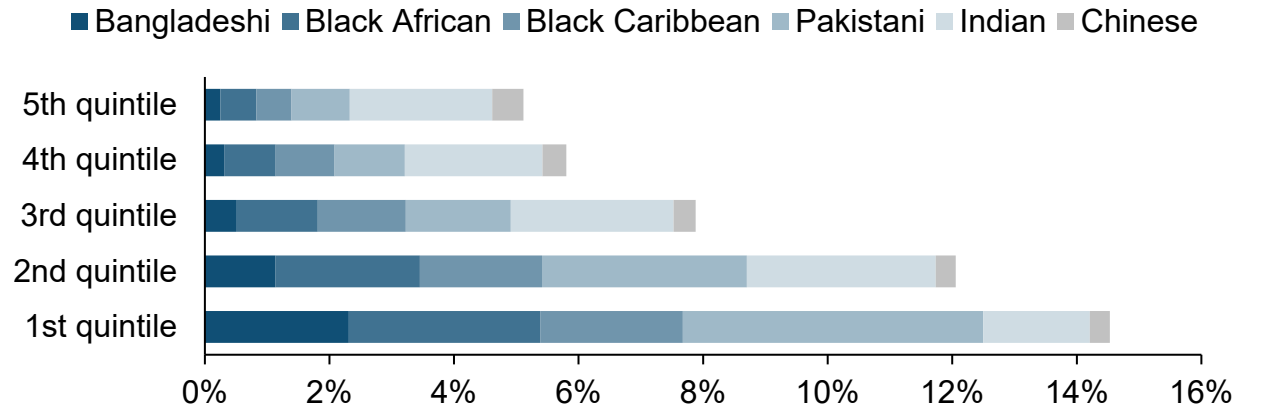

Source: Authors' analysis using Longitudinal Education Outcomes data

Very similar proportions from each quintile attended a school in London (Figure 2: 14 per cent for the 2nd quintile, 12 per cent for all others). However, this is not the case for the rest of the South, making up only 11 per cent of those from the lowest SES quintiles, but 39 per cent of the highest SES quintile. In total, around half of the highest SES quintile went to school in the South or London; in contrast almost half of those from the lowest SES quintile went to school in the North.

## Ethnicity

The ethnic composition of each socioeconomic quintile is different. Figure 3 shows the proportion of individuals in each quintile by ethnic group. For clarity of presentation, white British is omitted (as the majority ethnic group) and only the other six ethnic groups used in the accompanying ethnicity chapter are shown.

**Figure 3: Proportion of each SES quintile by ethnic group (selected)**



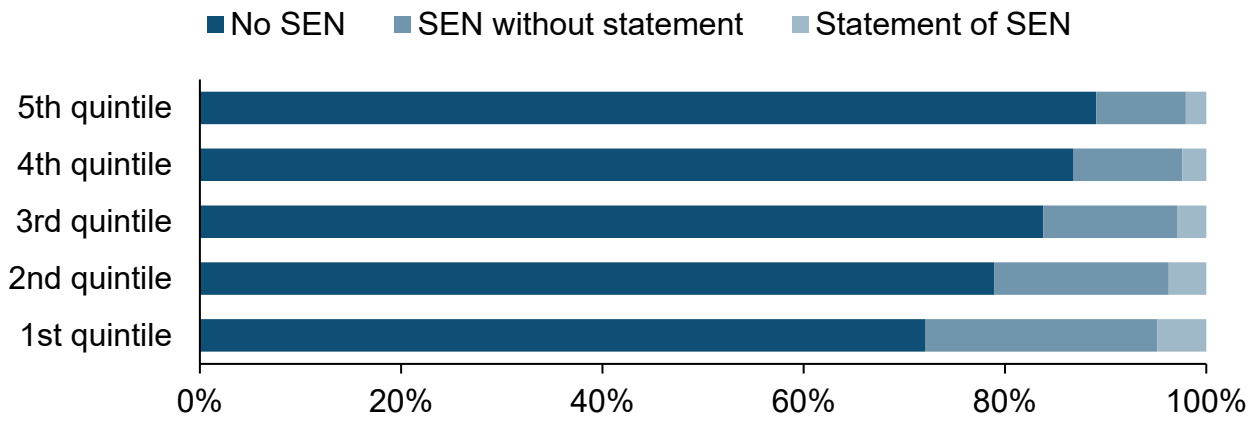

Source: Authors' analysis using Longitudinal Education Outcomes data

The lowest SES quintile has the largest proportion of minority ethnic groups in total, and the highest SES quintile has the lowest proportion. However, looking at individual ethnicities, we see that the higher SES quintiles have larger proportions of those from the Indian and Chinese ethnic groups, and the lower quintiles have larger proportions of those from the Bangladeshi, Pakistani, black African and black Caribbean ethnic groups.

## Special educational needs

The analysis on special educational needs in this report uses the SEN Code of Practice[18] which came into effect on 1 January 2002, before the introduction of Education, Health and Care (EHC) plans. Under this Code of Practice, a child or young person could be identified in one of three categories: statement of SEN, School Action or School Action Plus. A statement of SEN is when a formal assessment has been made which sets out the child's need and the extra help they should receive. For this analysis, the SEN categories School Action and School Action Plus are combined into 'SEN without statement'. Figure 4 shows the proportions of each SEN category in each SES quintile.

**Figure 4: Proportion of each SES quintile by SEN**




Source: Authors' analysis using Longitudinal Education Outcomes data

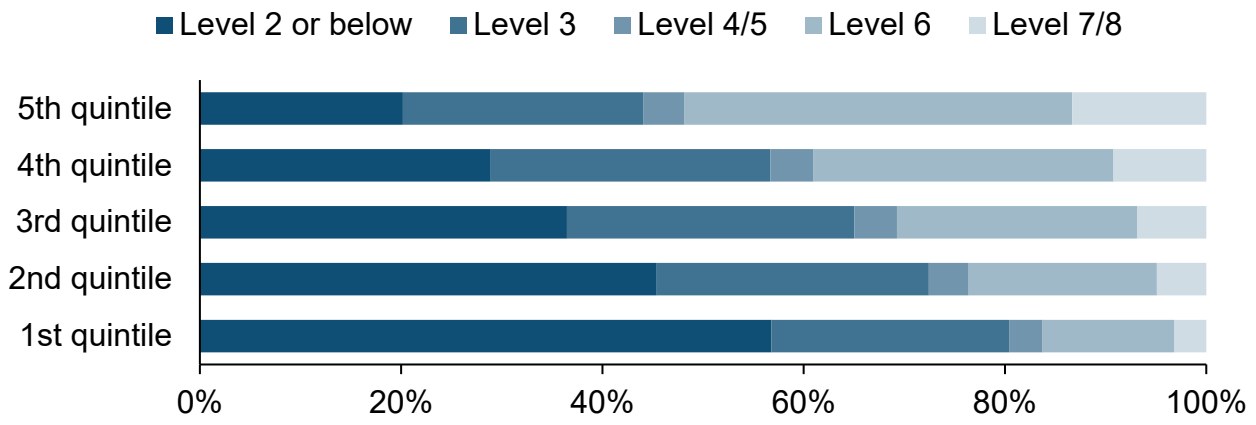
Special educational needs is strongly associated with socioeconomic status: the lower the SES quintile, the higher the percentage of individuals in that quintile who are identified with SEN. Almost a quarter of those from the lowest SES quintile are identified with SEN without a statement, with a further 5 per cent having a statement of SEN, whereas less than a tenth (9 per cent) of those from the highest SES quintile have SEN without a statement and only 2 per cent have a statement of SEN.

18 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/273877/special_educational_needs_code_of_practice.pdf

## Highest level of education

The highest level of education is measured in the same year as labour market outcomes i.e. it is the cumulative highest level of achievement by the age of 24 to 31. The proportions of each SES quintile achieving each level is shown in Figure 5. Level 2 is the equivalent of five GCSEs at A*-C, level 3 is the equivalent of two A levels, and level 6 is a first degree.

**Figure 5: Proportion of each SES quintile by highest level of education**



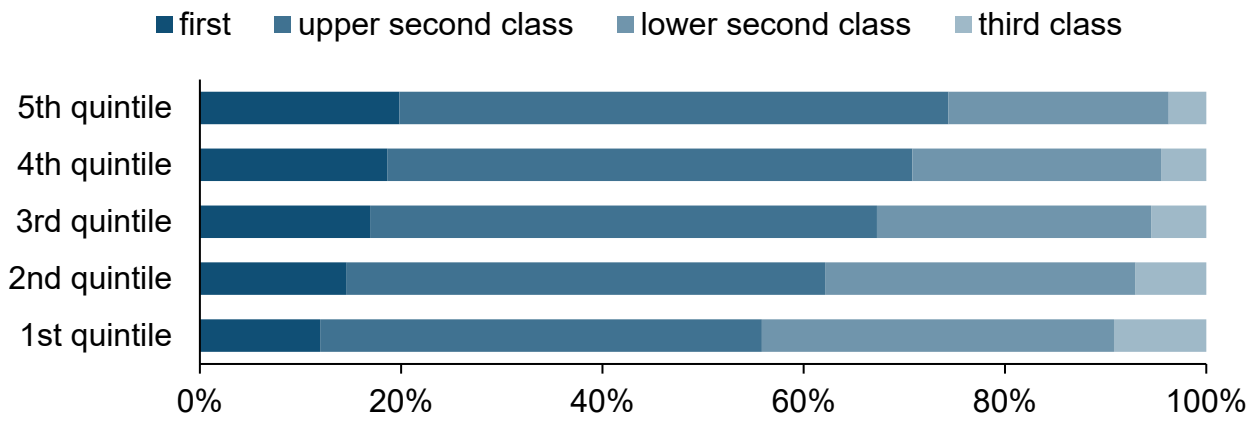

Source: Authors' analysis using Longitudinal Education Outcomes data

The attainment profile of each SES quintile is quite different, with the majority (57 per cent) of those from the lowest SES background only achieving level 2 or below and only 16 per cent achieving degree level or higher. Each increase in quintile from the lowest sees higher levels of education, with smaller proportions with level 2 or below as the highest level, and higher proportions with degree level achievements. Just over half (52 per cent) of those from the highest SES quintile are educated to at least degree level and only around a fifth with level 2 or below. All SES quintiles have similar level 4 and 5 achievements at around 3 to 4 percent of the cohort.

## Higher education: degree classification

The classification of first degree awarded to those educated to at least degree level is shown in Figure 6.

**Figure 6: Proportion of each SES quintile by degree classification**



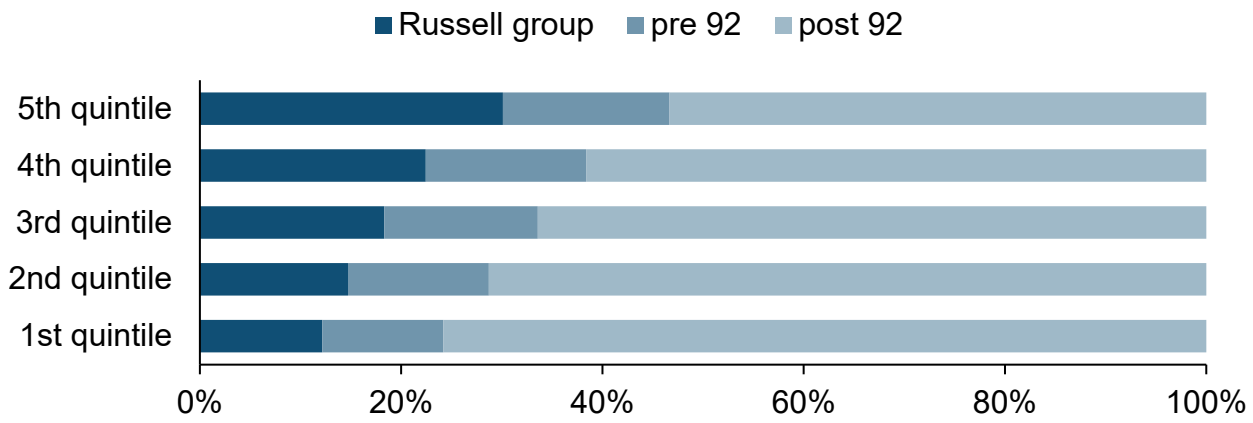

Source: Authors' analysis using Longitudinal Education Outcomes data

As well as different proportions achieving a degree for each SES quintile, the distribution of degree classifications awarded also differ. For the highest socioeconomic group, 17 per cent were awarded a first class degree and 32 per cent were awarded a lower second or third class. With each lower SES quintile, the proportions achieving a first or upper second class decrease and the proportions achieving a lower second or third class increase. Around 12 per cent of the lowest SES quintile were awarded a first class degree, and 44 per cent a lower second or third class. This means that not only are those from lower socioeconomic backgrounds less likely to get a degree but are also less likely achieve the best results.

## Higher education: type of institution

Those from lower socioeconomic backgrounds are less likely to participate in higher education and less likely to achieve as highly as those from higher SES quintiles, but they are also less likely to attend more prestigious institutions (Figure 7).

**Figure 7: Proportion of each SES quintile by type of institution (first degree only)**




Source: Authors' analysis using Longitudinal Education Outcomes data

Around three-quarters (76 per cent) of those from the lowest SES background received their first degree from post 92 institutions, compared to just over half (53 per cent) of those from the highest SES quintile. Over a quarter (30 per cent) of those from the highest SES quintile went to Russell group institutions but only 12 per cent of those from the lowest. Each increase in quintile results in an increase in proportions attending Russell group or pre 92 institutions and a decrease in attending post 92.

## B: Observed labour market differences

The section above gives some insight into the distribution of various demographic and education factors across the quintiles of socioeconomic status used in the chapter. For example, in Figure 5 we saw that far fewer individuals from the lowest SES quintile achieved at least degree level. This section describes the proportion of each quintile in good or poor labour market outcomes broken down by each of these demographic and education factors. This tells us whether or not those with the same characteristics from different quintiles have the same labour market outcomes.

### Methodology

The percentage of individuals in each SES quintile in the good labour market outcome category has been calculated for males, females, and males and females combined for each of the factors listed in the Definitions section. A selection of these has been presented in this chapter.

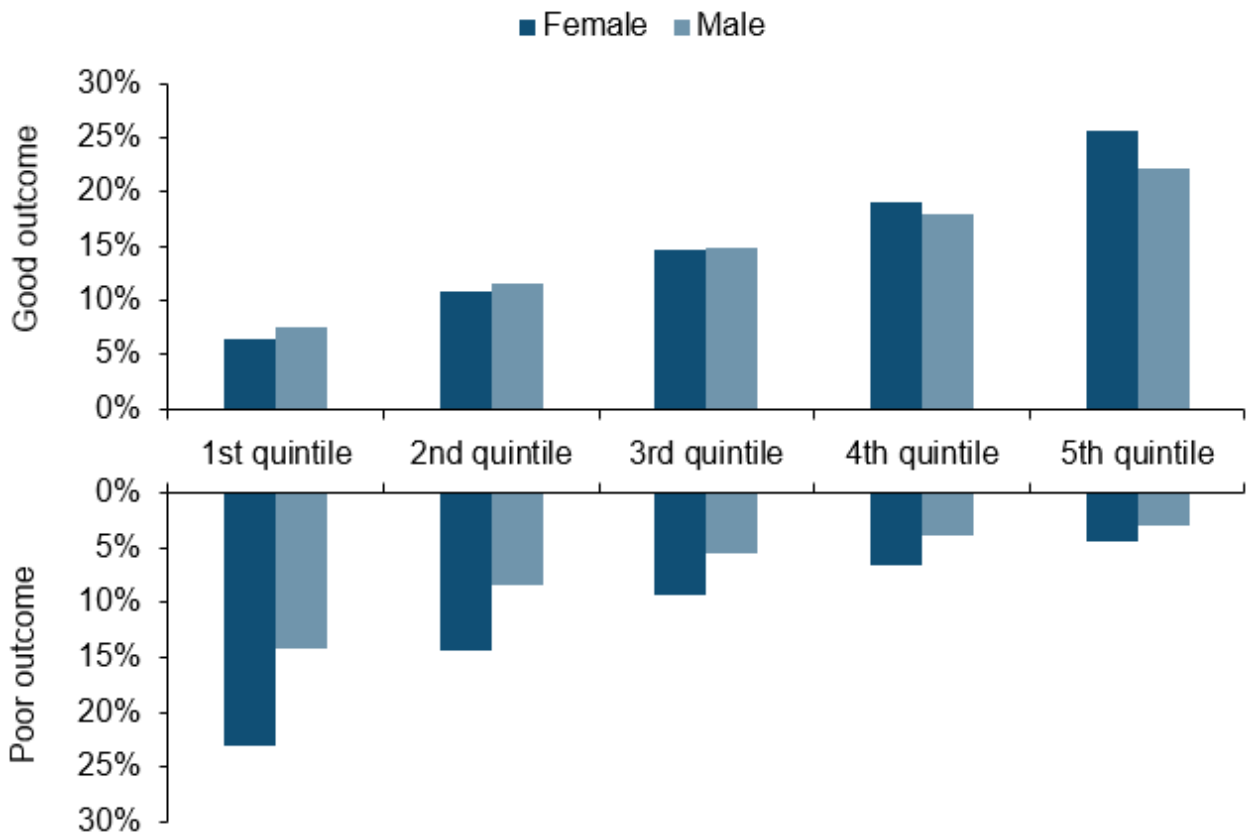
This process has been repeated for poor labour market outcome.

A good outcome has been defined as being in sustained employment and an upper quartile earner. A poor outcome has been defined as claiming out-of-work benefits for at least a six month period (see Definitions for more detail).

## Gender

Figure 8 shows the proportion of male and females in each socioeconomic quintile who are in the good outcome group (top half of the chart), and the proportion in the poor outcome group (bottom half of the chart).

**Figure 8: Proportions in Good Outcome and in Poor Outcome by SES quintile and gender**




Source: Authors' analysis using Longitudinal Education Outcomes data

For both men and women, those from higher SES backgrounds are more likely to be in a **good outcome** than those from lower SES backgrounds, with a gradual progression in likelihood from the lowest to the highest SES quintile. For the highest SES quintile, females are more likely to be in a good outcome than males, but the opposite is true for the lowest SES quintile. This means that females have a larger disparity across socioeconomic status with those from the highest SES quintile being more than 4 times more likely to be in a good outcome than those from the lowest SES quintile.

For both men and women, those from lower SES backgrounds are more likely to be in a **poor outcome** than those from higher SES backgrounds. The gaps between quintiles are larger at the lower SES end of the scale and for all quintiles, females are more likely to be in a poor outcome than males. For both males and females, those from the lowest SES quintile are around 5 times more likely to be in a poor outcome than the highest SES quintile.

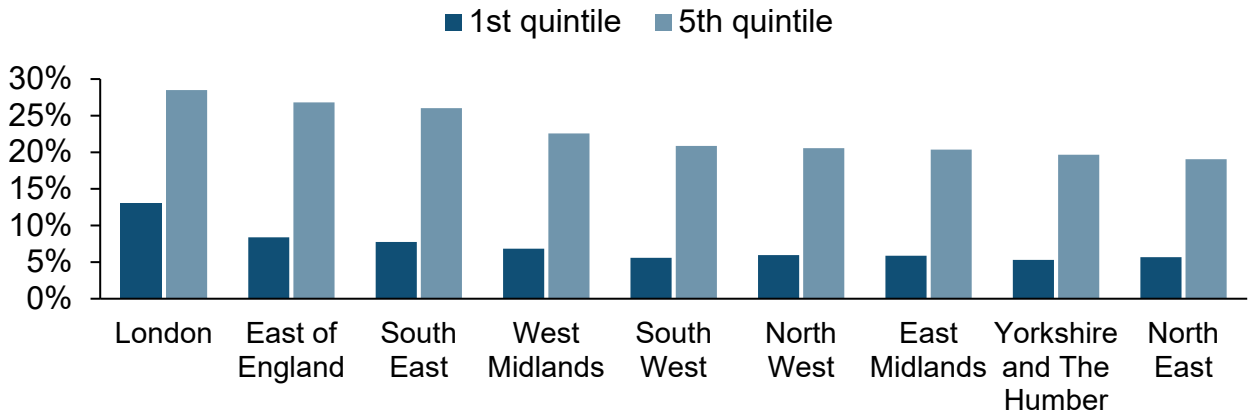
It is also interesting to note that there are more females than males in either outcome category in every quintile.

## Location

For clarity, the charts showing region of school only show the lowest (1st) and highest (5th) socioeconomic quintiles. For every region of England, there is an even increase in the proportions in good outcome from each quintile to the next. For poor outcome for all regions except London, however, there is a large drop in the proportions with this outcome from the lowest socioeconomic status to the second quintile; thereafter the gaps in poor outcome become smaller. For London, the gaps between quintiles are much more even.

Throughout this section, regions are consistently ordered according to decreasing proportions in good outcome for the highest SES quintile.

**Figure 9: Proportions in Good Outcome by SES quintile and region of school**



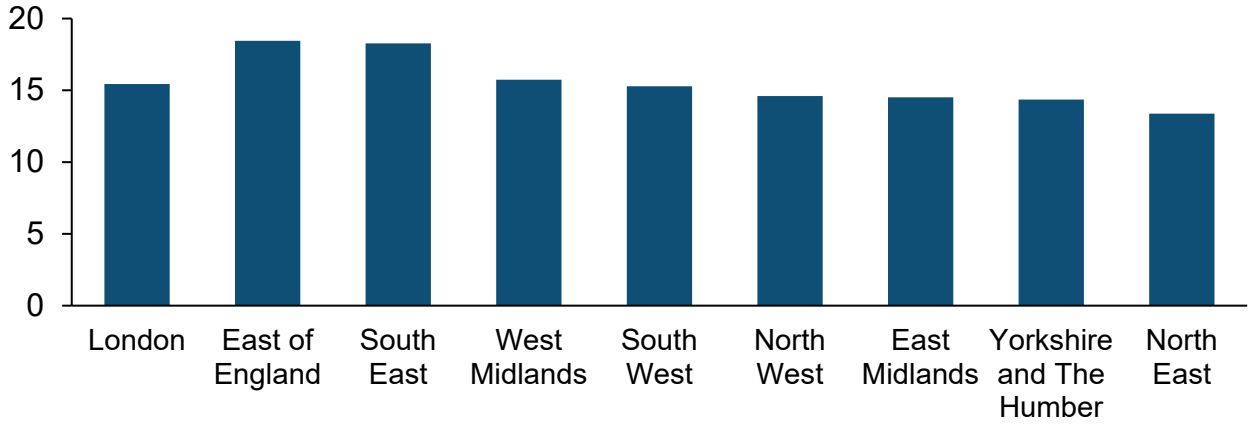

Source: Authors' analysis using Longitudinal Education Outcomes data

Within the same region, those from the highest SES quintile are far more likely to be in a good outcome than those from the lowest SES quintile (Figure 9). London has the highest proportions in good outcome for both SES quintiles, but particularly so for lowest SES quintile: in London 13 per cent of this quintile are in a good outcome and the next highest are South East and East of England at 8 per cent. The North East has one of the

lowest proportions in good outcome for the lowest SES group and also the lowest good outcome for the highest SES quintile.

The difference in proportions in good outcome between the highest and lowest quintiles in each region is shown in Figure 10.

**Figure 10: Percentage point gap in Good Outcome between the top and bottom socioeconomic quintiles by region of school**

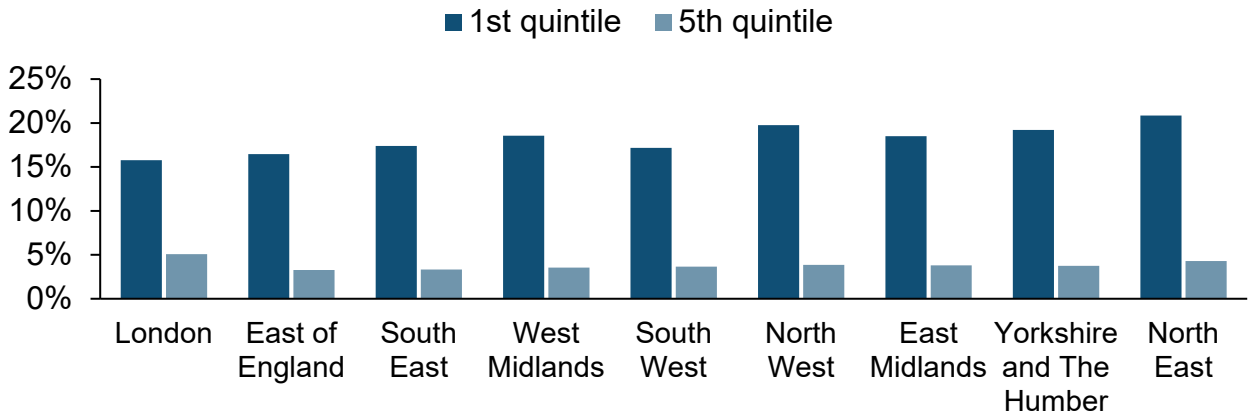



Source: Authors' analysis using Longitudinal Education Outcomes data

East of England has the highest disparity between those from the most and least deprived backgrounds, closely followed by the South East: those from the lowest SES quintile in these regions do not fare better than in other regions. The North East has the smallest disparity in good outcome between the highest and lowest SES groups; outcomes are poorer here across all socioeconomic groups.

The proportions in each region in poor outcome (Figure 11) show less variation than for good outcome. In general, however, those regions with higher good outcome also tend to have lower poor outcome.

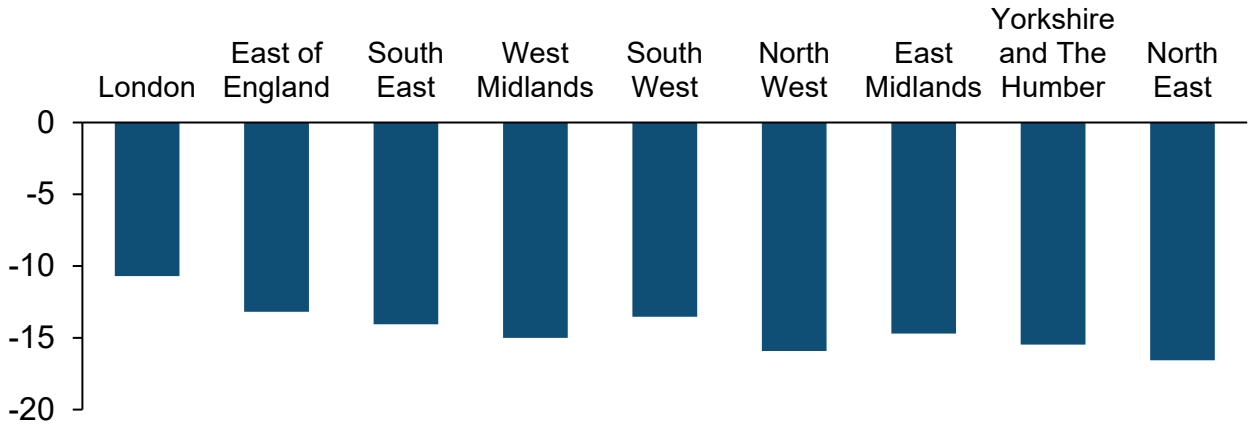
**Figure 11: Proportions in Poor Outcome by SES quintile and region of school**




Source: Authors' analysis using Longitudinal Education Outcomes data

Most regions show very similar proportions of the highest SES quintile in poor outcome, apart from London. The North East has the highest proportion of those from the lowest SES quintile in poor outcome and London has the lowest.

**Figure 12: Percentage point gap in Poor Outcome between the top and bottom socioeconomic quintiles by region of school**




Source: Authors' analysis using Longitudinal Education Outcomes data

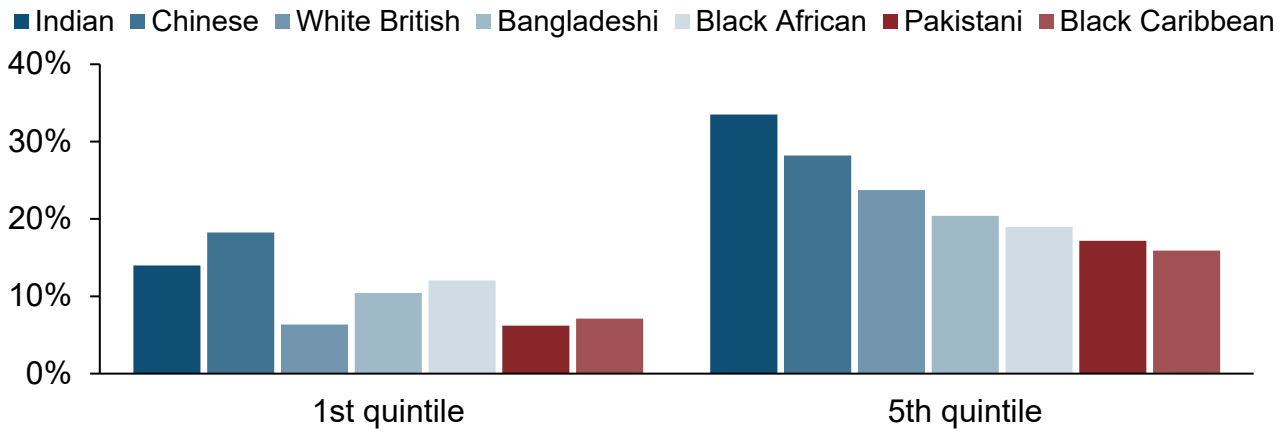
For poor outcome we also see large gaps between the highest and lowest SES groups across all regions (Figure 12). The smallest gap between quintiles is seen for London and the biggest gap in the North East.

## Ethnicity

Different ethnic groups have different chances of being in a good or poor labour market outcome and this is true within SES quintiles, although the relationship is not consistent

across ethnicities. To highlight this, the ethnic groups shown in Figure 13 and Figure 14 have been ordered by decreasing good outcome for the highest SES quintile. This section shows good and poor outcomes for the 7 ethnic groups used in the accompanying ethnicity chapter (Bangladeshi, Indian, Pakistani, black African, black Caribbean, Chinese and white British).

**Figure 13: Proportions in Good Outcome by SES quintile and ethnic group (selected)**

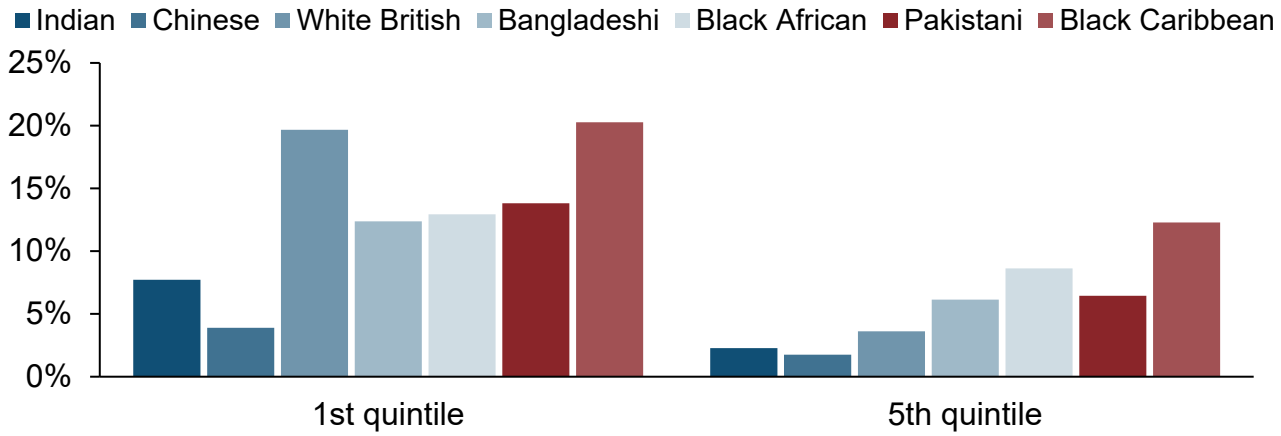



Source: Authors' analysis using Longitudinal Education Outcomes data

The Indian and Chinese ethnic groups have the highest proportions from both the highest and lowest socioeconomic groups in good outcome (Figure 13). The Pakistani and black Caribbean ethnic groups have the lowest proportions in good outcome in the highest SES group, and some of the lowest proportions in the lowest SES group (along with white British).

Ethnic groups also have different likelihoods of being in a poor labour market outcome (Figure 14) and the chance of poor outcome increases with lower socioeconomic status but to different extents for different ethnic groups.

**Figure 14: Proportions in Poor Outcome by SES quintile and ethnic group (selected)**

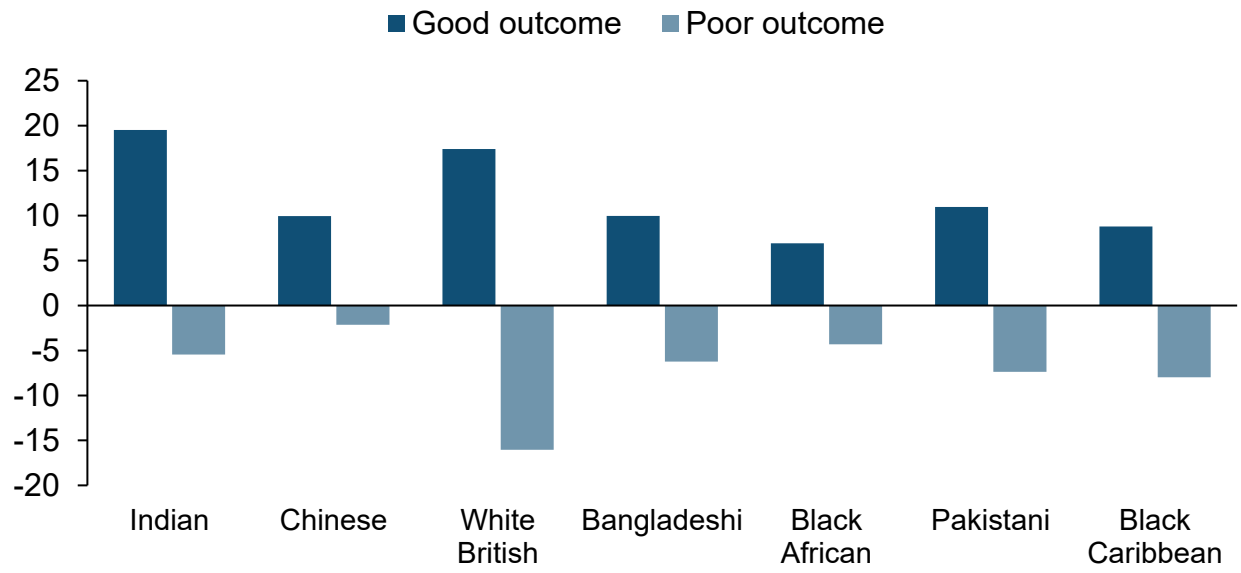



Source: Authors' analysis using Longitudinal Education Outcomes data

In general, those ethnic groups from either socioeconomic quintile which are more likely to be in good outcome are less likely to be in a poor outcome (Chinese and Indian), and those which are less likely to be a good outcome are more likely to be in a poor outcome (such as Pakistani and black Caribbean).

For all ethnic groups, there is a disparity between labour market outcomes for those from the highest and the lowest SES groups, and these gaps are shown in Figure 15 for both good and poor outcome.

**Figure 15: Percentage point gap in Good Outcome and Poor Outcome between the top and bottom socioeconomic quintiles by ethnic group (selected)**



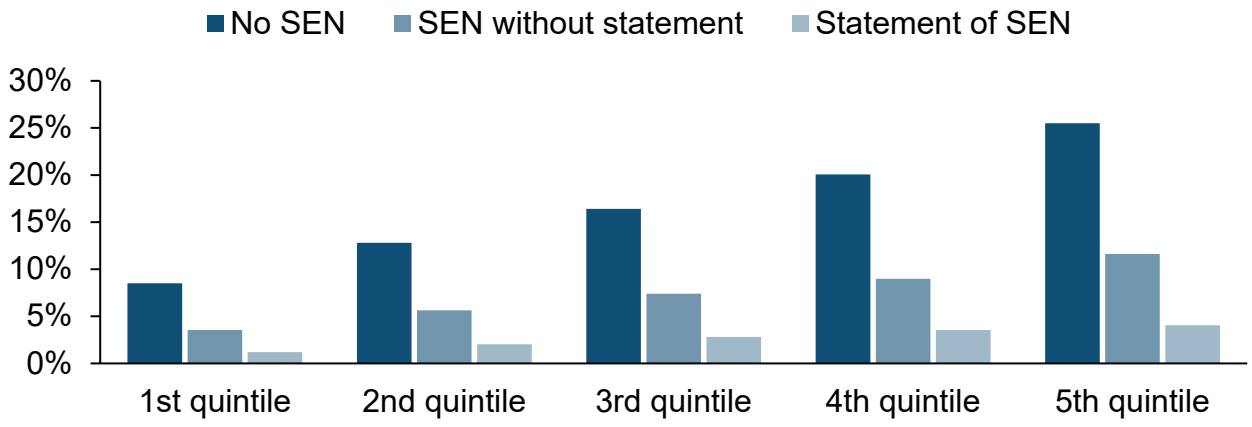

Source: Authors' analysis using Longitudinal Education Outcomes data

The largest gap in good outcome between SES backgrounds is for the Indian ethnic group despite the generally higher outcomes of this group. The white British group has the next largest disparity in good outcome, but this group also has, by far, the largest disparity in poor outcome between the most and least advantaged. Differences in good outcome between the highest and lowest socioeconomic groups are smallest for the black African ethnic group, and the Chinese group has the smallest disparity in poor outcome across SES quintiles.

## Special educational needs

The analysis on special educational needs in this report uses the SEN Code of Practice[19] which came into effect on 1 January 2002, before the introduction of Education, Health and Care (EHC) plans. Under this Code of Practice, a child or young person could be identified in one of three categories: statement of SEN, School Action or School Action Plus. A statement of SEN is when a formal assessment has been made which sets out the child's need and the extra help they should receive. For this analysis, the SEN categories School Action and School Action Plus are combined into 'SEN without statement'.

**Figure 16: Proportions in Good Outcome by SES quintile and special educational needs (SEN)**



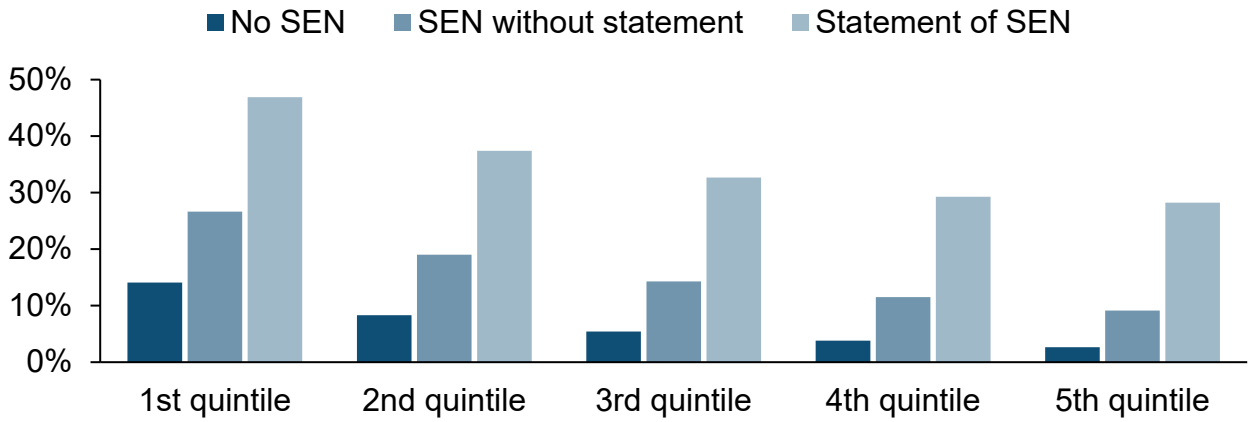

Source: Authors' analysis using Longitudinal Education Outcomes data

The proportion of individuals with SEN in a good outcome is, overall, very low particularly for those with a statement. We can see from Figure 16 that the difference in good outcome for those with special educational needs compared to those without SEN is consistent across SES quintiles. The disparity across quintiles means that those from the

19 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/273877/special_educational_needs_code_of_practice.pdf

highest SES quintile with SEN without a statement are more likely to be in a good outcome than those without SEN from the lowest SES quintile.

**Figure 17: Proportions in Poor Outcome by SES quintile and special educational needs (SEN)**



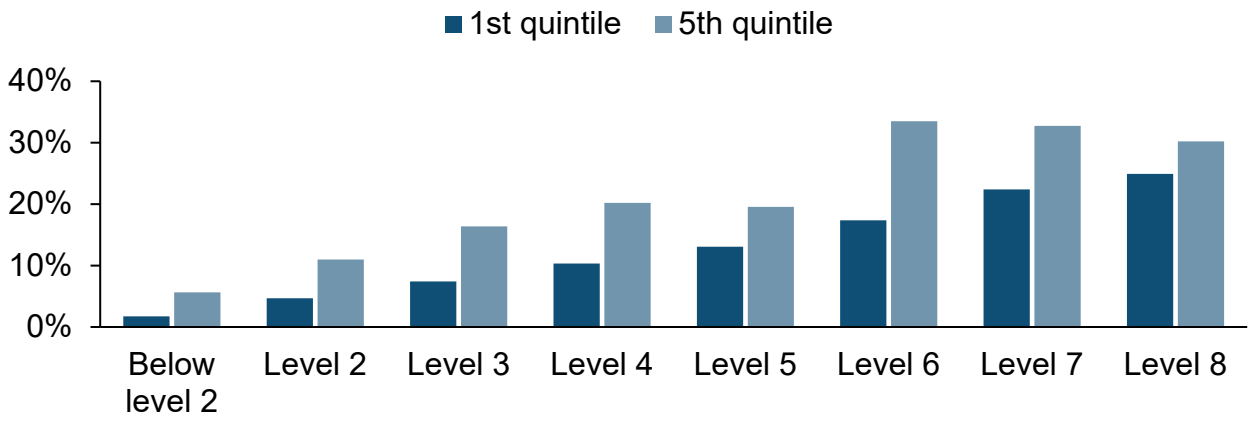

Source: Authors' analysis using Longitudinal Education Outcomes data

Having a special educational need is particularly associated with poor outcome and this is reflected across all SES quintiles (Figure 17). The disparity between those without SEN and those with SEN without a statement is relatively consistent across SES backgrounds, so that those with SEN without a statement from the highest SES quintile are less likely to be in a poor outcome that those with no SEN from the lowest SES quintile. However, the association between SES and SEN with a statement for poor outcome is less strong, although it is still more likely for those from the lowest SES quintile. This is not unexpected as the out-of-work benefits used to define poor outcome in this analysis include benefits to help those with a disability or health conditions which affect how much they can work.

## Highest level of education

For this section, the comparison is made between the highest and lowest SES quintiles only to make the trends between socioeconomic status easier to see across education levels.

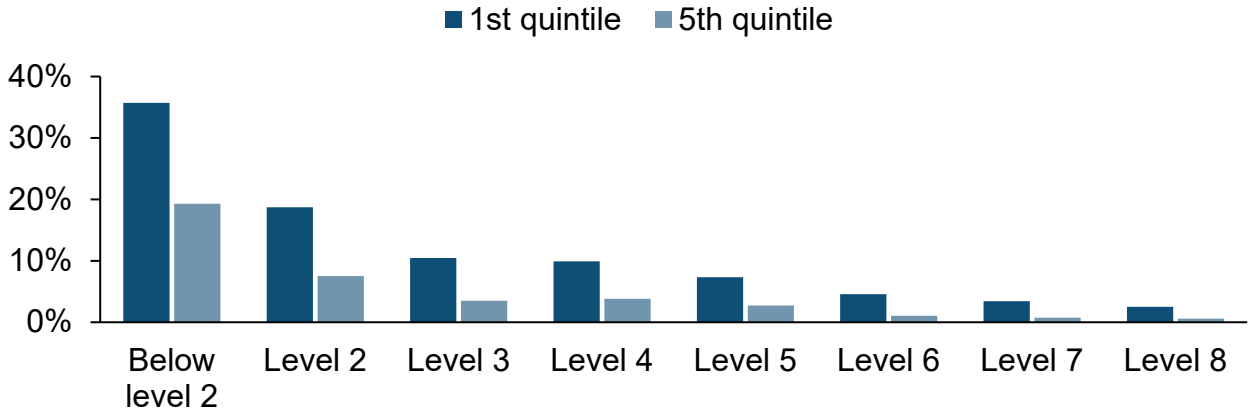
**Figure 18: Proportions in Good Outcome by SES quintile and highest level of education**




Source: Authors' analysis using Longitudinal Education Outcomes data

Figure 18 shows the relationship between good outcome and highest level[20] of education for the highest and lowest SES quintiles. At all levels of education, those from the highest SES quintile are much more likely to be in a good outcome than those from the lowest SES quintile. For example, those with a degree level qualification from the highest SES quintile are twice as likely to be in a good outcome than those with a degree from the lowest SES quintile (and they are far more likely to have a degree).

**Figure 19: Proportions in Poor Outcome by SES quintile and highest level of education**




Source: Authors' analysis using Longitudinal Education Outcomes data

Across all education levels, the chance of poor outcome is substantially higher for those from the lowest SES quintile than those from the highest. For both groups, not achieving

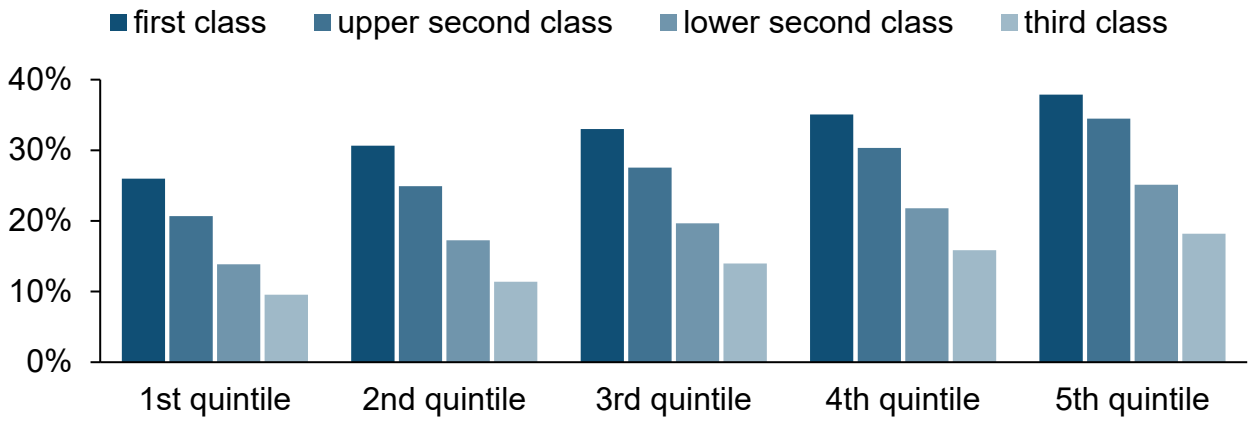
[20] What qualification levels mean: England, Wales and Northern Ireland - GOV.UK (www.gov.uk)

at least 5 A*-C GCSEs or equivalent (below level 2 achievement) is strongly associated with higher likelihood of being in a poor outcome (Figure 19), as is achieving below level 3 (equivalent to 2 A levels), but to a lesser extent. The lowest proportions claiming benefits for 6 months are seen for those achieving degree level or above.

## Higher education: degree classification

Achieving a degree is associated with a higher chance of good outcome, although Figure 18 shows us that this varies by SES. For those with degrees, a higher classification is associated with higher proportions of good outcome (see Figure 20) and this is reflected across SES quintiles.

**Figure 20: Proportions in Good Outcome by SES quintile and classification of first degree**

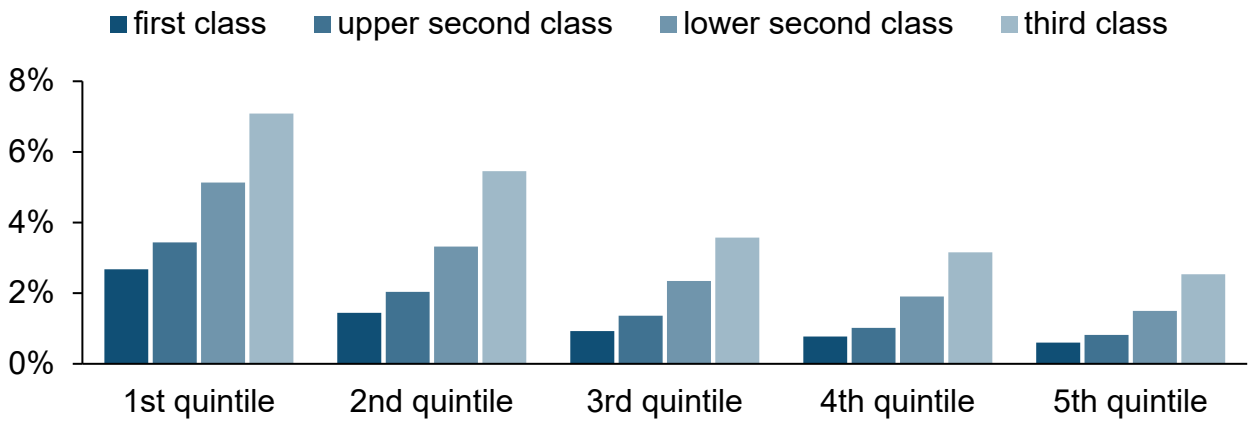



Source: Authors' analysis using Longitudinal Education Outcomes data

The disparity in good outcome across SES quintiles, however, means that those from the highest SES backgrounds with a lower second class degree are as likely to be in a good outcome as those from the lowest SES quintile with a first class degree.

Similarly, achieving a degree is associated with low chances of poor outcome, and the higher the classification of the result the lower the chance.

**Figure 21: Proportions in Poor Outcome by SES quintile and classification of first degree**



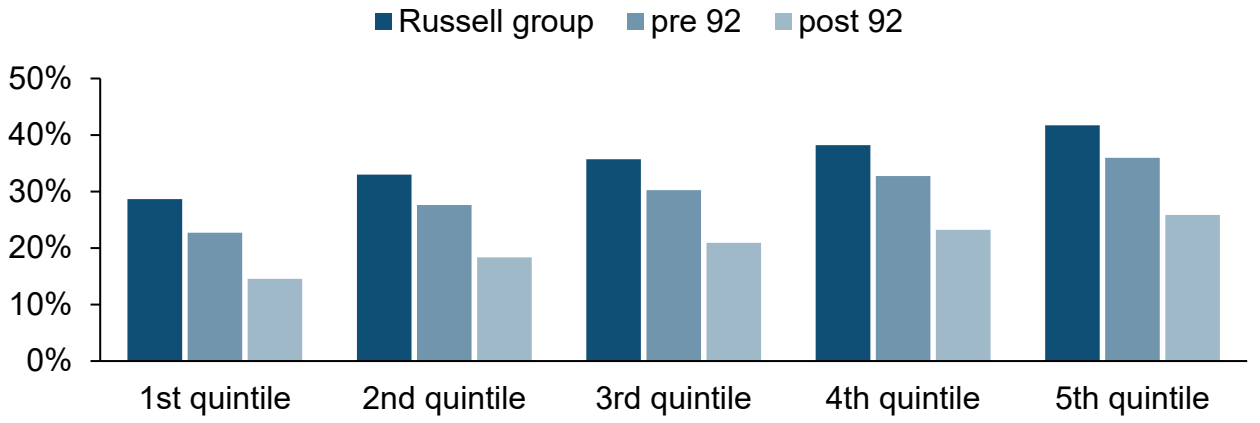

Source: Authors' analysis using Longitudinal Education Outcomes data

In Figure 21, however, we can see that those from the highest SES quintile with any class of degree are as likely or less likely to be in a poor outcome as those from the lowest SES quintile with a first class honours degree.

## Higher education: type of institution

The type of institution attended to study a first degree is also associated with different chances of good outcome. This effect is similar across all SES quintiles as seen in Figure 22.

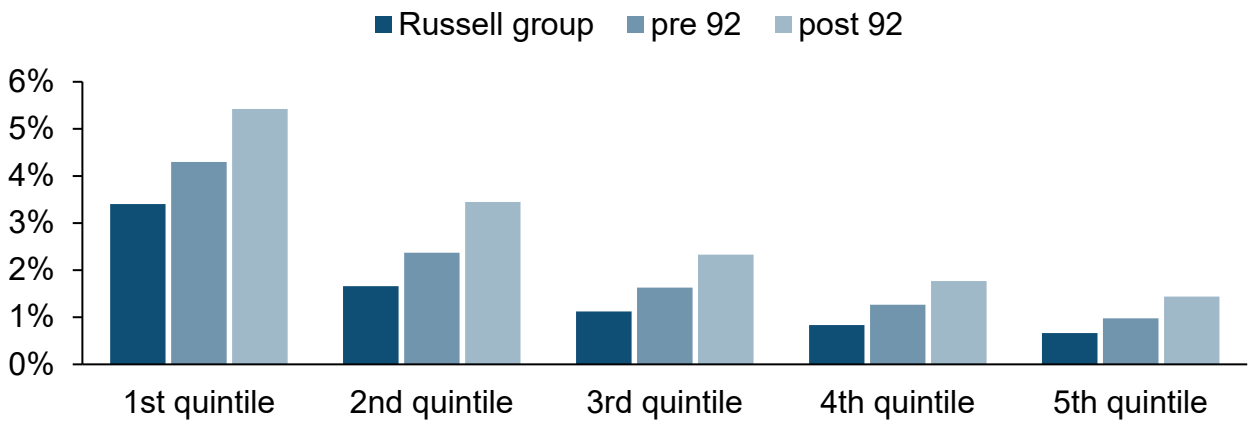
**Figure 22: Proportions in Good Outcome by SES quintile and type of higher education institution (first degree)**




Source: Authors' analysis using Longitudinal Education Outcomes data

The disparities we see across quintiles are still evident here, so that those from the lowest SES quintiles have the lowest chance of good outcome regardless of institution type attended.

**Figure 23: Proportions in Poor Outcome by SES quintile and type of higher education institution (first degree)**




Source: Authors' analysis using Longitudinal Education Outcomes data

The type of higher education institution attended is also associated with different chances of poor outcome. For all SES quintiles, those attending a Russell group institution are less likely to be in a poor outcome than those attending pre 92 or post 92 institutions. For those from the highest SES quintile, the chances of poor outcome for those with degree level qualifications are so low that the difference in poor outcome by institution type is small.

# Section 3: Effect of introducing controls

The analysis in Section 2 shows us that good labour market outcomes are more likely for those from a higher socioeconomic status, and poor labour market outcomes more likely for those from lower SES and that this is true across different education and demographic characteristics. We also saw that the proportions of these characteristics are different across quintiles; in general, those from lower SES participate less (in education) or have greater disadvantage. Regression analysis enables us to control for the factors found in the LEO administrative data to determine the extent to which the disparity between SES groups is due to the differences in these factors. This approach allows us to hold these demographic and education factors (covered in the previous section and mentioned in the Definitions section) constant at once, so we can compare the outcomes of different SES groups on a like for like basis and isolate the relationship between socioeconomic status and labour market outcomes.

## Methodology

Probit regression methods have been used to estimate the probability that an individual with particular characteristics will fall into the measured outcome category (e.g. good outcome).

A brief description of the method used is provided in this section. Full details can be found in the technical report.

Different socioeconomic groups have different labour market outcomes, but this could be driven by differences in underlying demographic and educational factors which vary between SES groups and which we know influence labour market outcomes. Regression analysis allows us to hold these factors constant so we can compare on a more like for like basis. This enables us to isolate the relationship between socioeconomic status and different labour market outcomes by calculating how much of the observed difference between quintiles is due to the factors we can observe in the administrative data, and how much cannot be explained by these factors. When referring to controls, it is a reference to these demographic and education factors that are held constant, allowing more like for like comparisons.

The dependent variable (good labour market outcome, or poor labour market outcome) is binary (an individual is either in a good outcome, or is not) so a binary regression model is used. Probit regression has been used which estimates the probability of an individual falling into the outcome category, for example having a good (or poor) labour market outcome.

The average marginal effect is then calculated: this is the average change in the probability of having a good or poor outcome compared to a baseline (or reference) group. For this chapter we measure the average difference in probability of good (or poor) outcome for each SES quintile compared to the lowest (1st quintile). For example, if the average marginal effect for good outcome for males from the highest SES quintile is 2.4, then the probability of achieving a good outcome for a male from the highest SES quintile is 2.4 percentage points higher, on average, than for a male from the lowest SES quintile.

The regression model was run twice, once without any controls and once with all demographic and education controls. The first results show the raw differences between SES groups in the labour market (i.e. the observed differences in outcomes between groups, before controlling for any other factors). The second results show the differences between SES groups after controlling for the factors in the model. These differences are what you would see if you looked at the data and accounted for a wide range of other factors that could affect labour market outcomes and differ by socioeconomic status.

## Results

The charts show the marginal effects of socioeconomic status on labour market outcomes, both with and without controls[21].

In all cases, the results for all other quintiles are significantly different from the lowest quintile (that is, the change in probability of being in the outcome group between quintiles is significantly different from zero).

Error bars show 95% confidence intervals. Confidence intervals provide an indication of the uncertainty of the estimates produced. Large intervals mean less precise estimates and smaller intervals indicate more certainty. There is a 95% chance that the true value for the population will fall between the upper and lower confidence limits. The error bars for these results are all very small, so we can be confident of the estimates produced.

Results of the regression are shown as a percentage point change in outcomes between the lowest SES quintile and each higher SES quintile. For good labour market outcome, a positive margin means those in that SES quintile are, on average, more likely to be in a good outcome than those in the lowest SES quintile. Similarly, for poor labour market outcome, a negative margin means those in that SES quintile are, on average, less likely to be in a poor outcome than those in the lowest SES quintile. The uncontrolled values are equivalent to the differences we observe in the previous section. Adding the controls

[21] see Figure 23 for an example

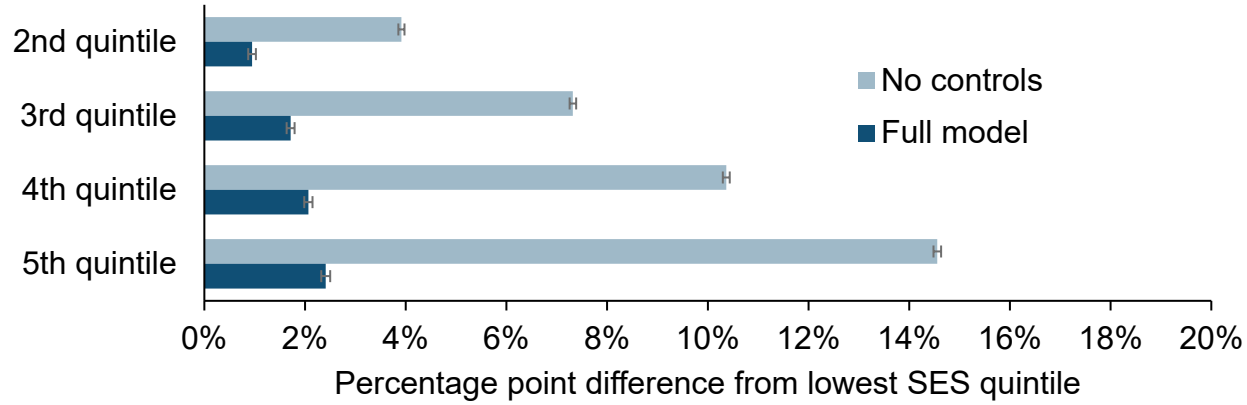
to the regression allows us to take into account the differences in level of education, subjects studied, ethnicity, geography etc between the socioeconomic groups.

## Good outcome

### Males

Results from the regression on males in each socioeconomic group on good labour market outcome are shown in Figure 24. Before controlling for demographic and education factors, we observe that males from the highest SES quintile are around 14.5 percentage points more likely to achieve a good labour market outcome than those from the lowest SES quintile. Those from the second lowest quintile are around 4 percentage points more likely, with the other quintiles evenly between these.

**Figure 24: Males - Marginal effects of SES on good labour market outcome**




Error bars represent 95% confidence intervals

Source: Authors' analysis using Longitudinal Education Outcomes data

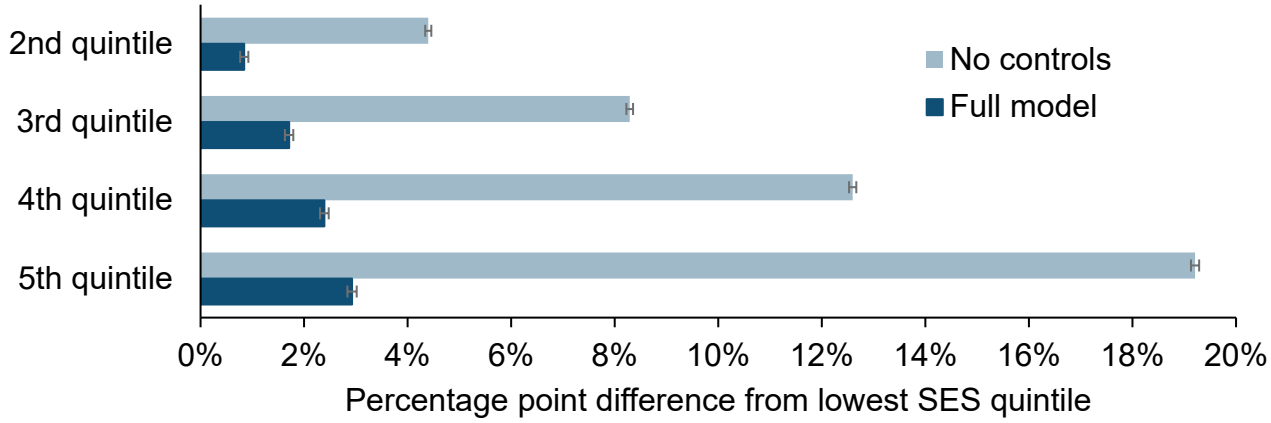
After controlling for the range of demographic and education factors in the full model, we see that males from all other socioeconomic groups are still more likely to be in a good outcome than those from the lowest socioeconomic quintile, i.e. for men with similar characteristics and education profiles, a higher socioeconomic status during GCSEs means they are more likely to be employed and have earnings in the upper income quartile.

### Females

For females and good outcome, the picture is similar to that for males (Figure 25). Again, we see that before we add the demographic and education controls, women from higher socioeconomic status groups are more likely to be in a good outcome than those from lower socioeconomic backgrounds, although the differences are much bigger than for

males. Females from the highest SES quintile are observed to be around 19 percentage points more likely to be in a good outcome than those from the lowest SES quintile.

**Figure 25: Females - Marginal effects of SES on good labour market outcome**



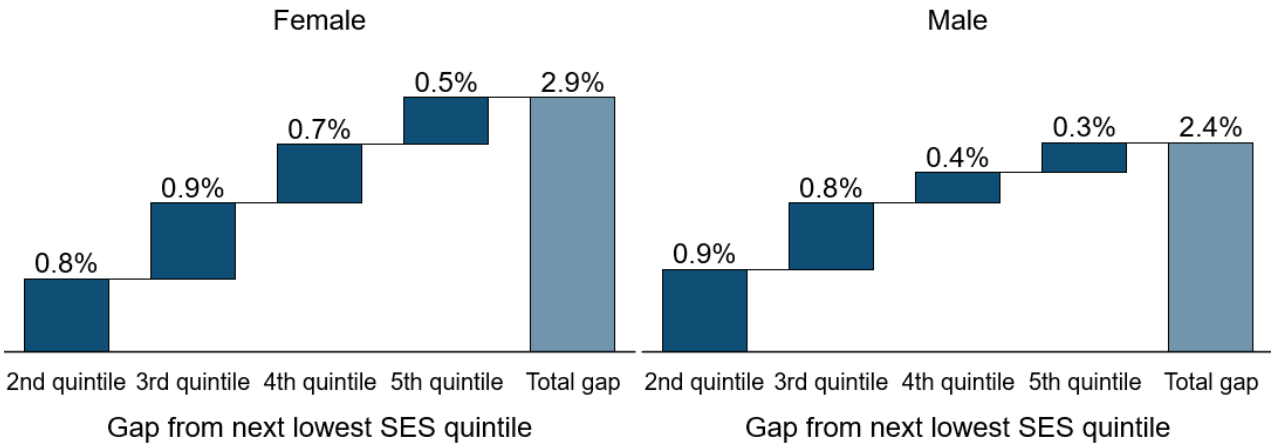

Error bars represent 95% confidence intervals

Source: Authors' analysis using Longitudinal Education Outcomes data

As with men, after controlling for the full range of demographic and education factors, females from all SES quintiles are still more likely to be in a good outcome than those from the lowest. This means that women from the lowest SES backgrounds are less likely to be employed and upper quartile earners than similar women from higher SES backgrounds.

These results in the previous two figures show the difference in good outcome for each quintile from the lowest. Figure 26, however, shows the difference in likelihood of good outcome (after controlling for all demographic and education factors) between each quintile and the next for both males and females i.e. the columns marked '5th quintile' shows the difference in chance of good outcome between the highest SES quintile and the 4$^{th}$.

**Figure 26: Difference in chance of Good Outcome from next lowest SES quintile: males and females**




Source: Authors' analysis using Longitudinal Education Outcomes data

We can clearly see that the gaps between the first and second quintiles and the second and third quintiles are bigger than the gaps between the higher socioeconomic status quintiles, particularly for men. This means that more of the disparity between those from different socioeconomic backgrounds is for the more disadvantaged groups.

## Poor outcome

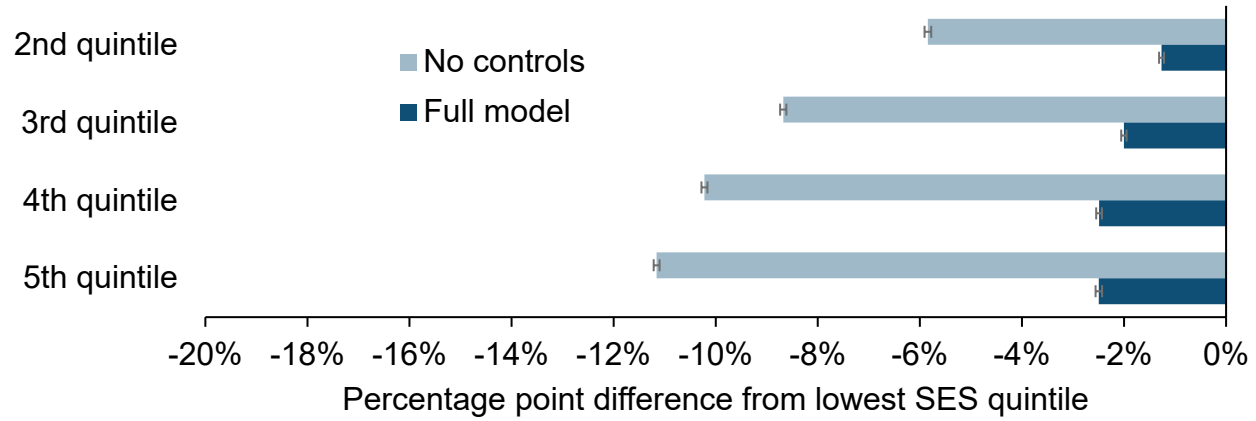
When looking at poor labour market outcome, a positive percentage difference indicates higher proportions claiming out of work benefits for six months in a year, which means worse labour market outcomes. With poor outcome in general we see trends consistent with good outcome in that, on a like for like basis, those from lower SES backgrounds are more likely to have a poor labour market outcome, particularly for females.

### Males

Males from higher SES background are less likely to be in a poor outcome than those from lower SES backgrounds (Figure 27) by around 11 percentage points between the highest and lowest SES quintiles and nearly 6 percentage points between the lowest and second lowest SES quintiles.

When we add the demographic and education controls the gaps in poor outcome between SES quintiles are reduced, but those from lower SES backgrounds are still more likely to be in a poor outcome than those from higher SES backgrounds (around 2.5 percentage points between the highest and lowest SES quintiles).

**Figure 27: Males - Marginal effects of SES on poor labour market outcome**




Error bars represent 95% confidence intervals

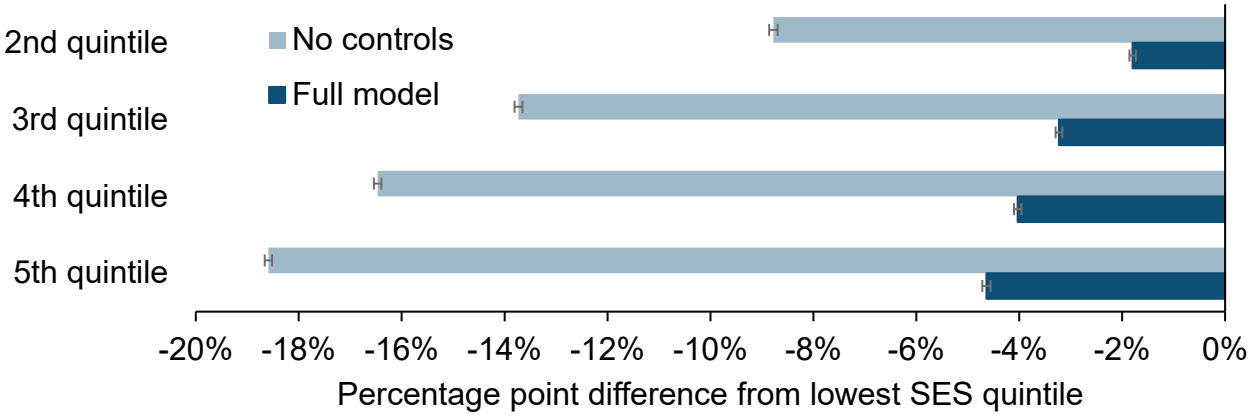
Source: Authors' analysis using Longitudinal Education Outcomes data

This means that when we compare males with similar characteristics and education levels, those from the lowest SES quintile are more likely to be claiming out of work benefits for at least 6 months than males from other quintiles.

### Females

The gap in poor labour market outcome for females from different SES backgrounds is much larger than for males, with the observed gap between highest and lowest SES around 18.5 percentage points (Figure 28).

**Figure 28: Females - Marginal effects of SES on poor labour market outcome**



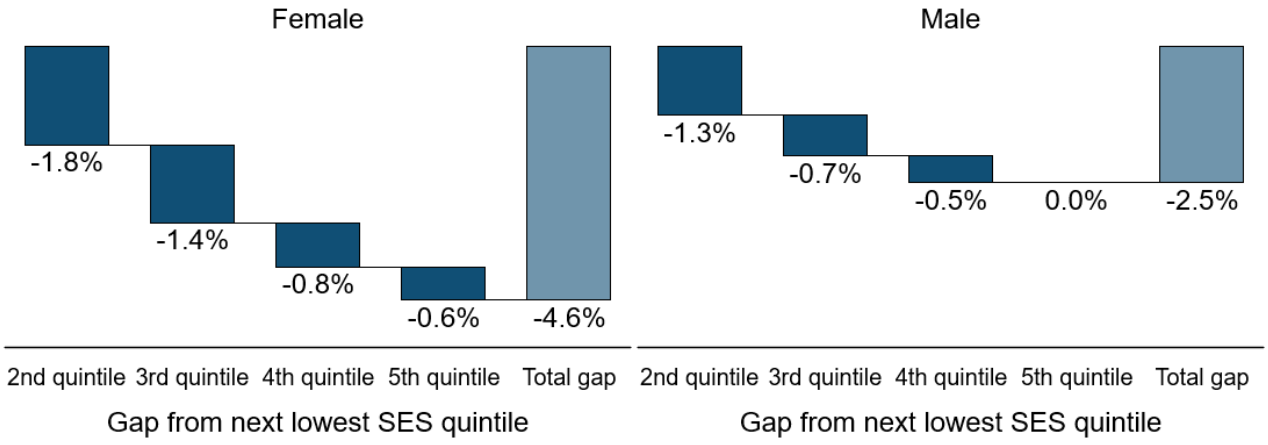

Error bars represent 95% confidence intervals

Source: Authors' analysis using Longitudinal Education Outcomes data

After controlling for demographic and education factors, the gap between SES quintiles is reduced, but there is still a gap of nearly 5 percentage points.

The results in the previous two figures show the difference in poor outcome for each quintile from the lowest. Figure 29, however, shows the difference in likelihood of poor outcome between each quintile and the next (after controlling for all demographic and education factors) for both males and females i.e. the columns marked '5th quintile' shows the difference in chance of poor outcome between the highest SES quintile and the 4th.

**Figure 29: Difference in chance of Poor Outcome from next lowest SES quintile: males and females**




Source: Authors' analysis using Longitudinal Education Outcomes data

We can clearly see that the gaps between the lower quintiles are bigger than the gaps between the higher socioeconomic status quintiles for both males and females. This means that more of the disparity between socioeconomic status groups is for the more disadvantaged groups. The disparity in the chance of a poor outcome for females is bigger than it is for males between all quintiles. For males, there is no gap in the probability of poor outcome between the highest and 4th SES quintiles.

In summary, for both men and women and for good and poor labour market outcome, we observe better outcomes in the labour market with higher socioeconomic status. Holding demographic and education factors constant reduces these gaps but they are still present.

# Section 4: Relative importance of controls

The regression results in Section 3 illustrate that there are gaps in good and poor labour market outcomes between males and females from different socioeconomic backgrounds and that only part of these gaps can be explained by the demographic and education controls.

Demographic and education factors all play a part in determining whether someone is employed, how much they earn, and whether they claim benefits, and the information in Section 2 shows us that SES groups differ for each of these: there are variations in the composition of these, and how good and poor outcome are affected varies too. This raises questions as to whether some of these factors are more important than others, and how much of a contribution each makes in explaining outcome gaps.

This section presents the use of decomposition methodology to attempt to quantify the contribution each demographic and education factor makes towards the gaps in outcomes between the highest and lowest SES quintiles.

Decomposition analysis is a way of isolating the importance of each factor: to determine how much of the gap between the two socioeconomic groups is explained by each factor (or group of factors). A discussion follows on the multivariate decomposition methodology used and the interpretation of results to provide further insight into this.

## Methodology

Multivariate decomposition analysis is used to give insight into the importance of factors in explaining the difference in average outcomes between two groups. For this analysis, it is used to look at the difference in probability of a good labour market outcome (or poor labour market outcome) between those from the **highest SES quintile** and those from the **lowest SES quintile** and to quantify the contribution of each of the factors that may be driving the difference between these two groups. As the analysis focuses on either end of the labour market spectrum, so too this part of the analysis focuses on the top and bottom socioeconomic groups.

A brief description of the method used is provided in this section. Full details can be found in the technical report.

The *mvdcmp*[22] command in STATA was used to run a probit regression model and then 'decompose' or split the raw difference in outcomes between the groups into two parts:

- **Characteristics**: the proportion of difference due to the different compositional makeup (demographic and education) of the two groups
- **Returns**: the proportion of difference which cannot be explained by accounting for the differences in characteristics, i.e. due to the different behaviours, experiences and returns to those behaviours, of individuals in these two groups with the same characteristics

Each of these parts is then further broken down to show the proportion of the difference explained by **each factor** in the model. The **Returns** component also includes the contribution from the constant term: this is the difference in outcomes which cannot be explained by the factors in the model. This will be referred to as the **Unexplained** component in this chapter.

The decomposition analysis was carried out using good labour market outcome as the dependent variable for males and for females separately. The explanatory variables (or factors) used in the model are based on those used as controls in the full probit regression model discussed in Section 3 (as we want to ascertain the importance of each of these in explaining the differences between socioeconomic status groups). For ease of presentation and interpretation, some of the factors have been grouped together in the results charts to indicate the total contribution from related factors.

The groups of variables presented in the charts are (see Definitions):

- Ethnicity
- SEN - total contribution from special educational need
- English as an additional language (EAL)
- Region of key stage 4 school - local authority is used as a control in the probit regression, however the breakdowns for some ethnic groups are too small for local authority to be used
- Pre-16 attainment – key stage 2 maths level, key stage 4 performance points
- School factors (school type, progress, demographics and attainment)

[22] Mvdcmp: Multivariate Decomposition for Nonlinear Response Models (sagepub.com)

- Variables on level and type of educational achievement, as well as institution type, classification and subject studied at degree or above have been combined to give the following post-16 education factors (presented separately). Below level 2 learning is the reference value for these:
    - Below degree level – achievements at levels 2 to 5
    - Lower second class, third class or unclassified degree
    - 1st class or upper second class degree
    - Postgraduate level achievements

This process was then repeated using poor labour market outcome as the outcome variable of interest.

The assumptions and methodology for the probit regression and decomposition differ, and the explanatory variables in the decomposition differ slightly from the controls used in the probit regression (such as using region rather than local authority). This results in slight differences in the explanatory power of these models.

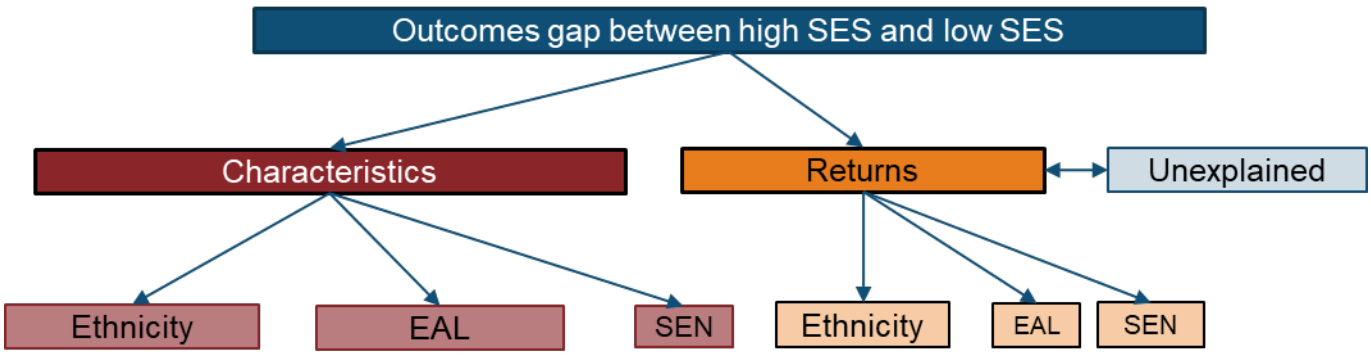
Further details on the explanatory variables and more detail on the decomposition methodology can be found in the technical report.

## Interpretation

The gap in outcomes between the two groups of interest (e.g. the difference in probability of good outcome between males from the highest and lowest SES quintiles) is broken down to show the relative contribution of each factor.

To illustrate this, Figure 30 shows a simplified representation of a model with three explanatory variables.

**Figure 30: Representation of decomposition analysis**




Each factor in the **Characteristics** component, the **Returns** component and the **Unexplained** term makes up a percentage of the total gap (100 per cent) in outcomes

between these two groups. The percentage contribution from a factor can be either positive or negative depending on its association with the outcome variable.

## Results

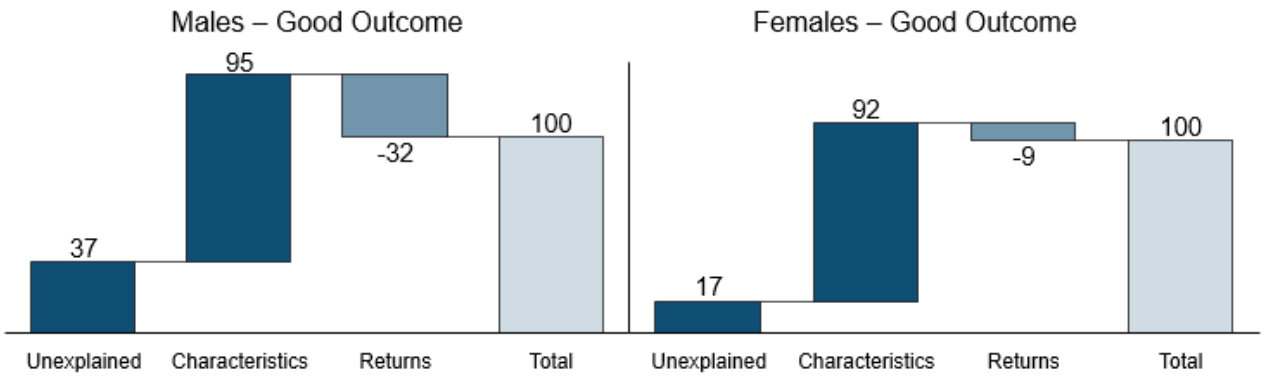
The decomposition results examine the gaps in good and poor outcome between the highest and lowest SES quintiles, for males and females separately. The first part of this section will look at the contributions of the components at a summary level, and the second part will examine more closely the contributions of the individual groups of factors.

### Good outcome summary decompositions

The chance of being in a good outcome is around 15 percentage points higher for males from the highest SES group than for males from the lowest SES group (Figure 24), and 19 percentage points higher for females from the highest SES group compared to females from the lowest SES quintile (Figure 25). The decomposition methodology assigns a high outcome group and compares this to the low outcome (reference) group. As the highest SES group (for both males and females) has a higher chance of good outcome than the lowest SES quintile, the highest SES quintile is treated as the high outcome group and the lowest SES quintile as the reference.

The summary results from the decomposition can be seen in Figure 31, which shows the total percentage contribution from the Unexplained, Characteristics and Returns components for both males and females for good outcome.

**Figure 31: Gaps in Good Outcome between highest and lowest socioeconomic status groups – component totals**




Components with a positive percentage are associated with an increase in the gap; those with a negative percentage are associated with a decrease in the gap.

Source: Authors' analysis using Longitudinal Education Outcomes data

The **Unexplained** component (similar to a constant term in a regression) accounts for that part of the gap in the outcomes of the two SES groups which cannot be explained by the demographic and education data included in the model. This can have either a positive or negative percentage, dependent on whether (overall) unmeasured factors have a positive or negative effect on outcomes (or, are associated with an increase or decrease in the gap between groups).

For the males comparison, the **Unexplained** component makes up 37 per cent of the gap i.e. 63 per cent of the difference in good outcome between these two groups is explained by the demographic and education variables in the model. This is a positive percentage and provides a substantial contribution to the higher outcomes seen for high SES males. For females, the **Unexplained** component is smaller, at 17 per cent. Again, this is a positive percentage so contributes to the higher outcomes of high SES females. This indicates that more of the gap in good outcome between high and low SES groups for females is explained by the demographic and education administrative variables than for males.

The **Characteristics** component accounts for a large proportion of the gap between high and low SES males. At 95 per cent, this suggests that the **composition** of the demographic and education characteristics of the two SES groups is contributing to the much higher chance of good outcome observed for high SES males. The negative percentage for the **Returns** component, however, suggests that the behaviours or returns to behaviours for those with the same characteristics are, overall, less likely to result in good outcome for males from high SES backgrounds. So even though the characteristics of high SES males are typically more favourable for labour market success, the difference in outcomes between these two groups is not as large as we might expect.

For females, the total **Characteristics** component is 92 per cent and therefore the different compositions of the high and low SES groups are leading to the higher chance of good outcome for high SES females too. Again, there is a negative **Returns** component so that the behaviours or returns to behaviours for those with the same characteristics are, overall, less likely to result in good outcome for females from high SES backgrounds.

## Good outcome detailed decompositions

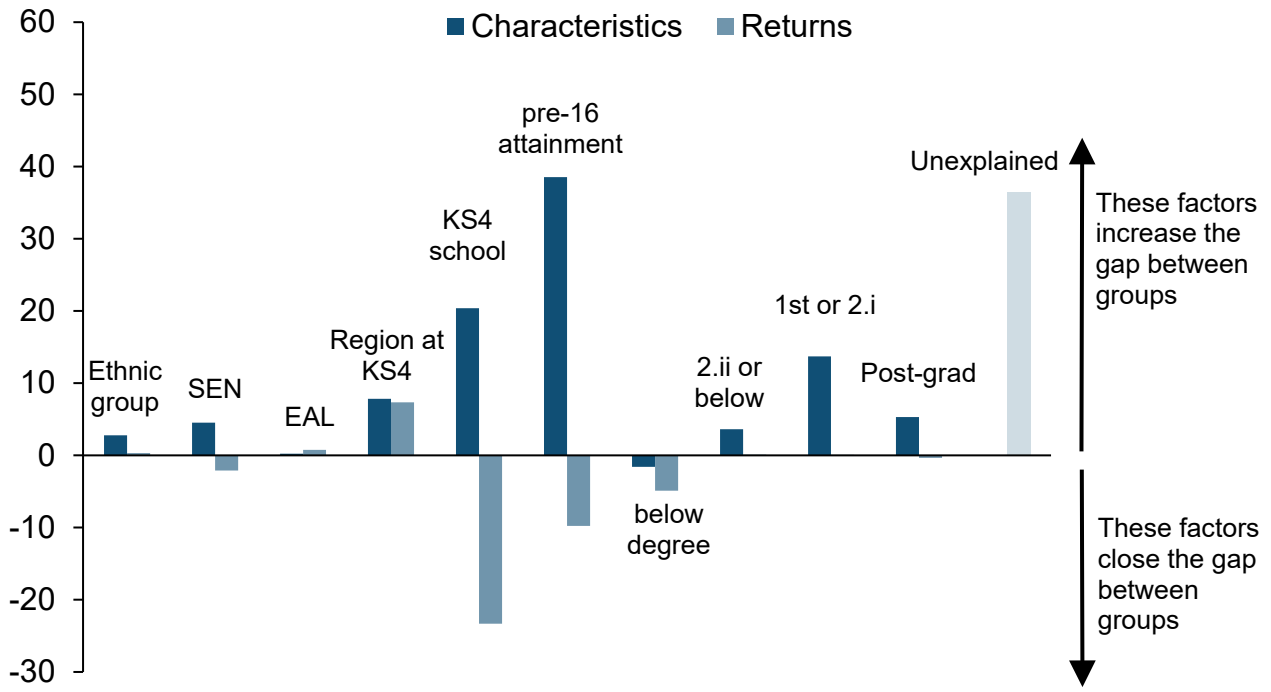
### Males

The results of the detailed decomposition for males from the highest and lowest SES quintiles and good outcome are shown in Figure 32 with contributions towards the gap for each factor or group of factors. The further breakdown of the Characteristics and Returns

components allows us to better understand which of these are driving observed differences in outcomes.

We saw from the summary decomposition in Figure 31 that the differences in **Characteristics** for these two groups accounts for most of the gap in good outcome. Further decomposition shows us that much of this disparity is explained by differences in education levels between the two SES groups. The percentage contribution from a factor can be either positive or negative depending on its association with the outcome variable.

**Figure 32: Decomposition analysis of Good Outcome for high SES compared to low SES males: percentage of gap explained by each factor or group of factors**




Source: Authors' analysis using Longitudinal Education Outcomes data

The factor group with the highest percentage contribution in the **Characteristics** component is Pre-16 attainment (composed of KS2 maths level and KS4 attainment). This has a positive value and suggests that the distribution of school attainment positively contributes to the differences between these two groups. Specifically, it implies that if high SES males were given the same distribution of KS2 and KS4 results as low SES males have, this would lower high SES males' outcomes – resulting in a decrease in the gap between the two groups.

The total contribution from the three higher education factor groups, '2.ii or below', '1st or 2.i' and 'Postgraduate' is the next biggest contribution in the Characteristics component, followed by KS4 school factors (type of school, school progress and peer demographics

and attainment). Again, these factors contribute positively to the gap between the high and low SES groups for males.

The ‘below degree’ factor, which includes academic and vocational qualifications below degree level, has a very small negative value in the Characteristic component, which suggests that the distribution of this factor provides a small negative contribution to the gap between the two groups. Specifically, it implies that giving high SES males the distribution of below degree level achievements as low SES males would slightly increase the chance of good outcome of high SES males and hence increase the gap between the two groups.

The region during GCSEs further increases the gap between males from the two SES quintiles, due to the differences in distribution of these groups across England (higher proportions of the high SES in the South, lower proportions in the North). The positive percentage in the **Returns** component for region suggests that the behaviours or returns to that behaviour of those in the same region of England also positively contribute to the differences between the groups. Specifically, if low SES males had the same behaviours or returns to behaviours as high SES males growing up in the same regions, the gap between the groups would become smaller, as low SES would have higher outcomes.

The demographic factors (such as SEN and ethnicity) are not important in explaining the gaps in outcomes between the two SES groups, but are associated with differences in attainment and achievement which will be reflected in these factors instead. That is to say that it is the lower attainment associated with being identified with SEN which is important in explaining the lower labour market outcomes.

The **Returns** component (which, overall, reduces the good outcome gap between high and low SES males as we saw in Figure 31) is mostly comprised of pre-16 education factors (KS2 maths and KS4 attainment and factors related to the KS4 school). The large negative percentage for these groups of factors indicates that, overall, the behaviours or returns to behaviours of those in similar schools and achieving similar levels in school, while taking into account later educational achievements, reduce the gap in good outcome between high SES and low SES males. Specifically, if low SES males had the same behaviours or returns to behaviours as high SES males for these factors, the gap would between the groups would become bigger, as low SES males would have lower outcomes. The high negative Returns component for the KS4 school effectively cancels out the high positive Characteristics component for these same factors.

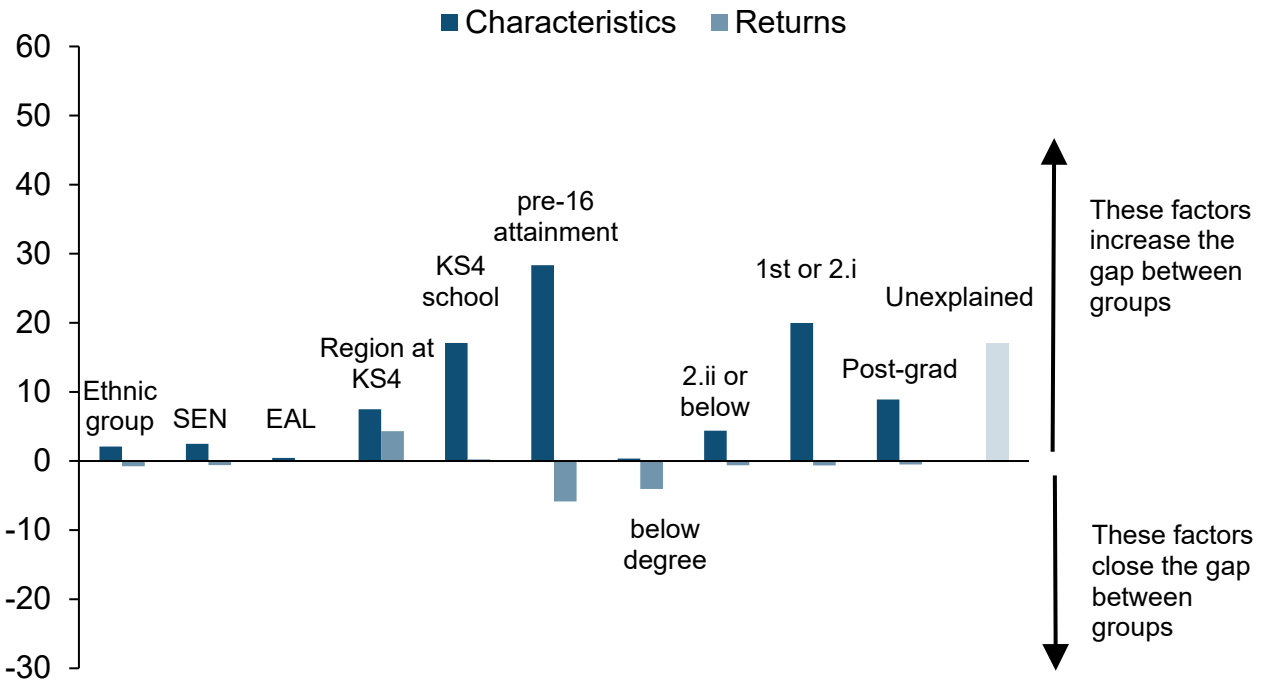
### Females

The results of the detailed decomposition for females from the highest and lowest SES quintiles and good outcome are shown in Figure 33 with contributions towards the gap for each factor or group of factors. The further breakdown of the Characteristics and Returns

components allows us to better understand which of these are driving observed differences in outcomes.

The largest contributors to the **Characteristics** component are similar for females as they were for males, although for females the total contribution from post-16 attainment is higher than the contribution from pre-16 attainment. It can still be said though that the distribution of pre- and post-16 attainment positively contributes to the differences between these two groups. Specifically, it implies that if high SES females were given the same distribution of attainment as low SES females have, this would lower high SES females' outcomes – resulting in a decrease in the gap between the two groups.

**Figure 33: Decomposition analysis of Good Outcome for high SES compared to low SES females: percentage of gap explained by each factor or group of factors**




Source: Authors' analysis using Longitudinal Education Outcomes data

The only negative contributions to the gap in good outcome between females of the highest and lowest SES quintiles are seen in the **Returns** components for below degree achievement and pre-16 attainment. This suggests the behaviours or returns to behaviours of those achieving similar levels in school and achieving below degree level post-16 reduce the gap in good outcome between high SES and low SES females. Specifically, if low SES females had the same behaviours or returns to behaviours as high SES females for these factors, the gap between the groups would become bigger, as low SES females would have lower outcomes.

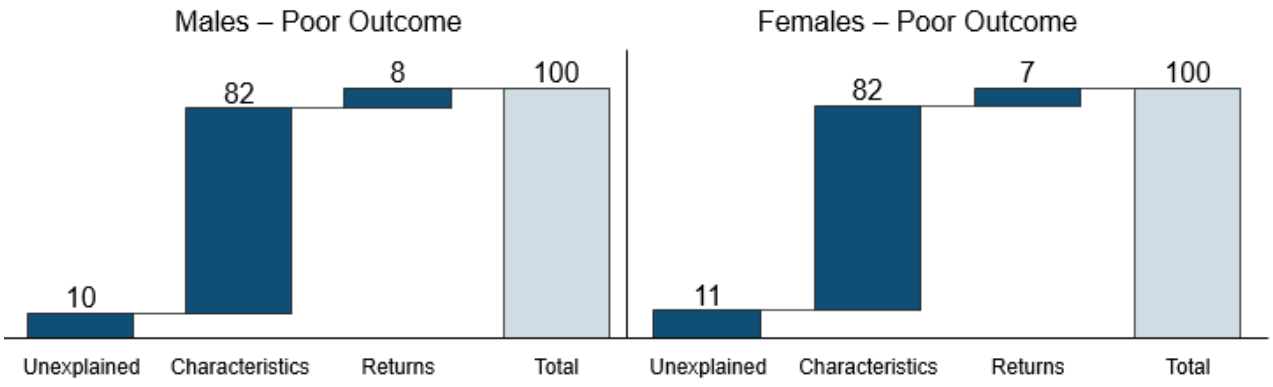
## Poor outcome summary decompositions

The previous sections examined the differences in good outcome between the highest and lowest SES groups and the importance of demographic and education factors in explaining this gap. This section looks at the summary and detailed decompositions for the gap in poor outcome between these two groups.

The chance of being in a poor outcome is around 11 percentage points higher for males from the lowest SES group than males from the highest SES group (Figure 27). The equivalent figure is 19 percentage points for females (Figure 28). The decomposition methodology assigns a high outcome group and compares this to the low outcome (reference) group. As the lowest SES quintile (for both males and females) has a higher chance of poor outcome than the highest SES quintile, in this case the lowest SES quintile is treated as the high outcome group and the highest SES quintile as the reference.

The summary results from the decomposition can be seen in Figure 34, which shows the total percentage contribution from the Unexplained, Characteristics and Returns components for both males and females for poor outcome.

**Figure 34: Gaps in Poor Outcome between lowest and highest socioeconomic status groups – component totals**




Components with a positive percentage are associated with an increase in the gap; those with a negative percentage are associated with a decrease in the gap.

Source: Authors' analysis using Longitudinal Education Outcomes data

The sizes of the contributions to the poor outcome gap between low and high SES groups due to the Unexplained, Characteristics and Returns components are very similar for males and females.

The **Unexplained** component for males is 10 per cent (11 per cent for females), suggesting that 90 per cent of the poor outcome gap is explained by the demographic

and education factors in the model. The **Unexplained** component has a positive percentage so is contributing to the higher chance of poor outcome seen for those from low SES backgrounds.

The **Characteristics** component is 82 per cent, for both males and females, indicating that the composition of the low and high SES groups is contributing most of the gap in poor outcome between these two groups.

In contrast to good outcome, the **Returns** component (for both males and females) is a **positive** percentage. This means that the behaviours or returns to behaviours for those with the same characteristics are, overall, more likely to result in poor outcome for males and females from low SES backgrounds, further compounding the higher chance of poor outcome for this group; the characteristics of those from low SES groups are typically less favourable for labour market success as are the returns to these characteristics.

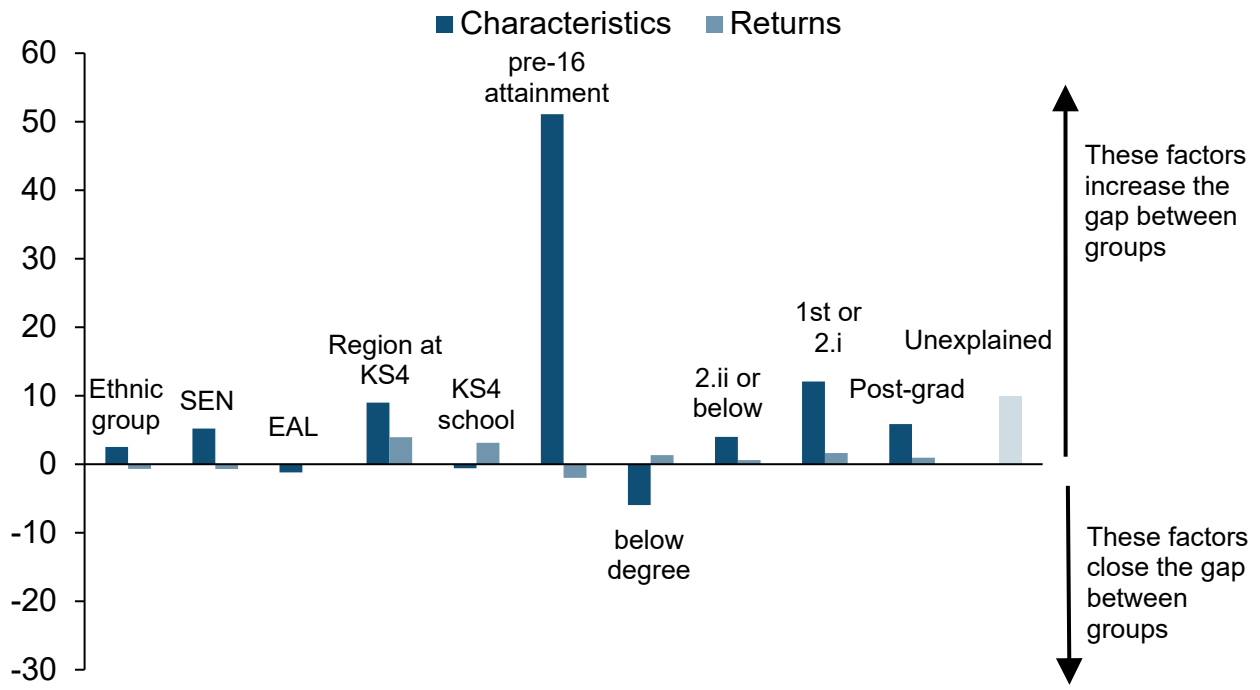
## Poor outcome detailed decompositions

### Males

The results of the detailed decomposition for males from the highest and lowest SES quintiles and poor outcome are shown in Figure 35 with contributions towards the gap for each factor or group of factors. The further breakdown of the Characteristics and Returns components allows us to better understand which of these are driving observed differences in poor outcome.

We saw from the summary decomposition in Figure 34 that the differences in **Characteristics** for these two groups accounts for a large proportion of the gap in poor outcome. Further decomposition shows us that much of this disparity is explained by differences in education levels across the two SES groups. The percentage contribution from a factor can be either positive or negative depending on its association with the outcome variable.

**Figure 35: Decomposition analysis of Poor Outcome for high SES compared to low SES males: percentage of gap explained by each factor or group of factors**




Source: Authors' analysis using Longitudinal Education Outcomes data

A positive percentage for a factor represents one which is associated with increasing the gap between the high outcome and low outcome group. **Note that for poor outcome, the high outcome group has a higher chance of poor outcome, therefore this group has poorer labour market outcomes.** As the lowest SES quintile (for both males and females) has a higher chance of poor outcome than the highest SES quintile, in this case the lowest SES quintile is treated as the high outcome group and the highest SES quintile as the reference.

It can be clearly seen in Figure 35 that the factor group with the highest percentage contribution in the **Characteristics** component is Pre-16 attainment (KS2 maths level and KS4 attainment). This has a positive value and suggests that the distribution of school attainment positively contributes to the gap between these two groups. Specifically, it implies that if low SES males were given the same distribution of KS2 and KS4 results as high SES males have, this would decrease low SES males' chance of poor outcome – resulting in a decrease in the gap between the two groups.

The total contribution from the three higher education factor groups, '2.ii or below', '1st or 2.i' and 'Postgraduate' is the next biggest contribution in the Characteristics component, followed by region of KS4 school. Again, these factors contribute positively to the gap between the high and low SES groups for males.

The ‘below degree’ factor, which includes academic and vocational qualifications below degree level, has a very small negative value in the Characteristic component, which suggests that the distribution of this factor provides a small negative contribution to the gap between the two groups. Specifically, it implies that giving low SES males the distribution of below degree level achievements as high SES males would slightly increase the chance of poor outcome of low SES males and hence increase the gap between the two groups.

In contrast to that seen for good outcome, the group of factors relating to the KS4 school (school type, progress, peer demographics and attainment) has very little importance in explaining the gap in poor outcome between low and high SES groups for males.

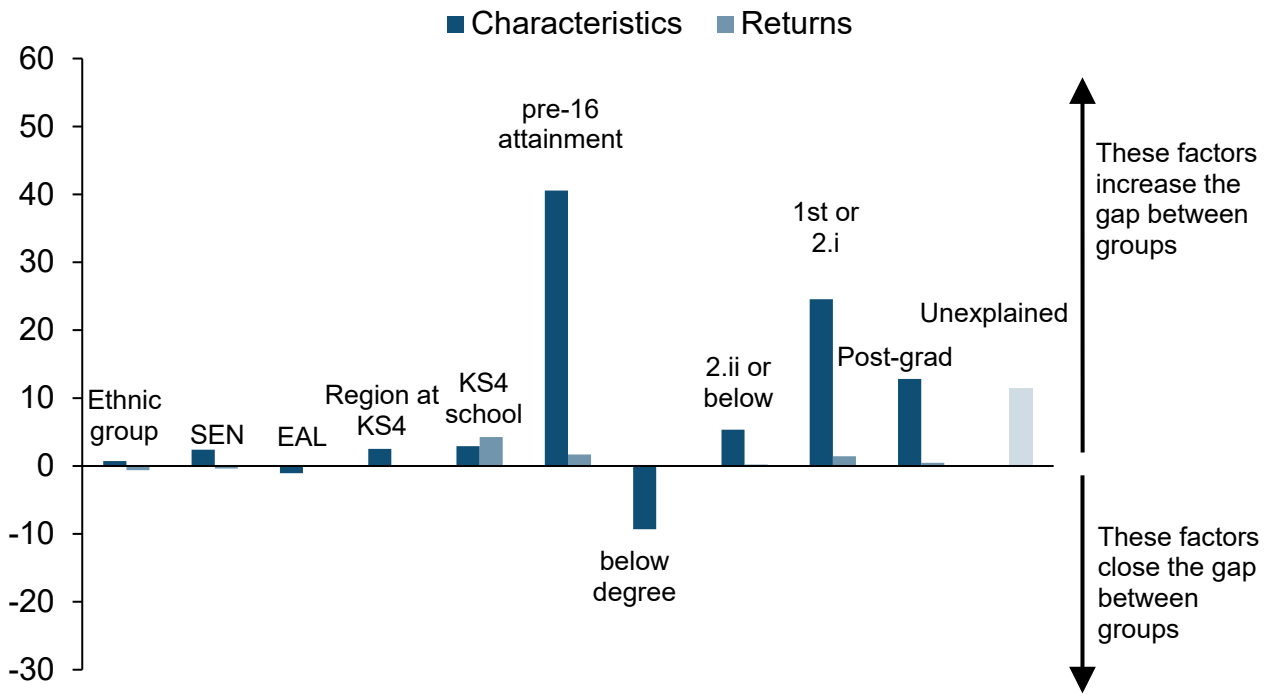
As we saw in Figure 34, the contribution from the **Returns** component is small and in total has a **positive** percentage and is therefore, overall, making a positive contribution to the gap in poor outcome for males from these two SES groups. This contribution is spread across most of the factors or groups of factors with little difference in importance.

### Females

The results of the detailed decomposition for females from the highest and lowest SES quintiles and poor outcome are shown in Figure 36 with contributions towards the gap for each factor or group of factors. The further breakdown of the Characteristics and Returns components allows us to better understand which of these are driving observed differences in poor outcome.

We see a similar picture for poor outcome for females from the lowest and highest SES groups as we did for males, in that differences in the **Characteristics** component for education at all stages provide the largest contributions to the gap between these groups. Pre-16 attainment is important, but the total contribution from the three higher education factor groups, ‘2.ii or below’, ‘1$^{st}$ or 2.i’ and ‘Postgraduate’ is slightly higher. Specifically, this implies that if low SES females were given the same distribution of attainment as high SES females have, this would raise low SES females’ outcomes – resulting in a decrease in the gap between the two groups.

**Figure 36: Decomposition analysis of Poor Outcome for high SES compared to low SES females: percentage of gap explained by each factor or group of factors**




Source: Authors' analysis using Longitudinal Education Outcomes data

The largest negative contribution to the **Characteristic** component for females and poor outcome is the 'below degree' factor, which includes academic and vocational qualifications below degree level. This suggests that the distribution of this factor provides a small negative contribution to the gap between the two groups. Specifically, it implies that giving low SES males the distribution of below degree level achievements that high SES males have, this would increase the chance of poor outcome of low SES males and hence increase the gap between the two groups.

As we saw for males, the **Returns** component makes very little contribution to the gap in poor outcome for females from the lowest and highest SES quintiles, so that most of the explained gap between these two groups is due to differences in compositional makeup and not due to differential returns to the demographic and education factors included in the analysis.

# Discussion

## Summary of key findings

The information presented in Sections 1 and 2 summarises how socioeconomic groups, or those from different socioeconomic backgrounds, differ in demographics and educational participation and achievements: factors which are known to affect labour market outcomes. The descriptive analysis in Section 2 shows that labour market outcomes are consistently better with higher socioeconomic status (higher chance of good outcome and lower chance of poor outcome) even when accounting for factors such as ethnicity, special educational need and highest level of educational achievement. The combination of compositions and outcomes result in large differences across socioeconomic status groups e.g. those from low SES backgrounds are less likely to participate in higher education, more likely to go to a lower quality institution if they do, and less likely to achieve a first or upper second class degree.

This suggests that multiple factors can compound the differences in labour market outcomes and make it hard to make a fair comparison across SES groups. The regression analysis seeks to address this by comparing socioeconomic groups on a like for like basis, taking into account demographic and education differences so we are comparing similar individuals who otherwise differ only in socioeconomic status.

We find that males and females from all (higher) socioeconomic status groups still have a higher chance of good outcome and a lower chance of poor outcome than those from the lowest SES backgrounds after controlling for demographic and education factors, although the gaps between higher quintiles are reduced.

The decomposition analysis provides further insight into this, showing that much of the gap between the most and least disadvantaged quintiles is due to the different compositional makeup of the two groups (the Characteristics component). The returns to each of these demographic and education characteristics (the Returns component) has little importance, except for good outcome for males.

For both males and females, when looking at good and poor labour market outcome, it is clear that the differences in pre-16 and post-16 educational achievements explain the majority of this gap, with the demographic factors having very little importance.

For good outcome, however, the school attended during GCSEs appears to be important in explaining the gap between the highest and lowest SES quintiles, but this is not the case for poor outcome. This suggests that a good school environment can improve labour market outcomes, but a poorer school environment isn’t associated with worse outcomes.

The analysis would suggest that reducing the attainment gap in school would go some way to reducing the disparities in labour market outcomes between socioeconomic groups. Over and above this, increasing the proportions of lower socioeconomic groups in post-16 education would also help to close these gaps. Since 2013 (after these individuals completed their schooling) all young people in England have been required to continue in education or training until the age of 18. There may be barriers to increasing participation at higher education for those from lower SES backgrounds but the analysis suggests that post-16 learning below degree level may help to close the gaps in labour market outcomes and be more accessible to those from more deprived backgrounds.

The regression and decomposition analyses show that there is a residual component not measurable in the administrative data and this is works against the lowest SES group, or in favour of the highest SES group. The analysis accounts for school attainment and school effects, as well as achievement in post-16 education, individual characteristics and geography. This means that the unmeasured differences are explained by other factors such as sociocultural factors (e.g. family/social or societal circumstances, aspirations and expectations, networks, personal choice) and/or discrimination in the labour market. These can be thought of as unobserved advantages or disadvantages of one group or the other.

This chapter highlights the demographic and education factors which are most important in understanding the labour market outcomes of those from different socioeconomic backgrounds. The analysis in this series provides valuable insight into the labour market outcomes of different groups, which is a fundamental part of delivering the government's mission to break down barriers of opportunity for all.

## Next steps

Further work is needed to try to understand what some of these remaining factors are in order to fully understand the differences leading to unequal outcomes in the labour market. Including more detailed labour market information, for example experience and occupation, would be one way to investigate labour market discrimination further. Employment spells and earnings (capturing experience) and sector worked is available in LEO, but not occupation data. There may be other opportunities as more administrative datasets are linked together. In terms of investigating the importance of sociocultural factors, linking to survey and cohort study data would be the best approach. For example, the annual survey of hours and earnings (ASHE) contains information on working patterns which could be important, especially when looking at gender differences. Equally, the Longitudinal Study of Young People in England (and other cohort studies) or PISA have information on family circumstances, motivations, wellbeing, parental aspirations etc.

Other follow up analysis such as heterogeneity analysis could be used to investigate whether the patterns hold for differ for sub groups, for example different locations, different ethnic groups or other demographic variables, or different education outcomes.


© Department for Education copyright 2024

This publication is licensed under the terms of the Open Government Licence v3.0, except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3.

Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned.

For any enquiries regarding this publication, contact www.education.gov.uk/contactus.

This document is available for download at www.gov.uk/government/publications.