

MEM – technical report

October 2018

Contents

1 – Introduction	3
2 – Equality in HE	3
3 – Multidimensionality	4
4 – MEM methodology	7
5 – Updated version of the MEM	11
5.1 – Updated MEM model	11
5.2 – Model testing	12
5.3 – Entry rates	15
5.4 – POLAR3 comparison	16
5.5 – MEM group assignments across other equality dimensions	21
5.6 – Coverage	23
6 – Flexibility	26
Glossary	31

1 – Introduction

The purpose of this report is to describe how the UCAS multiple equality measure (MEM) is constructed. UCAS is committed to the further development of the MEM. The intention is that by providing a public document describing the methodology behind the measure, we will engender confidence in the accuracy and usefulness of the MEM as a standard equality measure across the higher education (HE) sector.

The specific aims of the report are to:

- provide the methodology behind the construction of the MEM
- demonstrate the accuracy and usefulness of the MEM as measure of equality in HE
- produce an updated version of the MEM originally used in the [‘UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics’](#)
- demonstrate the flexibility of the MEM methodology in the analysis of equality in HE

The report begins with a brief discussion of the current equality picture in higher education, as detailed in the ‘UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics’ (Section 2), to provide the context for why a multidimensional approach to equality analysis is needed (Section 3). The full methodology behind the construction of the MEM is then provided (Section 4). This methodology is used to produce an updated version of the MEM, used in the ‘UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics’ (Section 5), and technical details of the MEM model provided (see the annex to this report, Section 5.1). This section continues with an analysis of the updated version of the MEM, looking at its correspondence with the observed entry rate to HE (Section 5.2), patterns of entry to HE (Section 5.3), its comparison to the POLAR3 measure (Section 5.4), its composition in terms of various equality measures (Section 5.5), and a discussion of its coverage (Section 5.6). The report concludes with an outline of the flexibility the MEM methodology provides in the analysis of equality, with examples of how this could lead to development of the MEM (Section 6).

2 – Equality in HE

UCAS regularly reports on equality in higher education, with its flagship publication for sector-level equality being the ‘UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics’. This report provides a detailed analytical picture of the most recent university entry cycle, reporting the entry rates – the proportion of 18 year old students in a specific group who entered higher education through UCAS – for several equality dimensions, including sex, POLAR3 quintile, Scottish Index of Multiple Deprivation (SIMD), ethnic group, and free school meals (FSM) status.

The ‘UCAS Undergraduate End of Cycle Report 2017’ demonstrates that while steady improvement has been made over recent years in the closing of the equality gap between the most advantaged and disadvantaged groups in terms of area background – based on the

POLAR3 measure in England and the SIMD measure in Scotland – looking at gender, ethnic group, and free school meals status, the equality gaps are remaining constant, or widening.

3 – Multidimensionality

Research first conducted by UCAS in 2015¹ demonstrated the limitation of focusing on a single equality dimension in equality assessment, as the following analysis makes clear.

Figure 2 shows the entry rate of subgroups of students in POLAR3 quintile 3. In isolation, a POLAR3 assignment of quintile 3 would not typically identify an individual as advantaged or disadvantaged. However, in this broad POLAR3 grouping, there are subgroups with a range of entry rates, some of which are very low. Those students who are in POLAR3 quintile 3 and are:

- white, male, and in receipt of free school meals
- white, female, and in receipt of free school meals
- mixed ethnicity, male, and in receipt of free school meals

have entry rates that are below the average for those students in POLAR3 quintile 1, a quintile that is frequently identified as disadvantaged.

These are therefore subgroups of students who, on the basis of their POLAR3 categorisation, would not be considered disadvantaged, but who are among the least likely to enter HE. The use of a single measure thus risks creating blind spots in the assessment of equality. This means the most disadvantaged students can be missed, because they are only identified as disadvantaged if multiple background characteristics are considered in combination.

This effect is not limited to the use of the POLAR3 measure. Figure 3 shows a similar chart to Figure 2, here looking only at those students who received free school meals (FSM) while at school – a group typically considered as disadvantaged. A range of entry rates is again seen among subgroups in this cohort. A reverse of the behaviour in Figure 2 can be found, in which there are subgroups of students who are very likely to enter university, some having entry rates approaching those of the average POLAR3 quintile 5 entry rate, but who nevertheless would be identified as disadvantaged if considered only on the basis of the free school meals measure.

These limitations of using single measures demonstrate the need for a multidimensional approach to equality analysis, in which multiple measures of equality are considered simultaneously.

¹ [UCAS 2015 End of Cycle report](#), Figures 96, 97

Figure 1: Entry rates for English 18 year old state school students in POLAR3 quintile 3, by ethnic group, sex, and FSM status at age 15 (2013 to 2017). Coloured vertical lines show the average POLAR3 entry rates

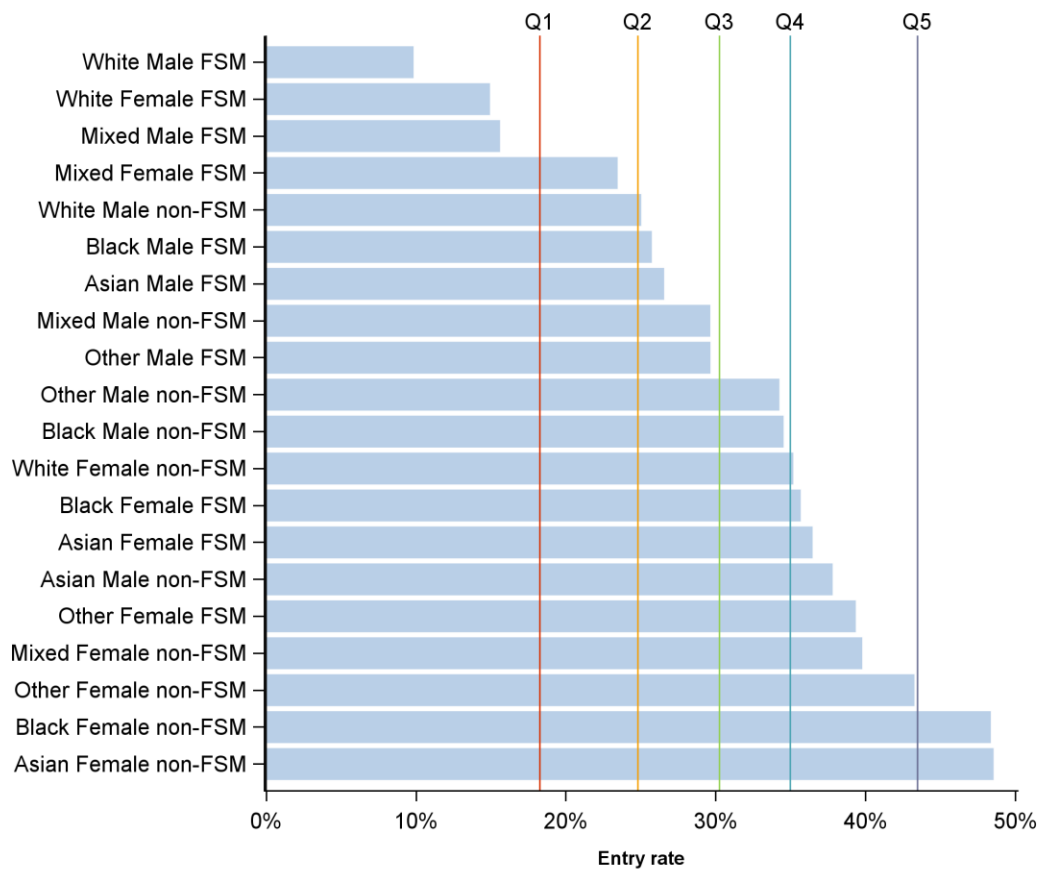
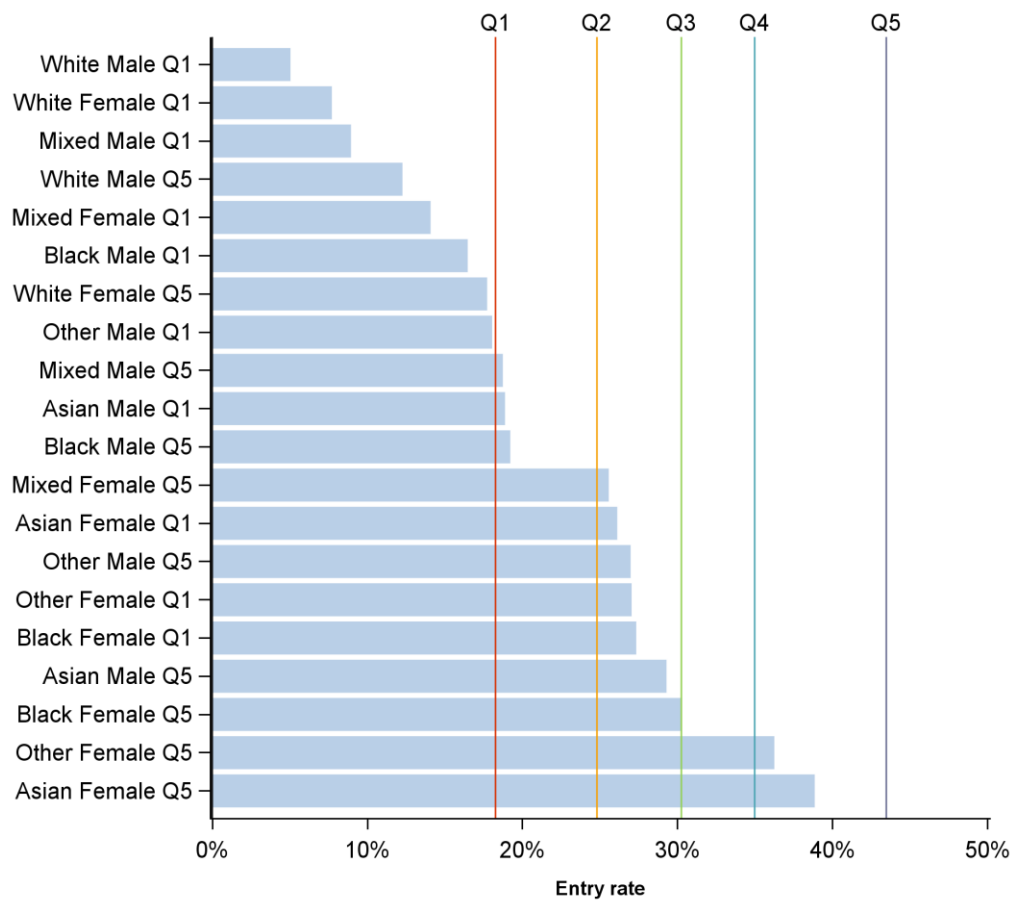


Figure 2: Entry rates for English 18 year old state school students receiving FSM at age 15 by ethnic group, sex, and POLAR3 quintile (Q1 and Q5) (2013 to 2017). Coloured vertical lines show the average POLAR3 entry rates



4 – MEM methodology

The UCAS multiple equality measure (MEM) is an equality metric for higher education, combining the effects of many of the measures currently used in the analysis of equality in HE into a single value. The result of applying the MEM methodology is to produce, through statistical modelling, a modelled likelihood of an individual entering HE, based on that individual's full complement of background equality characteristics. This modelled likelihood is evidence-based, in that it is based on effects that are present and significant in the underlying UCAS applicant data, and highly individualised, due to the use of multiple equality characteristics in its determination. This section will provide the methodology behind the construction of the MEM.

The MEM methodology encompasses the modelling approach required to produce the likelihood (expressed as a probability) of an individual student entering HE based on that individual's background equality characteristics, the selection of equality characteristics to be included, the years of data to be included in the model, and the final presentation of the modelled likelihoods.

The principles on which the MEM is based are as follows:

- **Accurate** – the measure must provide a clear definition of 'disadvantage', and, based on this definition, correctly quantify the level of disadvantage, through simultaneous analysis of multiple equality dimensions.
- **Data-driven** – the measure must be based on extensive and robust data, ensuring only those effects that are real and significant in the data are captured.
- **Individual-level** – the measure must be highly specific to an individual's set of background circumstances.
- **Flexible** – the methodology must allow the opportunity for development and improvement through the inclusion of new data.

Based on these principles, the following methodology is adopted:

Modelling approach

A logistic model is run on an individual-level dataset. This dataset contains equality variables (e.g. POLAR3 quintile, gender, and ethnic group) on each individual, as well as a flag indicating whether or not that individual entered full-time undergraduate higher education through UCAS. An example of such a dataset would be the National Pupil Database (Source: NPD, National Pupil Database), which contains information on the characteristics of students in English schools, linked to UCAS' data on progression to higher education.

The response variable modelled in the logistic model is whether or not each individual entered HE, and the covariates used are the selected equality variables. In this way, the model will produce, for each individual in the dataset, a modelled likelihood of that individual entering HE, based on their values of the equality variables used as model covariates.

This modelled likelihood of entering HE produced by the logistic modelling approach achieves the aims of the MEM:

- **Accurate** – the modelled likelihood is specific to higher education. It defines disadvantage in terms of the likelihood to enter HE (e.g. those individuals with a low modelled likelihood of entering HE would be classified as ‘disadvantaged’ in this context, and those with a high modelled likelihood of entering HE would be classified as ‘advantaged’ in this context), and as such, only those effects of the equality variables relevant to equality in HE will be accounted for. The modelling approach allows analysis of multiple equality characteristics simultaneously, ensuring the limitations of using single measures are removed.
- **Data-driven** – the model will only identify effects that are genuinely significantly present in the underlying input dataset.
- **Individual-level** – the use of an individual-level dataset means the model produces modelled likelihoods for each individual in that dataset. The modelling approach allows for the inclusion of multiple equality characteristics in the analysis, ensuring a high degree of individual specificity in the modelled likelihood.
- **Flexible** – the model approach provides flexibility both in the equality characteristics that are included, and the level at which equality is measured (see Section 6).

Selection of model covariates

The overriding criterion for which variables are selected as covariates in the model is that they are those variables in which there *should be no differential* in the likelihood of entering higher education (i.e. should not alter an individual’s likelihood of entering), but where *analysis of aggregate groups shows there is a differential*. Examples of such characteristics include POLAR3 (students from quintile 5 have a higher entry rate to HE than those from quintile 1), gender (women have a higher entry rate to HE than men), geography (students living in London have a higher entry rate to HE than those from other regions), and ethnic group (students from the white ethnic group have a lower entry rate to HE than those from other ethnic group)².

Beyond this initial selection criterion, the variables selected must comply with data use regulations. The impact of these regulations on the variable selection will depend on the final intended use of the MEM data. Specifically, if the MEM data are to be aggregated, the scope of variables available for inclusion is much broader than that available if the data are to be used at an individual level.

Beyond these two essential conditions, there are multiple criteria that are desirable:

- **Prior research/rationale** – the inclusion of a variable as a model covariate should be backed by evidence that there are differential outcomes in the likelihood of entering HE at an aggregate level, when split by values of this variable. An example would be

² [UCAS Undergraduate End of Cycle Report 2017](#) Figures 5.7, 5.17, 4.6, and 5.16

analysis from the 'UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics', which demonstrated entry rate differentials in POLAR3, sex, ethnic group, and free school meals status.

- **Statistical testing** – variables should be included on the basis that they have a statistically significant impact on the prediction of the model response. This is achieved through stepwise selection of variables, at a $p < 0.05$ significance level.
- **Data integrity** – the data from which the equality variables are obtained must be robust. This extends to ensuring the data source is reliable, hence the data is accurate, and the data is available for enough individuals to ensure both a large base dataset for model accuracy, and widespread applicability of the MEM. The variables used must also be constructed from consistent definitions across the period of use of the MEM.
- **Individual-level** – variables which are individual-specific, such as sex, ethnic group, and FSM status, will produce more individual-specificity in the MEM than those that are related to aggregated populations, such as school-level information.

Where assumptions or compromises are made among these desirable criteria, these should be offset by the public good benefits afforded.

Selection of data years to include in model

The selection of which data years to include in the model is an area of flexibility afforded by the MEM methodology. The years selected depend on the specific application of the MEM. If the MEM is to be used to produce a measure to allow tracking of the progress in equality over a period of time, such as in the version of MEM used in the 'UCAS Undergraduate End of Cycle Report 2017', as previously referenced, it is sensible to construct the model using historical data. This is because here the interest is in the progress made over historical time, in improving outcomes among those groups of students who were historically disadvantaged. Alternatively, if the MEM is being used to create a measure identifying those students who are currently most disadvantaged, then it is sensible to construct the model using the most recent data available. It is important to balance these considerations against the practical requirements of a large amount of input data, to ensure the model is robust.

Presentation of modelled likelihood

The output of the MEM modelling procedure is to produce, for each individual in the input dataset, a modelled likelihood of that individual entering HE, based on their equality characteristics. This output therefore has the form of a continuous number between 0 and 1. A useful alternative form of presentation of this statistic is as quintiles, with each quintile containing an approximately equal share of the population in the input dataset. The presentation of equality metrics as quintiles has been used in the POLAR3 and SIMD measures, and has proved successful in aiding usability and understanding. The procedure used to generate the quintiles for MEM, as in the 'UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics', is to rank all individuals in the input dataset by their values of the modelled likelihood of entering HE, and aggregate these into five groups

– MEM groups – based on this rank, such that each group contains an approximately equal number of individuals (and so represents around 20% of the total dataset population). In this way, each group contains individuals who have a broadly similar modelled likelihood of entering HE. The five MEM groups are named such that ‘group 1’ contains those individuals with the lowest modelled likelihoods of entering HE (‘most disadvantaged’ in this context), and ‘group 5’ contains those with the highest modelled likelihoods of entering HE (‘most advantaged’ in this context).

5 – Updated version of the MEM

The section will outline the use of the MEM methodology to update the version of MEM used in the ‘UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics’. This updated MEM will replace the version that is currently used in the [‘UCAS Undergraduate End of Cycle Report 2017’](#), the [‘UCAS Undergraduate reports by sex, area background, and ethnic group’](#), and UCAS’ provider-level MEM reports.

5.1 – Updated MEM model

The updated MEM model differs from the version used in the ‘UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics’, only in the equality variables that are used as model covariates.

Base dataset for modelling

The base dataset used to run the model is the same as in the version of MEM in the ‘UCAS Undergraduate End of Cycle Report: Patterns by applicant characteristics’ – students in English schools who were aged 18 between 2006 and 2010 (source: National Pupil Database and School Census, Department for Education), linked to UCAS’ admissions data.

Modelling approach

The model response variable used to run the model is the same as in the version of MEM in the ‘UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics’ – the response variable modelled in the logistic model is whether or not each individual entered full-time, undergraduate higher education through UCAS. The output of the model is, therefore, for each individual in the input data, a modelled likelihood of that individual entering HE.

Model covariates

The equality variables used as model covariates differ from those used in the version of MEM in the ‘UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics’.

The covariates to be included in the model for the updated version of MEM are selected using the conditions outlined in Section 4. From the variables available through the National Pupil Database and additional public sources, and after enforcement of the conditions of data integrity, the following variables were considered as candidates for inclusion:

- gender
- POLAR3 quintile
- ethnic group
- free school meals (FSM) status
- Index of Multiple Deprivation (IMD) rank

- school type
- region of domicile
- month of birth
- rural/urban area indicator
- drive time to nearest university

The variables highlighted in red (along with several interaction terms, see annex) are those that are identified as having a statistically significant effect following stepwise logistic regression ($p < 0.05$), so were selected for the updated MEM model.

Due the lack of available data on all but gender and school type of these variables for students in independent schools (8 per cent of the base dataset), it was necessary to separate the modelling procedure into two separate models: one for students in the base dataset who were in independent schools, using gender as the only model covariate, and one for all other students in the dataset, using the variables listed above as model covariates. The results of the two separate models were combined, so the final result of the modelling procedure was a modelled likelihood of entering HE for each student in the base dataset – both independent and non-independent school students.

The use of two separate models for students from independent and non-independent schools was not required in the construction of the version of the MEM in the 'UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics', but it was necessary in the updated version to improve the accuracy of the modelled likelihood of entering HE.

Data years included in the model

The years of data used to run the model are the same as in the version of MEM in the 'UCAS Undergraduate End of Cycle Report 2017: Patterns by applicant characteristics' – students who were aged 18 between 2006 and 2010 were included in the model for the updated MEM. The model was then applied to students who were aged 18 between 2011 and 2017 to generate modelled likelihoods of entering HE, and MEM groups, for these students.

Presentation

For the purposes of calibration testing (Section 5.2), the modelled likelihoods of entry were used directly. For the purposes of exploratory analysis of the updated version of MEM (Sections 5.3 – 5.5), MEM groups were created from the modelled likelihoods of entry, using the procedure outlined in Section 4.

5.2 – Model testing

Full details of the MEM model can be found in the annex to this report.

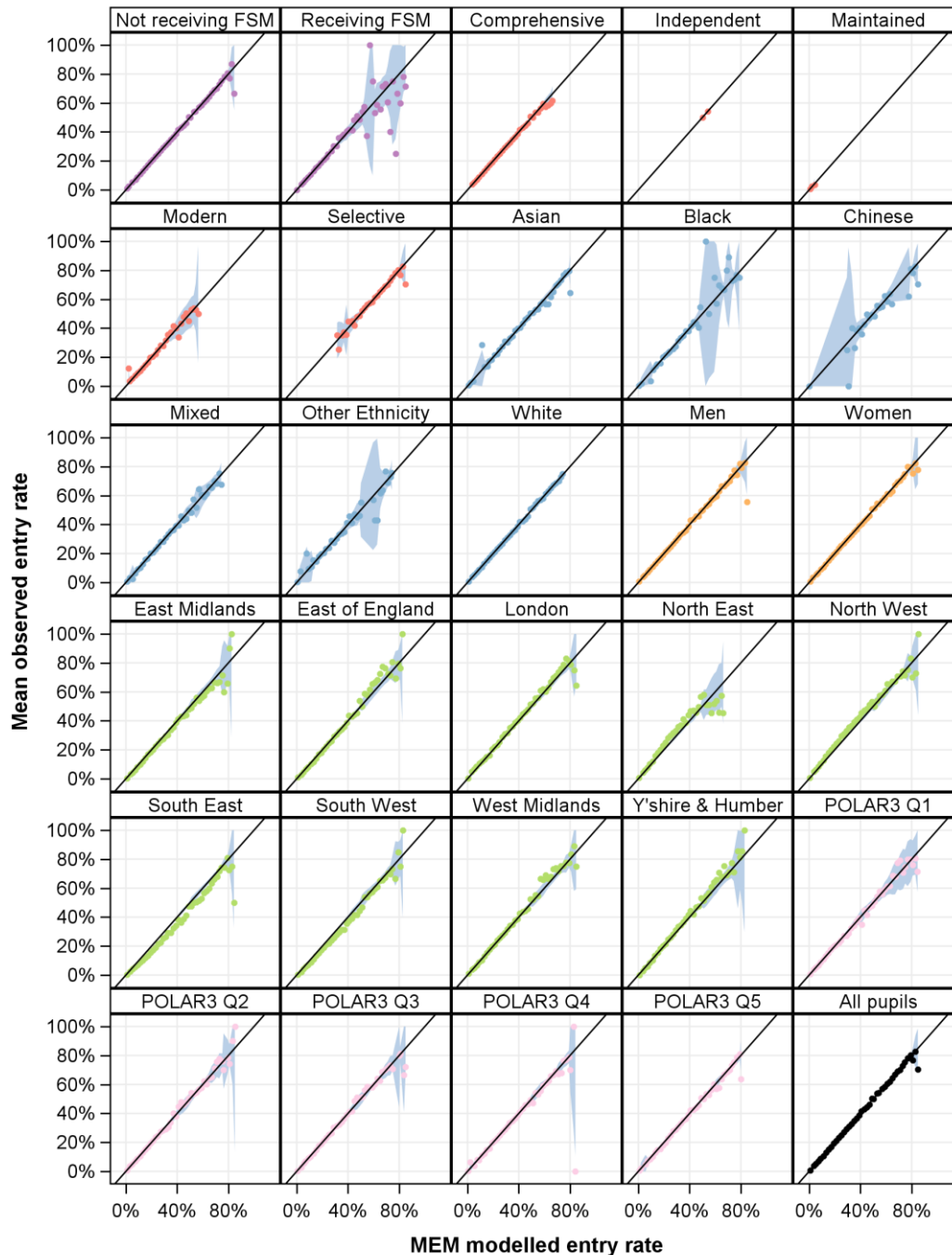
Qualitative assessment of the agreement between the modelled likelihood of entry from the MEM model, and the observed outcome of whether or not students entered HE, is shown in the calibration plots in Figure 3.

Students who were aged 18 between 2006 and 2010 are grouped into 50 bins, based on their values of the modelled likelihood of entering HE. For each of these bins, the mean modelled likelihood of entry, and the proportion of students in the bin who entered HE, are compared.

Figure 3 shows this comparison, with each point corresponding to a single bin of students. The charts also show a diagonal reference line, and the expected statistical variation given the number of students in each bin. Points lying within this interval are considered to have modelled and observed likelihoods of entering HE that correspond, within statistical variation. The comparison is shown both for the total student population, and split by different values of equality characteristics.

There is good correspondence between the modelled likelihood of entering HE (MEM modelled entry rate) and the proportion of students in the bins who actually entered higher education (mean observed entry rate), with no evidence of systematic bias.

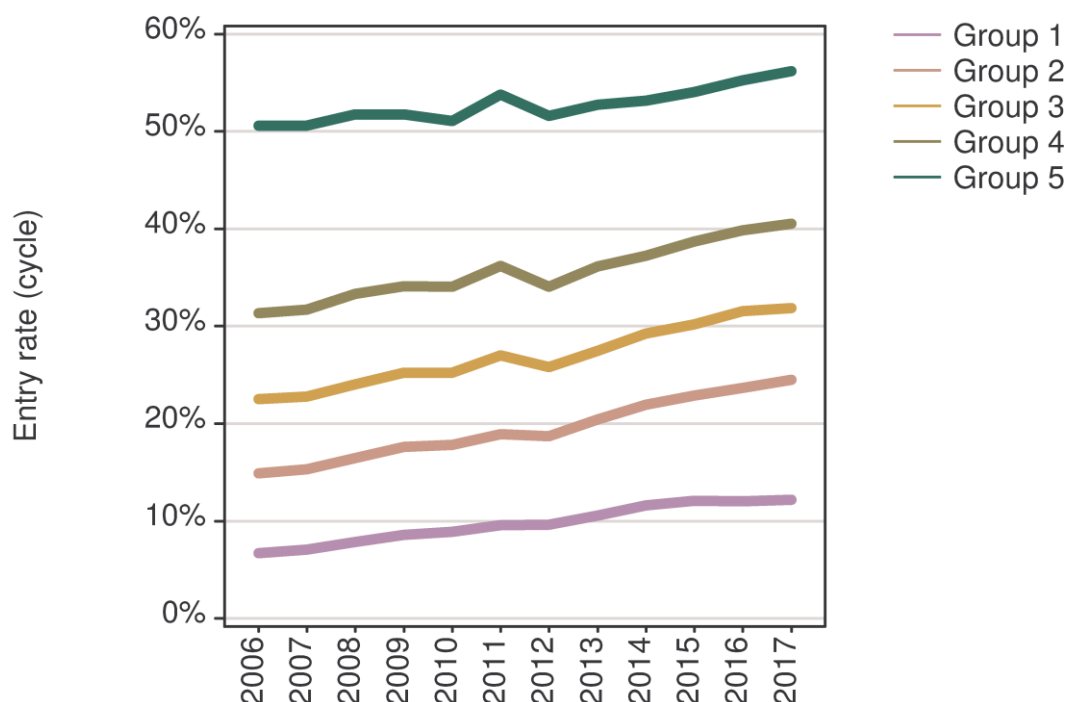
Figure 3: Mean observed entry rate and MEM modelled entry rate of 50 groups of students, split by FSM status (purple), school type (red), ethnic group (light blue), gender (orange), region (green), and POLAR3 quintile (pink). Expected statistical variation interval shown as blue band



5.3 – Entry rates

The entry rates – the proportion of the population entering higher education through UCAS – for English 18 year olds, split by the updated MEM groups, are shown in Figure 4. In 2017, the entry rate for MEM group 1 was 12.2 per cent, for MEM group 2 was 24.5 per cent, for MEM group 3 was 31.8 per cent, for MEM group 4 was 40.5 per cent, and for MEM group 5 was 56.2 per cent.

Figure 4: Entry rates for English 18 year olds by multiple equality measure (MEM) groups



5.4 – POLAR3 comparison

This section provides a comparison of the behaviour of the updated MEM and POLAR3, considering the entry rates, and assignments among students living in London of both measures. Figure 7 provides a summary table of the results. The analysis in this section is limited only to those students who attended non-independent schools, as POLAR3 information is unavailable in the MEM model base dataset for students who attended independent schools (see Section 5.1).

Entry rate comparison

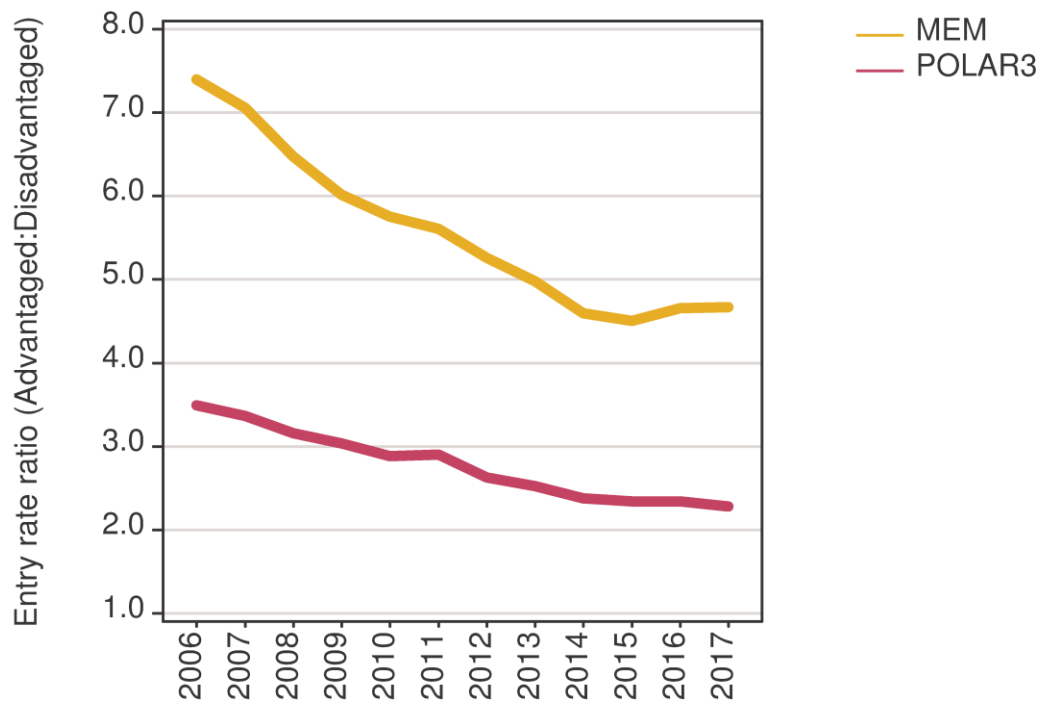
Figure 5 shows the entry rate ratio of the updated MEM (Group5:Group1) and POLAR3 (Quintile5:Quintile1). Examining the entry rate ratio provides a measure of the gap in equality of representation in HE. An entry rate ratio of 1 indicates equal chances of entering higher education for the two groups. Entry rate ratios greater than 1 indicate that students from the more advantaged group (group 5 for MEM, quintile 5 for POLAR3) are more likely to enter than those from the most disadvantaged group (group 1 for MEM, quintile 1 for POLAR3) – for example, an entry rate ratio of 2.5 means that students from the advantaged group are two and half times more likely to enter HE than the disadvantaged group.

In 2017, the entry rate ratio for the updated MEM was 4.7, meaning students from the most advantaged MEM group were 4.7 times more likely to enter university than those from the most disadvantaged group. By comparison, the most advantaged students on the POLAR3 measure were 2.3 times more likely to enter university than the most disadvantaged³. The size of the equality gap on the MEM is greater than that on POLAR3 alone, across the period. Since 2014, the equality gap on the POLAR3 measure has continued to narrow, while the MEM shows no progress has been made in reducing this size of this gap.

The analysis of multiple equality measures, including POLAR3, means the MEM has a higher granularity than POLAR3, and consequently can provide a more robust and specific predictor of whether or not an individual student enters higher education based on their background.

³ The POLAR3 entry rates and entry rate ratio have been calculated here using only those students who also have MEM group assignments. The values given will therefore differ from those found in other UCAS publications.

Figure 5: Entry rate ratio for MEM (Group5:Group1) and POLAR3 (Quintile5:Quintile1) for English 18 year old state school students



Comparison of assignment across state school students living in London

Figure 6 shows the distribution of assignment of both MEM group and POLAR3 quintile among English state school students living in London. A greater proportion of these students are assigned to MEM group 1 (11.5 per cent) and group 2 (18.6 per cent) than are assigned to POLAR3 quintile 1 (3.3 per cent) and quintile 2 (9.5 per cent), respectively. This is due to the overrepresentation of students domiciled in London receiving free school meals (22 per cent of those in London, compared to 14 per cent of the model base student population), a group that typically has low entry rates to HE⁴, and so is predominantly assigned to MEM group 1 and group 2 (see Section 5.5).

⁴ 16.9 per cent in 2017, [UCAS 2017 End of cycle report: patterns by applicant characteristics](#)

Figure 6: Assignment of MEM group and POLAR3 quintile among English 18 year old state school students living in London (over the period 2006 to 2017)

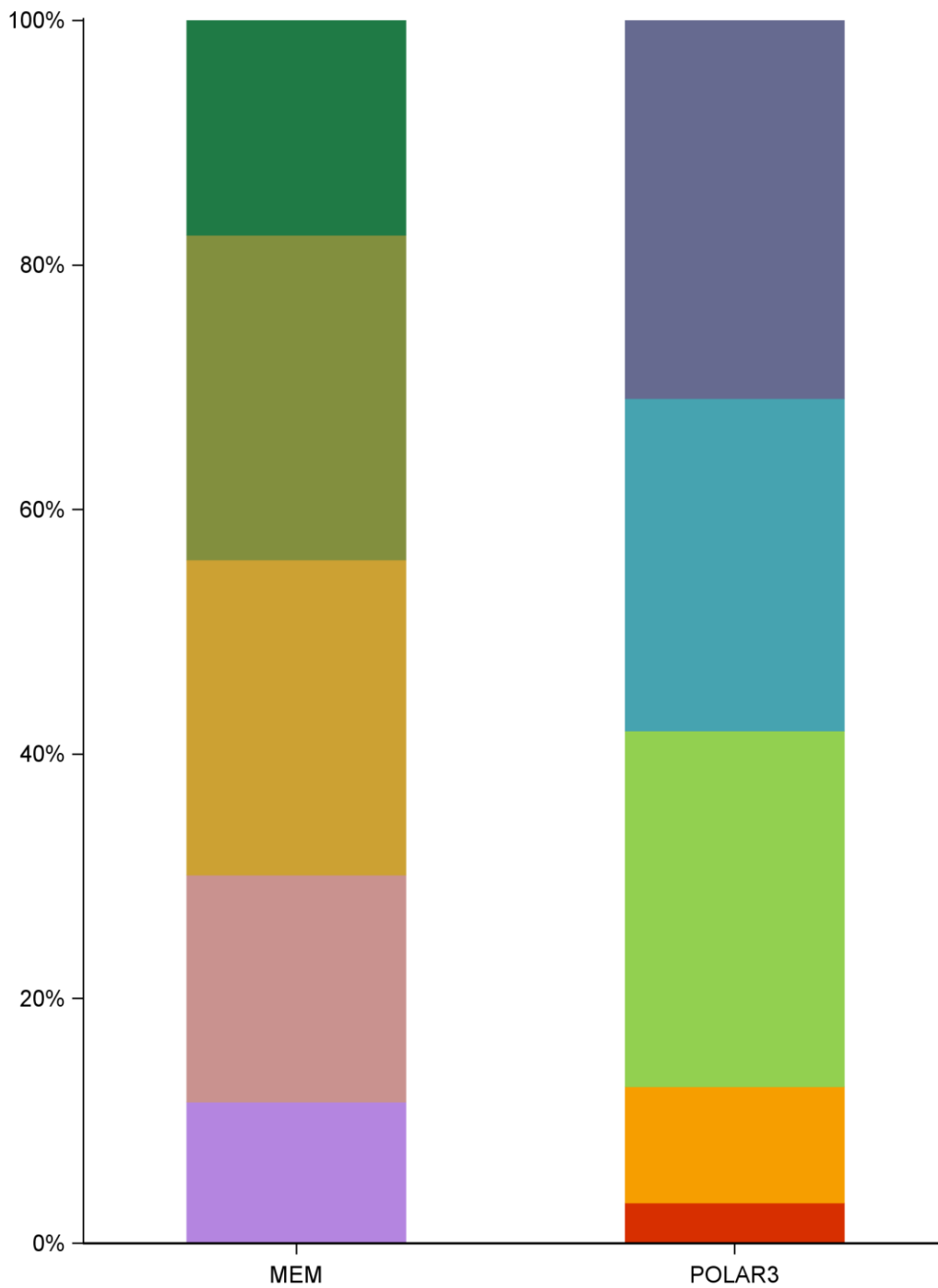


Figure 7: Summary table of comparison between MEM and POLAR3 (English 18 year old state school students). All values are for the 2017 cycle, unless otherwise stated

	MEM	POLAR3
Entry rate of disadvantaged group ⁵ (%)	12.2	19.9
Entry rate of advantaged group (%)	56.8	45.5
Equality gap (entry rate ratio [advantaged group:disadvantaged group])	4.7	2.3
Percentage change in equality gap since 2014 (%) (Positive change indicates equality gap has widened)	1.6	-4.1
Proportion of students living in London assigned to disadvantaged group (2006 to 2017) (%)	11.5	3.3
Proportion of students living in London assigned to advantaged group (2006 to 2017) (%)	17.6	30.9

⁵ For MEM, 'disadvantaged group' = MEM group 1, 'advantaged group' = MEM group 5. For POLAR3, 'disadvantaged group' = quintile 1, 'advantaged group' = quintile 5

5.5 – MEM group assignments across other equality dimensions

Figure 7 shows the MEM group assignments across different values of POLAR3, sex, ethnic group, school type, FSM status, and region of domicile (averaged over the 2006 to 2017 period). The analysis of POLAR3, ethnic group, FSM status, and region of domicile is limited only to those students who attended state schools, as information on these variables is unavailable in the MEM model base dataset for students who attended independent schools (see Section 5.1).

The MEM group assignments reflect the aggregate entry rate behaviour of the higher level demographics, and so correctly identify those higher level demographics that are advantaged (e.g. POLAR3 quintile 5) or disadvantaged (e.g. students in receipt of FSM) in their likelihood to enter HE.

Those groups that have lower entry rates to HE⁶, students who are:

- in POLAR3 quintile 1
- male
- from the white ethnic group
- in receipt of free school meals
- from maintained or modern schools
- domiciled in the North East

are overrepresented in MEM group 1 (above 20% assignment) and underrepresented in MEM group 5 (under 20% assignment), while those that have higher HE entry rates, students who are:

- in POLAR3 quintile 5
- female
- from the Chinese or Asian ethnic group
- not in receipt of free school meals
- from independent or selective schools
- domiciled in the South East

are overrepresented in MEM group 5 and underrepresented in MEM group 1.

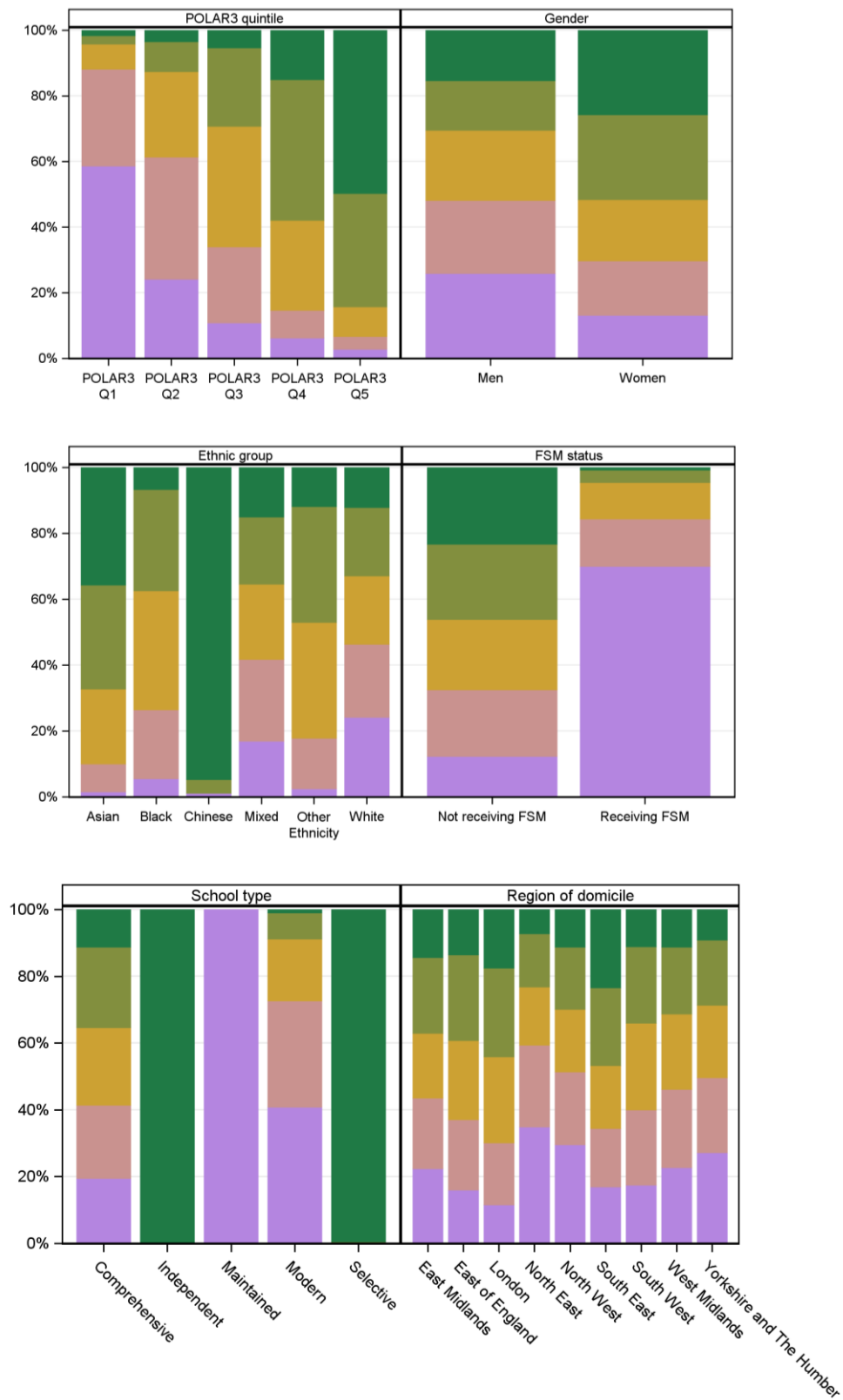
In addition to capturing the behaviours of higher level groups, the presence of a range of MEM group assignments across all higher level groups demonstrates that MEM is nuanced enough to capture effects that arise from the combination of multiple equality measures. As an example, although POLAR3 quintile 5 has an entry rate higher than all other POLAR3 quintiles⁷, there are still some students in POLAR3 quintile 5 who are assigned to MEM group 1 – these are those students who are disadvantaged in terms of their likelihood to enter HE on other equality measures (they have one or more of the following

⁶ See [UCAS 2017 End of Cycle report](#)

⁷ In 2017, the POLAR3 quintile entry rates for English 18 year olds are (%): Q1 – 20.4, Q2 – 26.9, Q3 – 32.5, Q4 – 37.6, Q5 – 47.1, [UCAS 2017 End of cycle report: patterns by applicant characteristics](#)

characteristics: male, from the white ethnic group, from maintained schools, or they were in receipt of FSM).

Figure 8: MEM group assignment by equality dimension for English 18 year old students (over the period 2006 to 2017)



5.6 – Coverage

This section outlines the student coverage of the updated version of the MEM.

English students only

The use of the National Pupil Database as part of the base dataset for the updated MEM model, means the MEM is restricted to English school students only. Expansion to other UK countries could be achieved with the use of similar datasets from these countries.

18 year old students only

The base dataset from which the updated MEM is constructed is limited to students who were aged 18 in the years 2006 to 2017. Therefore, only 18 year old students are assigned a modelled likelihood of entering HE, and so a MEM group.

The decision to limit to 18 year old students is because:

- the majority of prior equality research conducted by UCAS is limited to 18 year old students
- 18 year olds make up the largest single age group of English applicants to higher education through UCAS (47 per cent in 2017)

A substantial prior research base is a desirable criteria in the selection of equality variables to be included as covariates in the MEM model (see Section 4). Without this prior research base for age groups other than 18 year old students, it was deemed sensible to limit the MEM to 18 year old students only. Even with the limitation to 18 year old students, a MEM group assignment is still possible for nearly half of the English applicant population through UCAS.

Full-time undergraduate higher education entry only

The type of higher education that UCAS facilitates is limited almost exclusively to individuals entering into full-time undergraduate higher education, hence the data available for construction of the MEM is similarly limited to full-time undergraduate education. Expansion to other forms of higher education could be achieved under the MEM methodology, if suitable data were available.

Accuracy and completeness

There are several equality variables that, due to data integrity issues, are absent from this updated version of the MEM. These are variables that are often discussed in relation to equality in higher education, such as disability, care status, refugee status, and military child status. The flexibility of the MEM methodology means there is the potential to include these variables in future versions of the MEM, if suitable data are available.

Independent school students

The lack of available data for students from independent schools means that two models are constructed: one for those students from independent schools, with gender as the only

model covariate, and one for all other students (with a full set of equality variables as model covariates – see Section 4). This means the modelled likelihood of entering HE of those students from independent schools varies only on the gender of those students, so is it not possible to achieve as high a degree of individual-specificity in the modelled likelihood of entry, as is achievable for students from non-independent schools. This necessarily means that the MEM fails to capture the more nuanced effects of background on students from independent schools.

UCAS is keen to improve how independent school students are included in the MEM methodology, and the flexibility of the methodology means there is the potential to do so if suitable data are available.

Aggregate-level use

Due to data use restrictions, the specific set of equality variables used in the construction of the updated MEM limit its use to aggregate-level reporting and analysis – it cannot be provided or used at an individual level, to affect individual-specific decisions.

6 – Flexibility

The modelling-based approach of the MEM confers great flexibility at each stage of the MEM methodology (Section 4). By changing the level at which equality is measured, the equality variables that are included as model covariates, the years of data the model is based on, and the presentation of the modelled likelihood, it is possible to construct different versions of the MEM for different purposes. This section discusses the flexibility in each of these areas, with examples.

Level of equality

The response variable in the model for the updated version of the MEM outlined in this report, is whether or not an individual entered HE, from which the model produces a modelled likelihood of each individual entering HE. In this version of the MEM, equality is being investigated at the HE sector-level, as it is the likelihood of entering all HE that is being modelled.

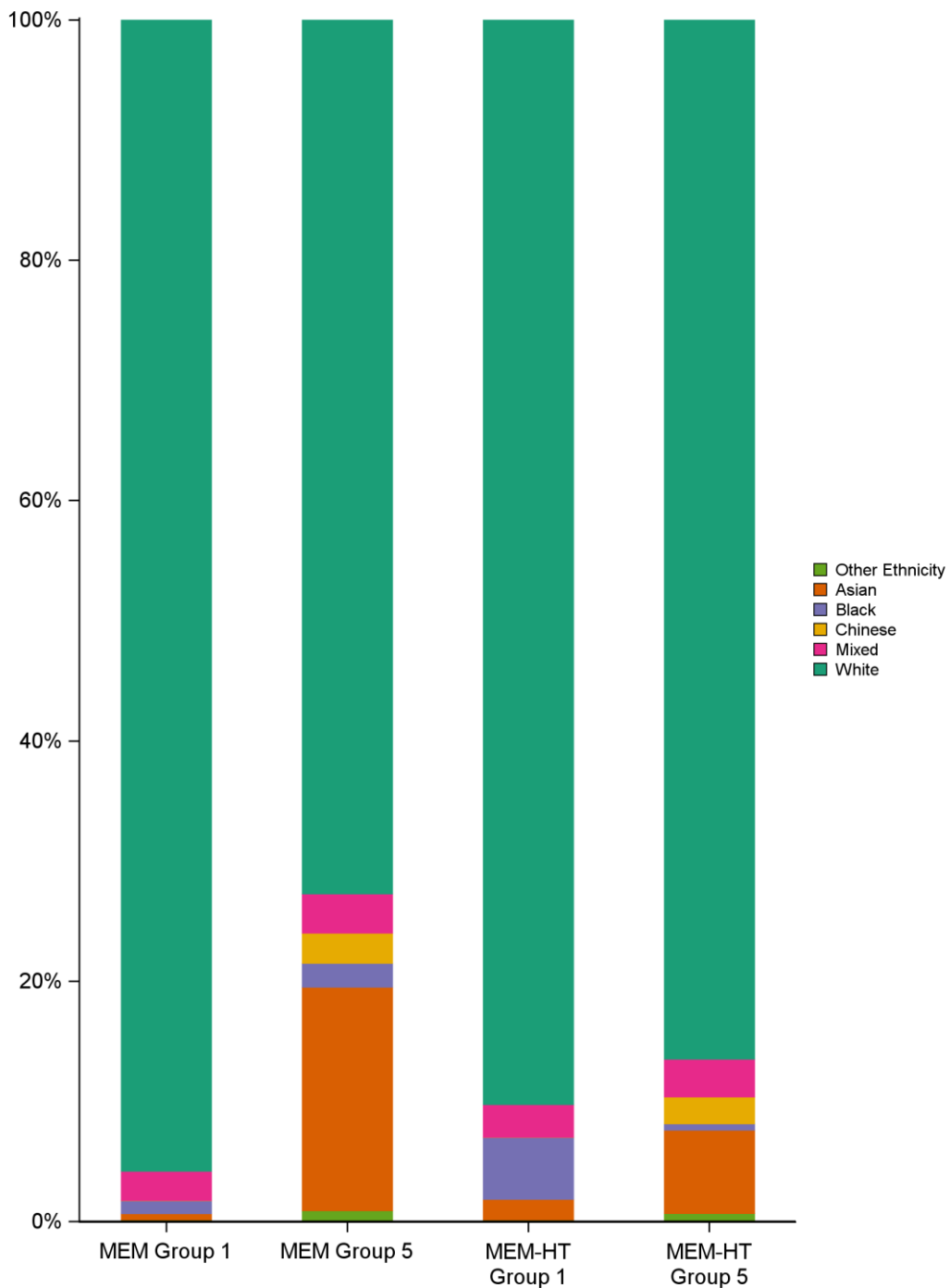
By changing the response variable of the MEM model, it is possible to change the modelled likelihood that is produced by the model, and so the level at which equality is measured. As an example, changing the response variable to whether or not an individual entered HE to study a particular subject (e.g. medicine), the model will produce a modelled likelihood of each student entering HE to study that subject, allowing the particular equality gaps in this subject in HE to be measured. Likewise, changing the response variable to whether or not an individual entered HE at a particular HE provider, means the model will produce a modelled likelihood of each student entering that particular provider, and so allows the equality gaps at that particular provider to be measured.

As an example for analysis, a version of the MEM was constructed that was identical to the updated version of MEM outlined in this report (with the same base dataset, equality variables, and data years), but where the response variable was whether or not a student entered HE at any higher tariff provider. The model produced a modelled likelihood of each student entering any higher tariff provider, and students were then aggregated into groups (MEM-HT group), as for the updated version of the MEM (see Section 6). In this way, MEM-HT is a version of the MEM that measures equality at higher tariff providers.

Figure 8 shows the percentage composition of groups 1 and 5 for the updated version of the MEM, and MEM-HT, by ethnic group. The proportion of group 1 composed of students from the black ethnic group is around five times greater in MEM-HT (5.1 per cent) than in MEM (1.1 per cent). This accurately reflects the differences in entry rates of different ethnic groups to higher tariff universities compared to the whole sector, where at higher tariff providers, the black ethnic group has the lowest entry rate of any ethnic group (6.7 per cent in 2017), compared to one of the highest across the whole HE sector (40.4 per cent in 2017)⁸.

⁸ Entry rates by ethnic group and provider tariff group can be found in the [Equality and entry rates data explorer](#)

Figure 9: Composition of group 1 and group 5 for MEM, and an alternative version of MEM for equality in higher tariff providers (MEM-HT), by ethnic group

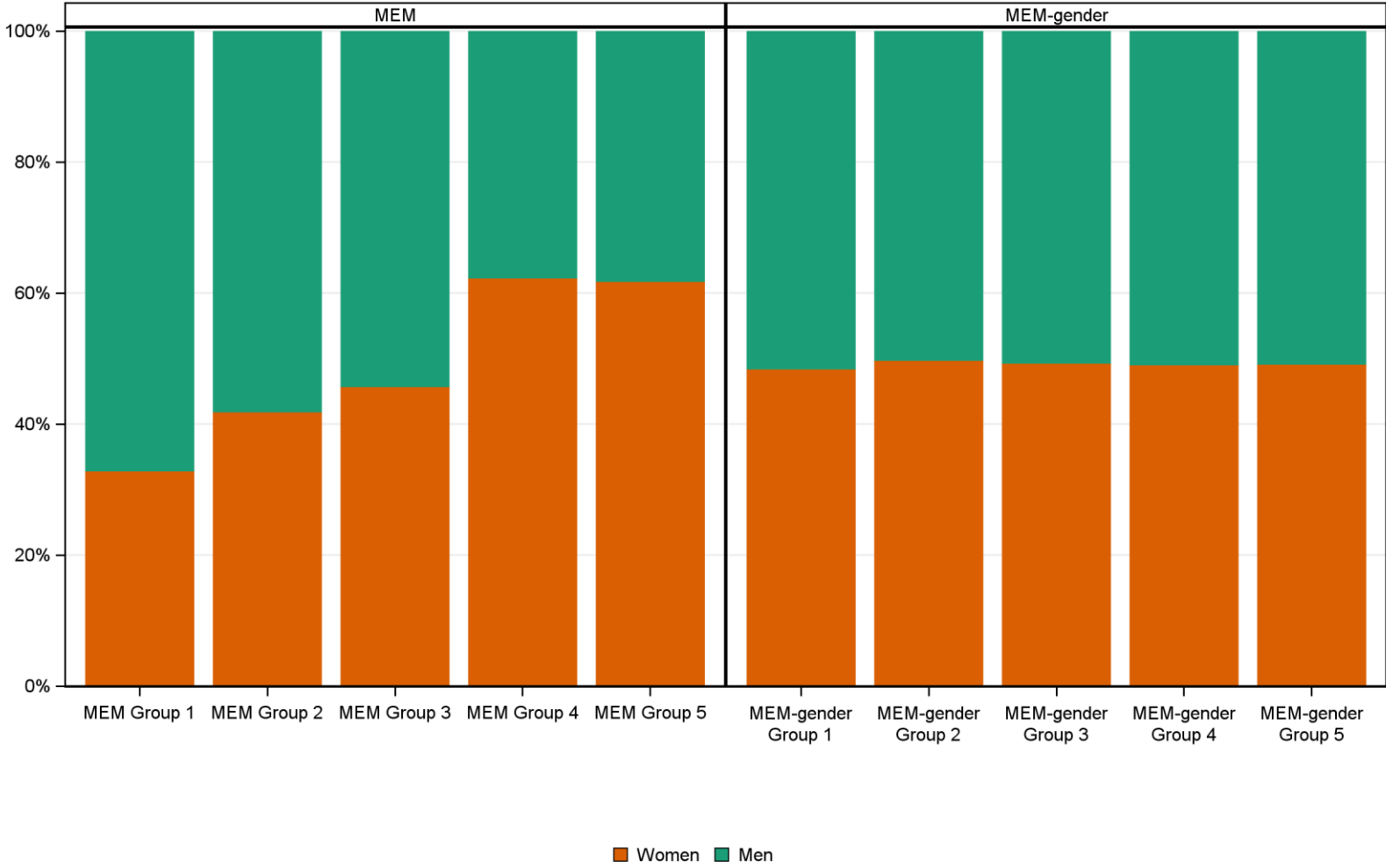


Within the statistical limitations of the modelling procedure used, it is possible to create a version of the MEM to analyse equality at any desired level. The use of multiple versions of the MEM in this way would help ensure that inequality at finer levels isn't masked by analysis of equality at broader levels.

Selection of model covariates

There is flexibility in the equality variables used as covariates in the MEM model, with the only essential criteria being that the variables selected for inclusion must be those where a differential in likelihood of entering HE (or indeed, the specific outcome being modelled – see ‘Level of equality’ in Section 6), would not be expected (see Section 4). This allows the continued development of the MEM when new equality variable data become available, ensuring the MEM is always accounting for as many relevant equality measures as possible. Changing the equality variables in the model also ensures only the effects of the specific equality dimensions that are of interest are accounted for. As an example, the Scottish Index of Multiple Deprivation (SIMD) is an equality measure that has greater relevance for providers in Scotland than those in other UK countries, so could be included as a model covariate in a version of the MEM for providers in Scotland, but omitted from one for providers in other UK countries. The omission of a variable from the MEM model means that any direct effects of the variable on an individual’s likelihood of entering HE are not accounted for, so have no impact on the MEM group assignment of an individual. This is shown in Figure 9, where removal of the gender variable from the updated MEM model has resulted in group assignments of the new version of MEM (MEM-gender) that are independent of gender, and simply reflect the distribution of gender across the entire student population.

Figure 10: Composition of groups 1 to 5 for the updated MEM, and a version of MEM which does not include gender as a model covariate (MEM-gender)



Due to data use restrictions, the specific set of equality variables used in the construction of the updated version of the MEM outlined in this report, limit its use to aggregate-level reporting and analysis. However, the use of an alternative set of equality variables in the MEM model would allow the production of a version of the MEM, suitable for use at an individual level. An example of such a MEM is the version created for the 2019 modernised contextual data service (MCDS) trial, which contains the equality variables gender, POLAR3, Index of Multiple Deprivation, and school type.

Years used for modelling

Due to the changing effects of equality variables on the likelihood to enter HE over time, the use of different data years in the MEM model will result in different modelled likelihoods of entry. The data years to include in the MEM model will depend on the purpose of the MEM being created – to track the progress of groups of students who were historically considered to be disadvantaged or advantaged, or to identify those groups of students who are currently advantaged or disadvantaged (see Section 4).

Form of MEM

Each MEM group contains students with similar values of modelled likelihood of entering HE, but each group encompasses a range of these values, therefore students at a range of disadvantage. Although the presentation of the MEM as a 1 to 5 grouping aids with usability and understanding of the measure, this necessarily results in a loss of resolution in the measurement of disadvantage. The underlying modelled likelihood of entry produced by the MEM methodology allows the possibility for presentation of the MEM at a higher resolution than is afforded by its presentation as groups, so a more precise assessment of disadvantage.

Glossary

Age	This analysis uses country-specific age definitions that align with the cut-off points for school and college cohorts in the different administrations of the UK. For England, ages are defined on 31 August.
Applicant	A person who has made an application in the UCAS Undergraduate system. Counts of applicants include those applying through the main scheme, late applicants direct to Clearing, and Records of Prior Acceptance (RPAs).
Domicile	Declared area of permanent residence.
Entry rate	Number of acceptances from a UCAS application cycle, divided by the estimated base population.
FSM	Free school meals – a means-tested benefit that can be used as an indicator of low income, which has been sourced from the National Pupil Database (NPD). Changes to the coverage of the free school meal indicator in the School Census for the 2013/14 academic year, affecting those applying in the 2016 cycle aged 18, have made it necessary to adjust identification of the FSM group in UCAS' data.
HE	Higher education.
Higher tariff provider	A provider that belongs to the higher tariff group, from the grouping of providers based on the average levels of attainment of their UK 18 year old accepted applicants (summarised through UCAS Tariff points) in recent cycles. The other two groups are medium tariff providers and lower tariff providers. Each group of providers accounted for around a third of all UK 18 year old acceptances in recent cycles.
Multiple equality measure (MEM)	<p>The multiple equality measure (MEM) brings together information on several equality dimensions, for which large differences in the probability of progression into higher education exist. These equality dimensions include sex, ethnic group, where people live (using the POLAR3 classification), secondary education school sector (state or private), and income background (as measured by whether a person was in receipt of free school meals [FSM], a means-tested benefit while at school).</p> <p>These equality dimensions are combined using statistical modelling techniques, and a linked dataset of students in English schools who were aged 18 between 2006 and 2010 (source: National Pupil Database and School Census, Department for Education). The probability of entry to higher education aged 18 is then calculated based on these equality characteristics and their combinations.</p> <p>These probabilities are used to aggregate students into groups, where group 1 contains those least likely to enter higher education ('most disadvantaged' in this context), and group 5 contains those most likely to enter higher education ('most advantaged' in this context).</p>

National Pupil Database (NPD)	The Department for Education's National Pupil Database (NPD) holds a range of information about students who attend schools and colleges in England. For this analysis, a link has been formed between individuals attending state schools and colleges in the NPD at 15 years old, and those who were 18 years old in UCAS' data three years later.
POLAR3	Developed by HEFCE, POLAR3 classifies small areas across the UK into five groups, according to their level of young participation in HE. Each of these groups represents around 20 per cent of young people, and is ranked from quintile 1 (areas with the lowest young participation rates, considered as the most disadvantaged) to quintile 5 (highest young participation rates, considered most advantaged).
Provider	A higher education provider – a university or college.
SIMD	The Scottish Index of Multiple Deprivation (SIMD) identifies small area concentrations of multiple deprivation across all of Scotland, providing a relative measure of deprivation among small areas (data zones).
UK	United Kingdom. Excludes the Channel Islands and the Isle of Man.