



# THE SUTTON TRUST CHAMPIONS SOCIAL MOBILITY THROUGH PROGRAMMES, RESEARCH AND POLICY INFLUENCE



## 25 years of the Sutton Trust

A new data asset on equality and fair access and how it can inform on long-term trends.

Dr Mark Corver, Founder and MD, dataHE  
Key findings for Sutton Trust, 13 April 2023

# What this project does

---



Annual Report 98

## 25 years of the Sutton Trust

---

The equality and fairness of young university entry:  
What the data says about how it has changed since the Trust started, and what the future holds

Revised proposal for Sutton Trust  
Mark Corver, dataHE. February 2022 (additions April 2022)

The Trust has always been distinctive by being anchored in data and evidence.

But the public data on equality and fair access is sparse, fragmented and increasingly lost.

This project does two things:

- (i) Recovers these data into a analytically strong data base
- (ii) Uses it to take long-term view on equality and fair access

# Some headlines

---

You have a data asset containing processed aligned public data on equality and fair access back to 1997.

Distributional fair access has been stubborn.

Trust focus, school type has seen some proportional improvements, though the “3,000” figure is up.

For areas and social class, no material improvement.

Young application and entry rates have both increased, with real reductions in inequality by area groupings.

But entry inequality on other dimensions, sex and ethnic group has been growing.

# What we will be covering today

---

1. Data asset

2. Fair access

3. Equality

4. Thoughts

---

# (1) Data asset

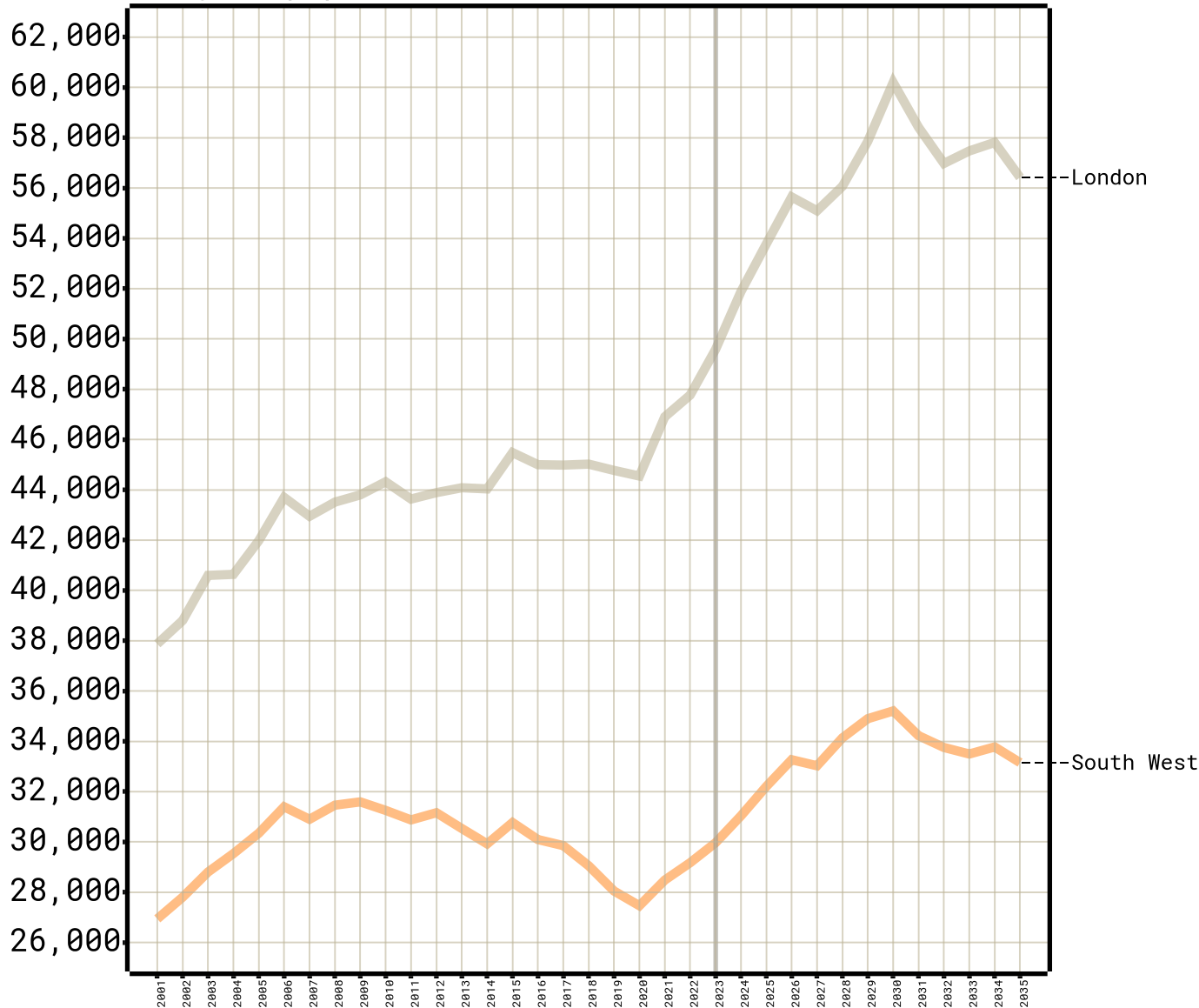
---

# What has been created?

| Dimension     | Population | Entrants (sector) | Entrants (university) |
|---------------|------------|-------------------|-----------------------|
| Neighbourhood | Yes        | Yes               | Yes                   |
| Region        | Yes        | Yes               | Yes                   |
| Ethnic group  | Yes        | Yes               | Yes                   |
| Sex           | Yes        | Yes               | Yes                   |
| School type   |            | Yes               | Yes                   |
| Social class  |            | Yes               | Yes                   |

# Populations: Region and sex

18 year old women by selected region  
School aligned population



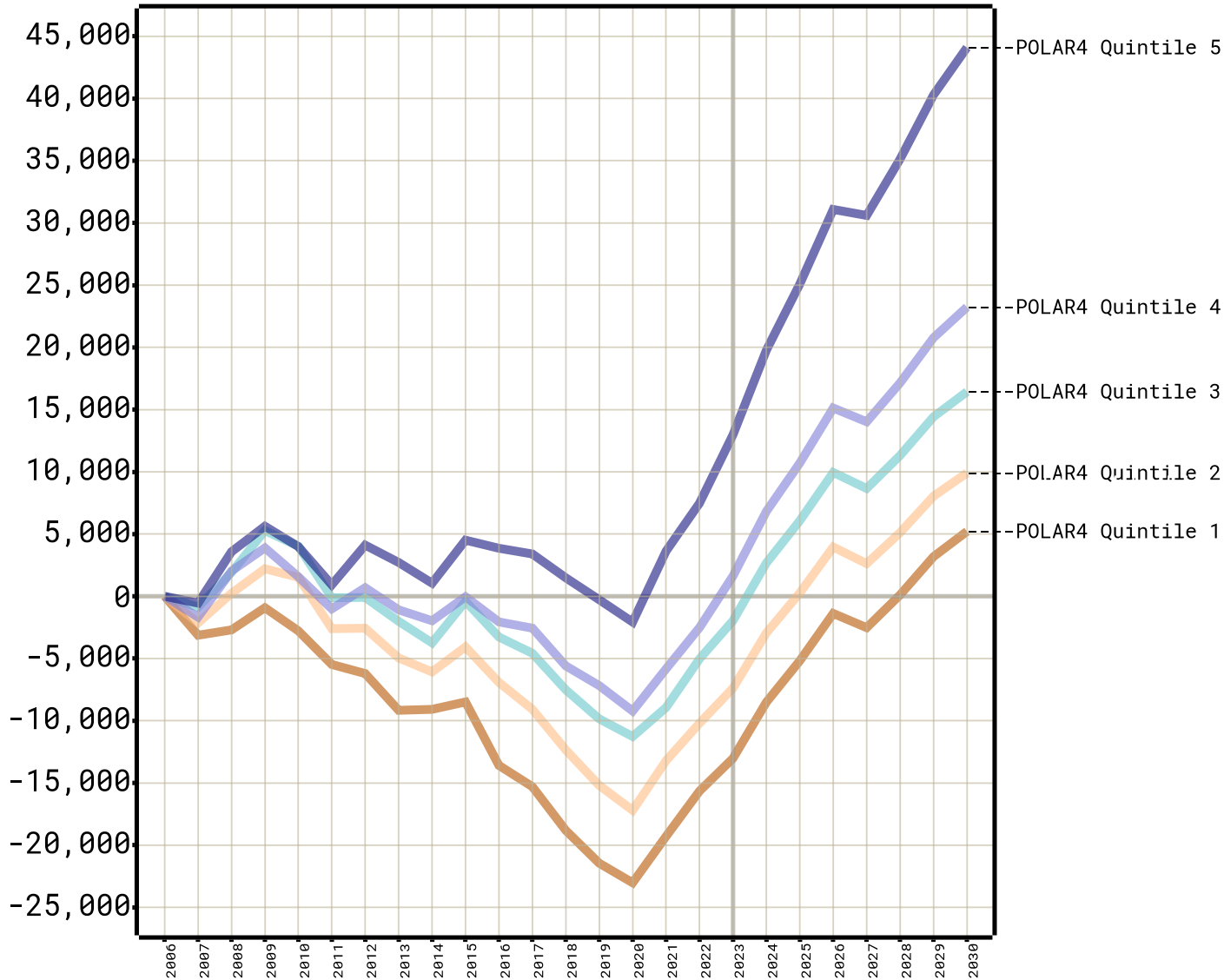
Region and sex come from our core work on populations.

ONS MYEs and projections (pre-pandemic referenced), aligned to match university cohort definitions.

Analysis by www.datahe.uk \* Student data fromucas.com, populations from un.org \* 1s1s1

# Populations: POLAR

(Difference from 2006) 18 year olds by POLAR group England



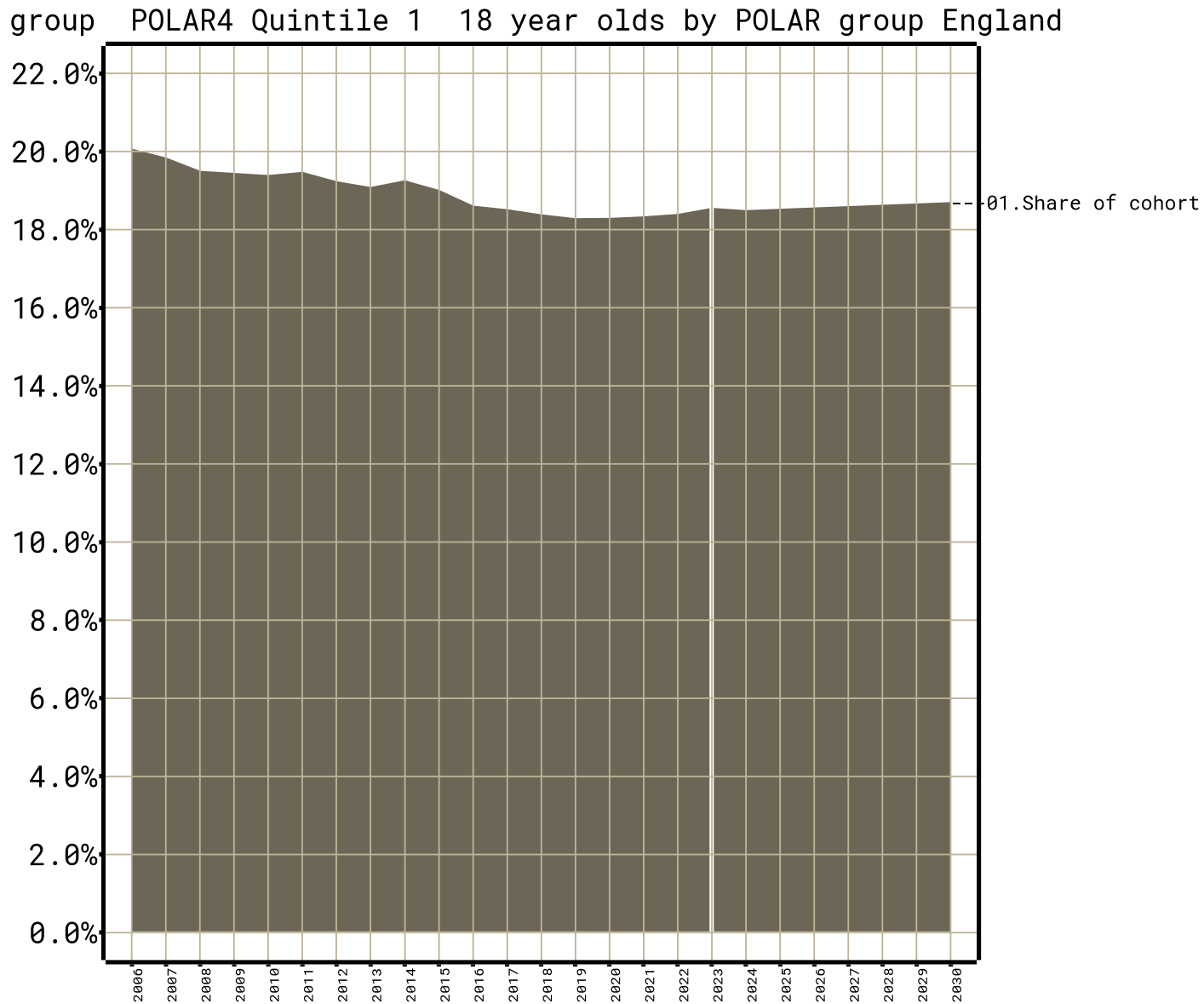
Analysis by www.datahe.uk \* Student data fromucas.com, populations from un.org \* p4eprI2

For POLAR we start with published national estimates.

We model (forward, and back) by projecting cohort proportion changes.



# Population shares: POLAR4 Q1



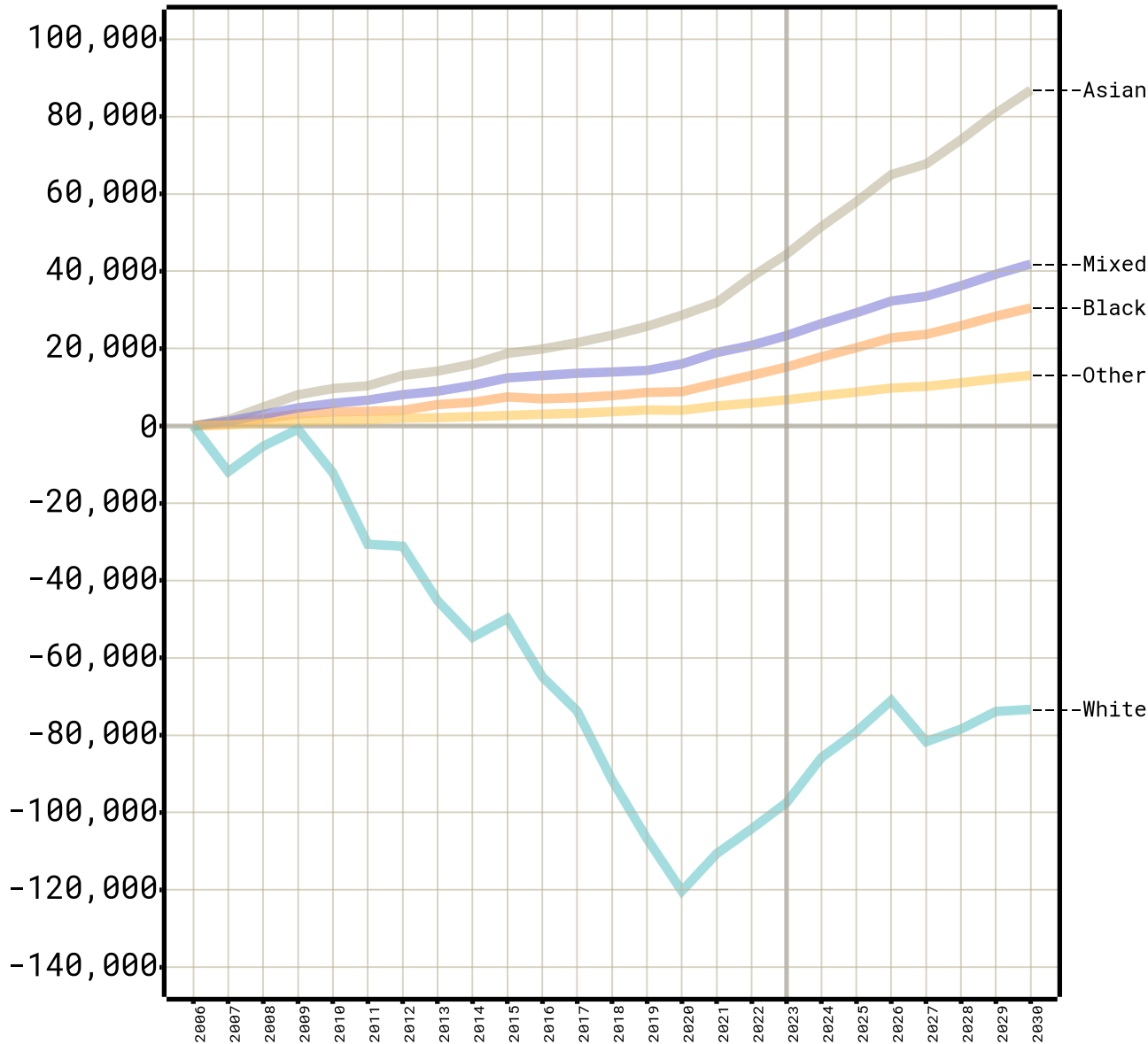
For fair access we look at proportions of entrants - population can matter here.

- For POLAR4 Q1 can take as near constant over the period, it is not a major influence.

Analysis by [www.datahe.uk](http://www.datahe.uk) \* Student data from [ucas.com](http://ucas.com), populations from [ONS/un.org](http://ONS/un.org) \* s4eG1C1

# Populations: ethnic group

(Difference from 2006) 18 year olds by Ethnic group England



Analysis by www.datahe.uk \* Student data fromucas.com, populations from un.org \* etepr12

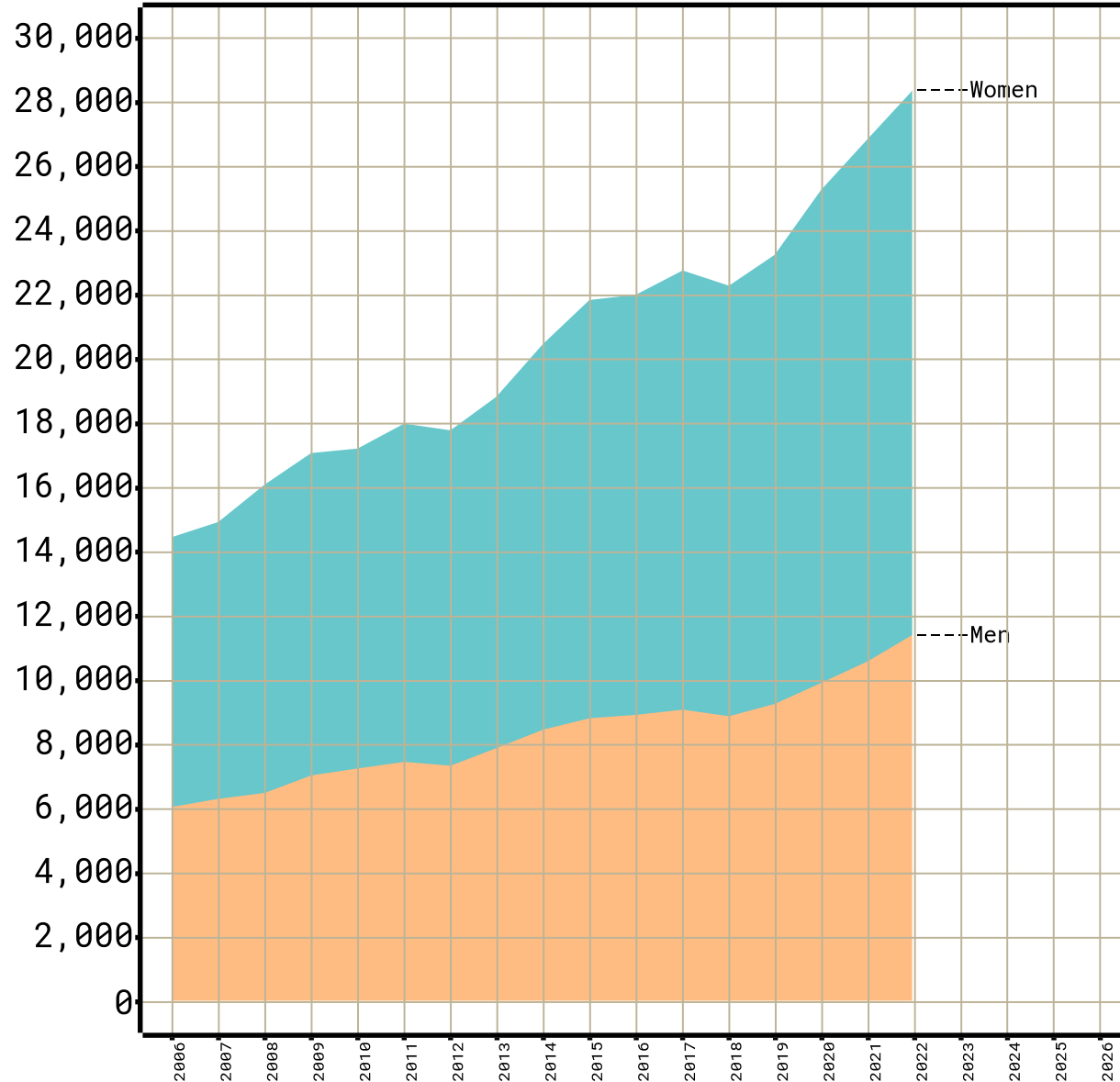
For ethnic group we start with published whole-population summary-level (i.e. not state school) UK estimates.

APS typical-mother-age data then used to redistribute to country

Modelling (forward and back) by cohort share projection.

# Entrants (sector): equality

P4 Q1 entrants by sex England  
MLX entrants, cumulative



Analysis by [www.datahe.uk](http://www.datahe.uk) \* Student data from [ucas.com](http://ucas.com), populations from [un.org](http://un.org) \* q1e1c1

The long-run sector equality data comes from the national UCAS end of cycle data.

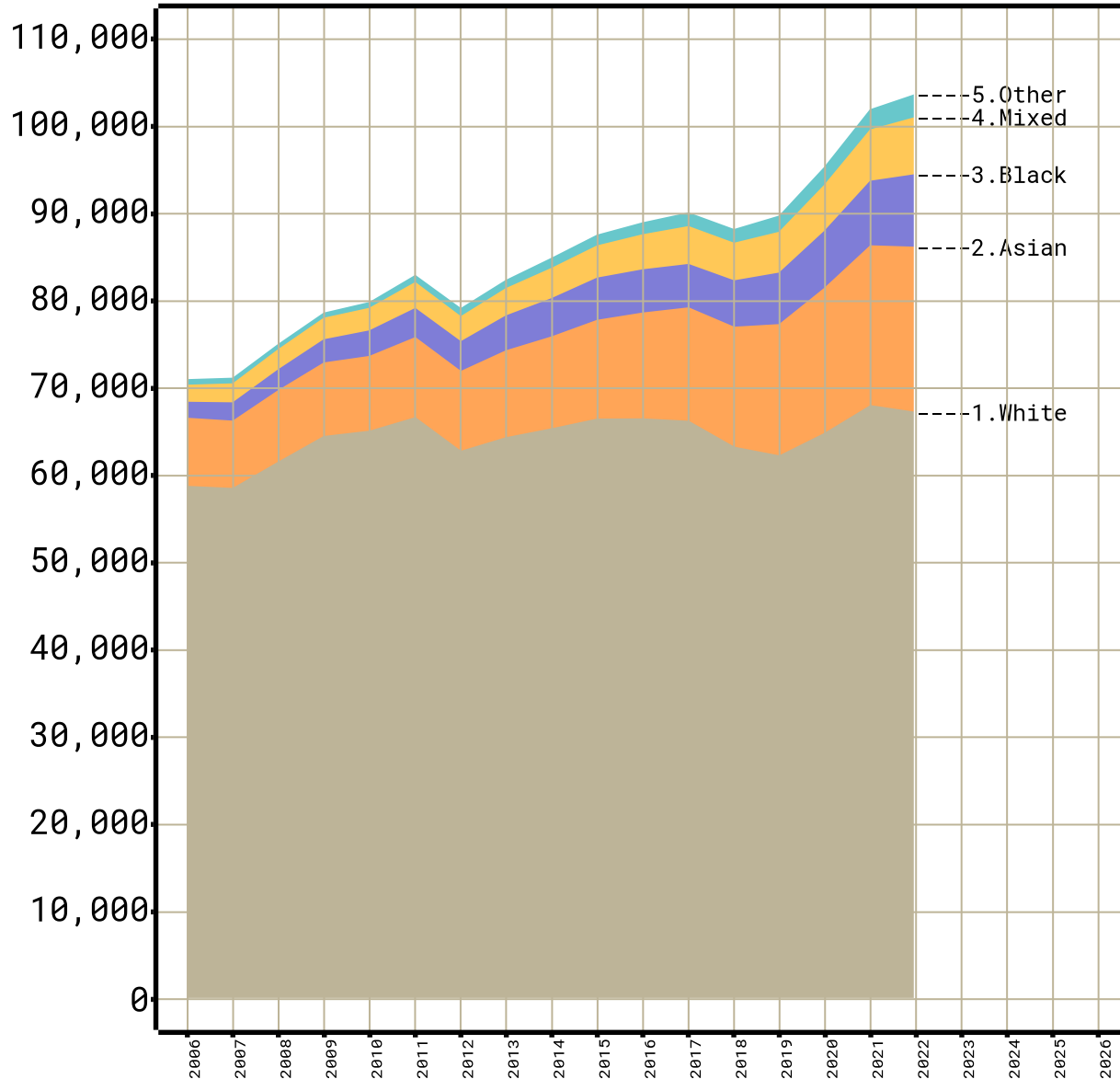
We process this to remove 'RPA's to get the best series.

We measure both demand and outcomes.

2006–2022 (17 years)

# Entrants (sector): equality

English entrants (men) by ethnic group  
MLX entrants, cumulative



Analysis by www.datahe.uk \* Student data fromucas.com, populations from un.org \* see1C1

The data files have further resolution by country and sex.

These are available to you in the data asset.



# Entrants (university): fair access

Widening participation summary: UK Performance Indicators

22 February 2022

## + Important - Changes to the UKPIs from 2020/21

Changes to definitions mean that the latest UKPI data differs from previous years' data. [Expand this box for more information.](#)

## + Impact of COVID-19 on UKPIs

There are known areas in which the pandemic has had some impact on the 2019/20 and 2020/21 Student data. [Expand this box for more information.](#)

This summary contains overview information on: [State school marker](#) | [Low participation neighbourhoods](#) | [Disabled Students' Allowance](#) for 2015/16 and 2020/21.

## UK Performance Indicators: Widening Participation



Open data licence: [CC-BY-4.0](#)

## UK Performance Indicators 2002/03: Widening participation

Release date: 30 September 2004

Our widening participation Performance Indicators explore the extent of specific groups' participation in higher education. Tables document the percentage of students:

- From state schools or colleges
- From specified socio-economic classes
- From low-participation neighbourhoods

HIGHER EDUCATION  
FUNDING COUNCIL  
FOR ENGLAND



HIGHER EDUCATION  
FUNDING COUNCIL FOR WALES  
CYMRU CYLLIDU ADYFFUS UWYCHYDOL



Scottish higher education  
funding council

Report 99/66

December 1999

## Performance Indicators in Higher Education in the UK

Note: This report is available in three parts. There are separate files for the [main text](#) of the report, the [Tables of indicators](#) and the [Annexes](#). (600K). Printed copies of the report are [available from HEFCE](#), price £25.

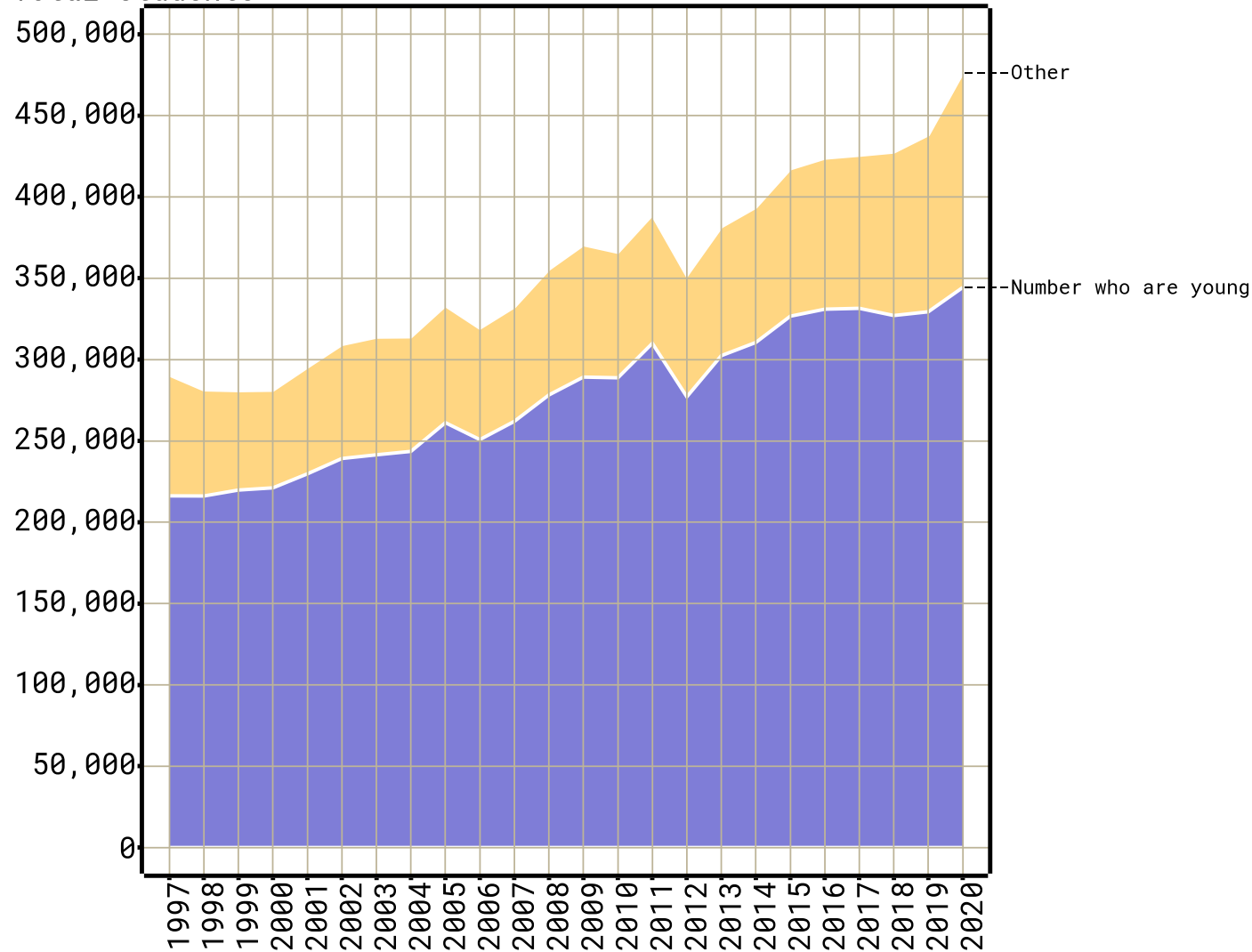
We use the official Performance Indicator sequence for long-run fair access.

Very difficult to recover. Multiple sources. Multiple formants. Specialised archives for earlier years.

Covers 1997 to 2020 (23 years)

# Young entrants in the PIs

UG/T1b Low participation: UK  
Total students



dataHE - Confidential for the Sutton Trust (20230203) tot01

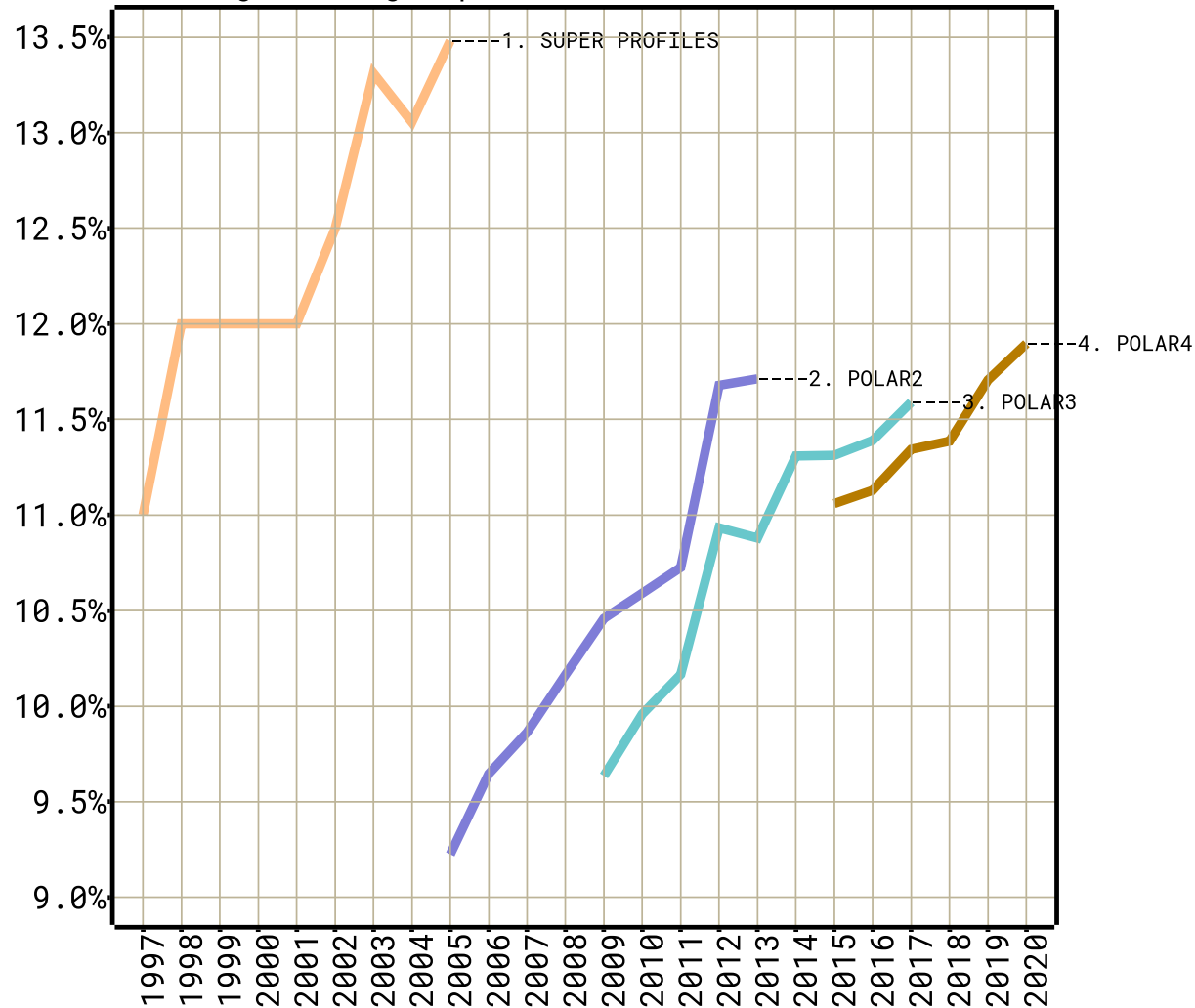
UK domiciled  
'young' entrants  
(aged up to 20).

There are three  
measures for  
young:

- Neighbourhood
- School type
- Social Class

# Neighbourhood: raw measures

FD/T1a/English providers/Low participation: POLAR methods  
05. Percentage from group



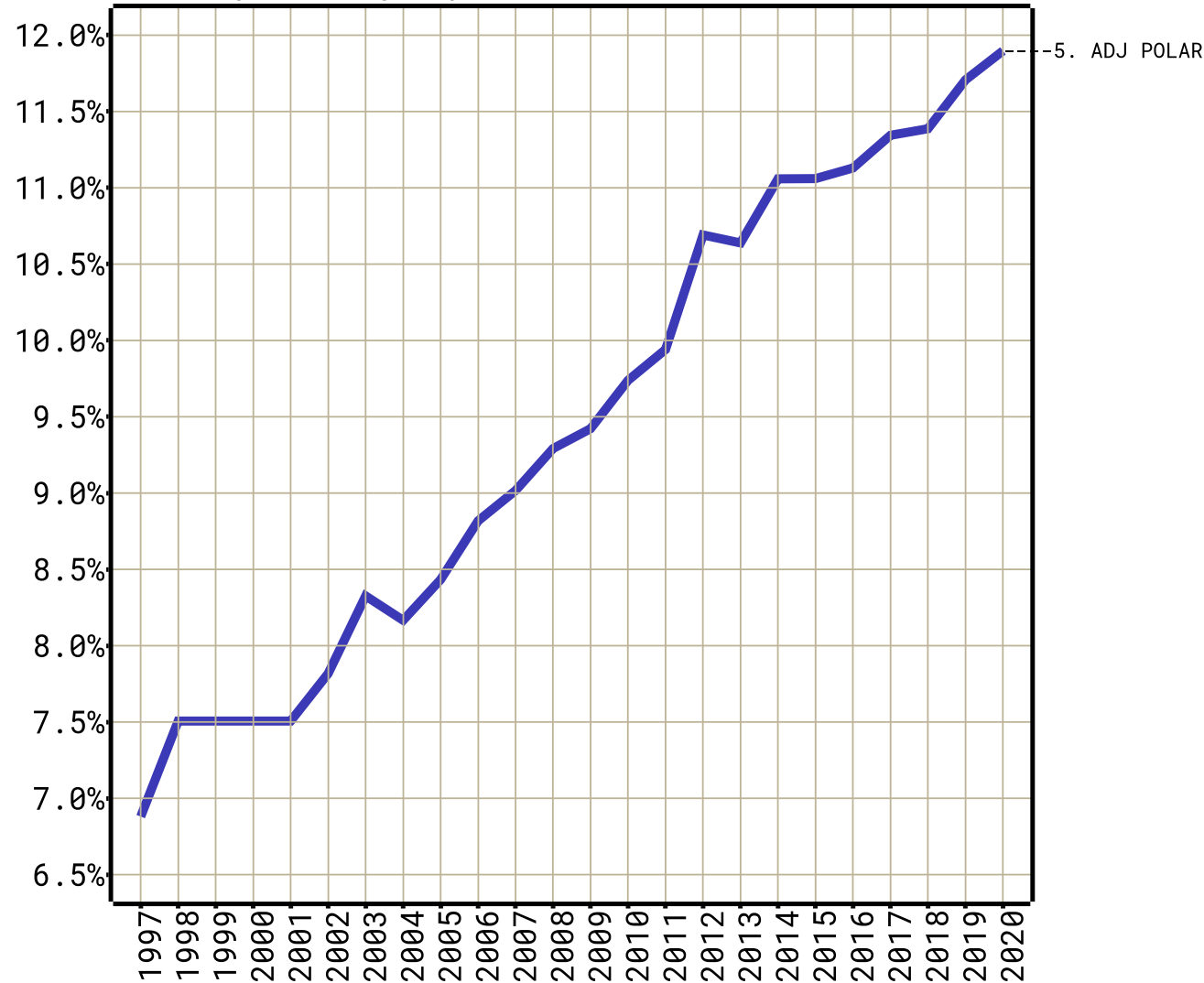
dataHE - Confidential for the Sutton Trust (20230203) RPR1rps1

Four measures through time.

A **narrowly based** measure in the PIs – e.g. just POLAR4 Q1 (20% of population)

# Neighbourhood: consistent share

FD/T1a/English providers/Low participation: POLAR methods  
05. Percentage from group



dataHE - Confidential for the Sutton Trust (20230203) RPR2rpS1

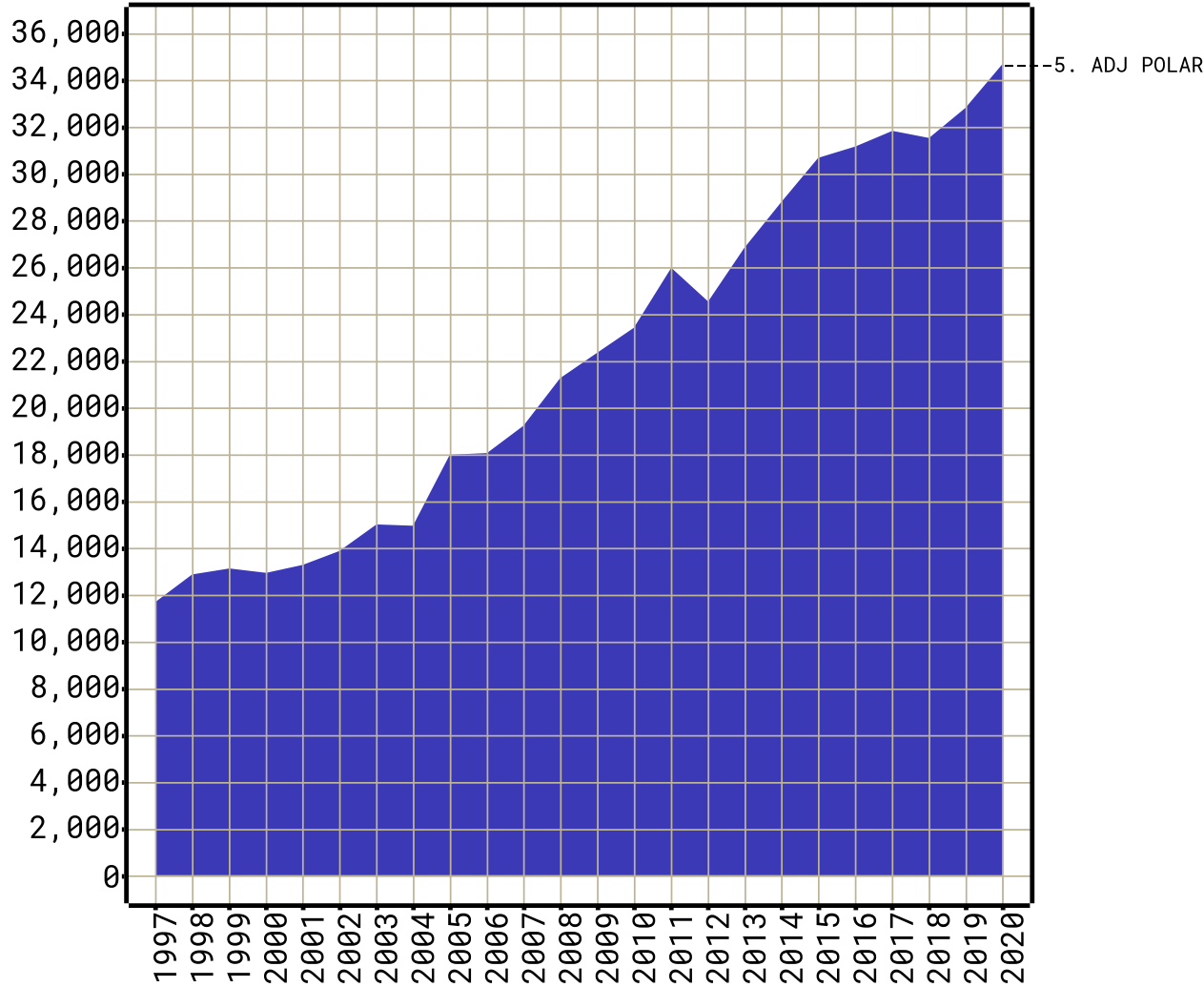
Using data overlaps and models we are able to remap the old definitions to the most recent (Q1 POLAR 4) to give a consistent entrant share measure

(Scottish providers dropped here as POLAR data for them suppressed in PIs)



# Neighbourhood: consistent numbers

FD/T1a/English providers/Low participation: POLAR methods  
02. Number from group, cumulative



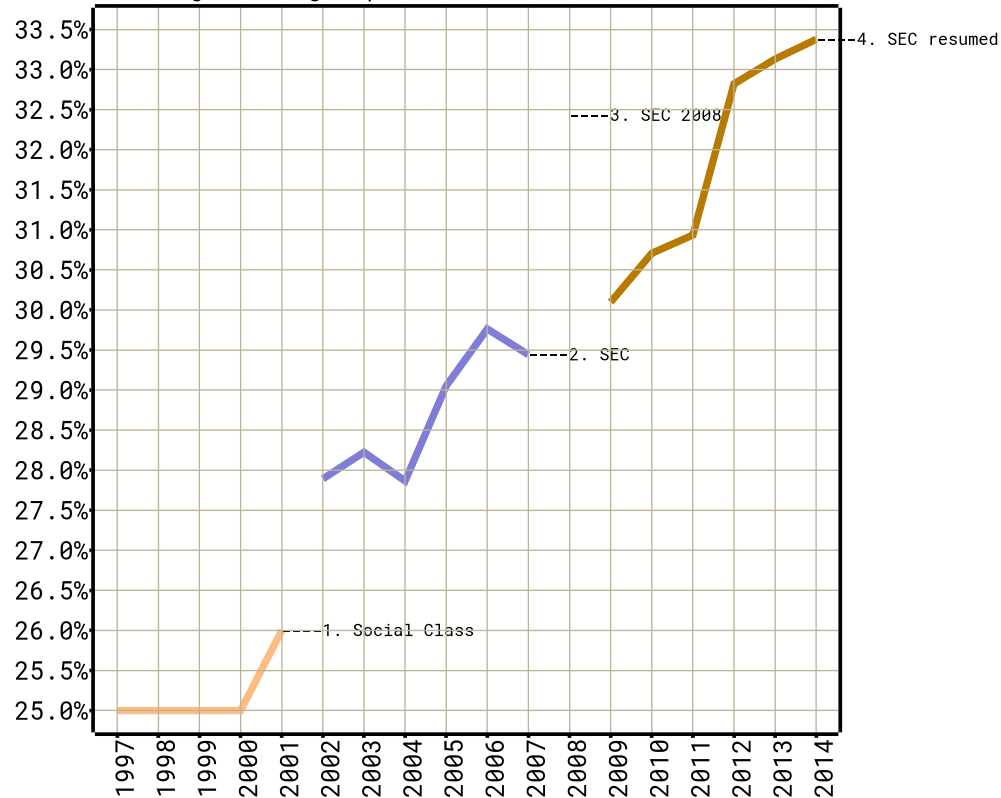
dataHE - Confidential for the Sutton Trust (20230203) NPR2rnc1

This also gives us a count of entrants on the same consistent basis.

All these items are in the data asset.

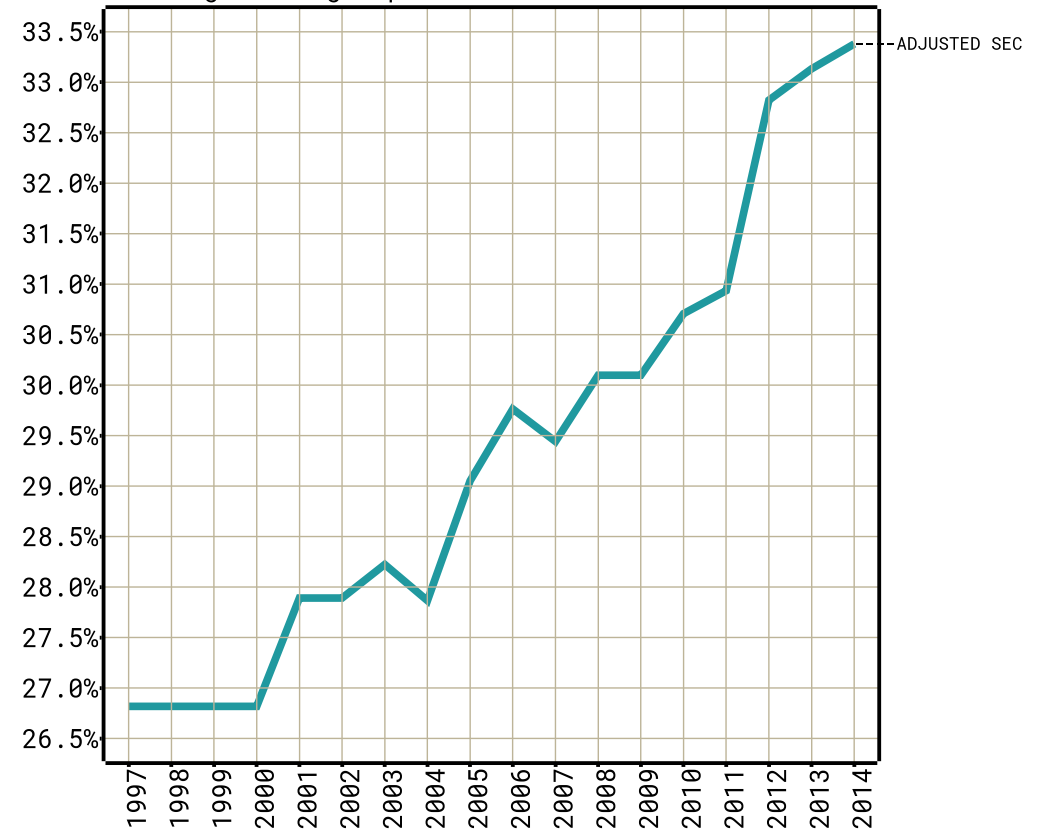
# Social class: consistent share measure

FD/T1a/English providers/Low social class: social\_class\_type  
05. Percentage from group



dataHE - Confidential for the Sutton Trust (20230203) RCR1rpS1

FD/T1a/English providers/Low social class: social\_class\_type  
05. Percentage from group

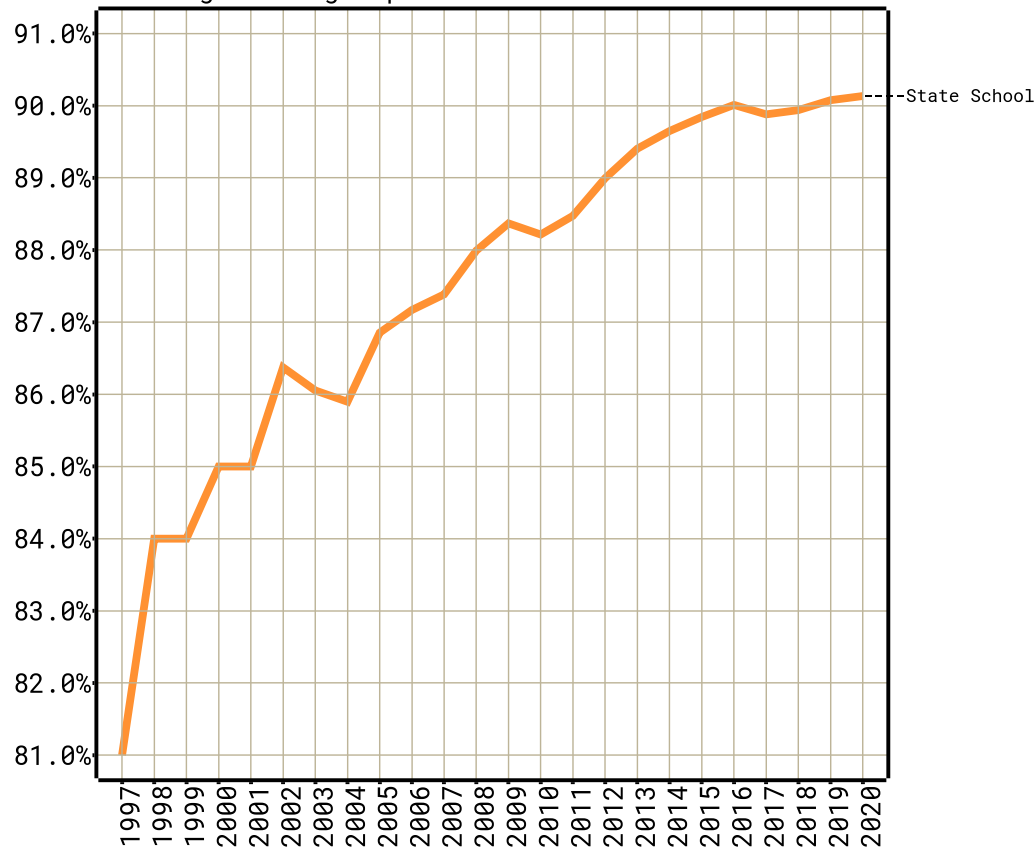


dataHE - Confidential for the Sutton Trust (20230203) RCR2rpS1

A similar process is used to get a consistent measure for parental occupational background ('social class'). These data only run 1997 to 2014 (18 years). **Balanced measure** in the PIs (around 40% of population)

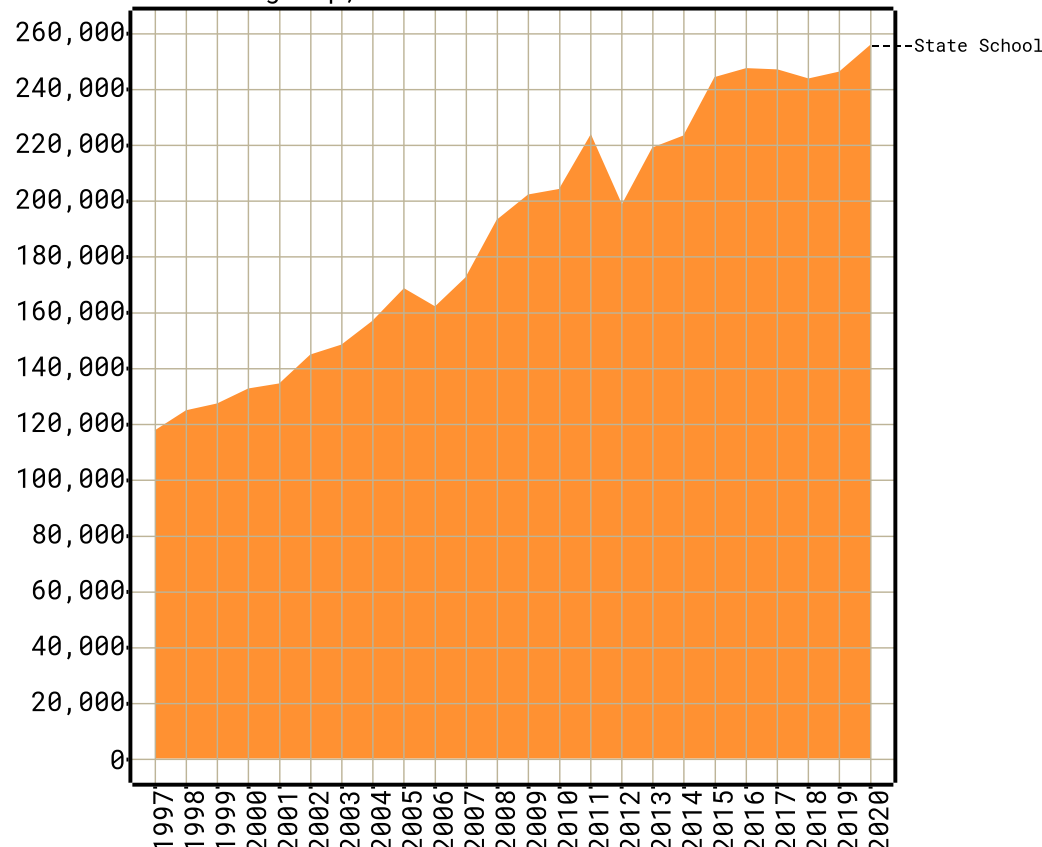
# School type: broad based

FD/T1a/English providers/State schools or colleges  
05. Percentage from group



dataHE - Confidential for the Sutton Trust (20230203) RSR1rps1

FD/T1a/English providers/State schools or colleges  
02. Number from group, cumulative

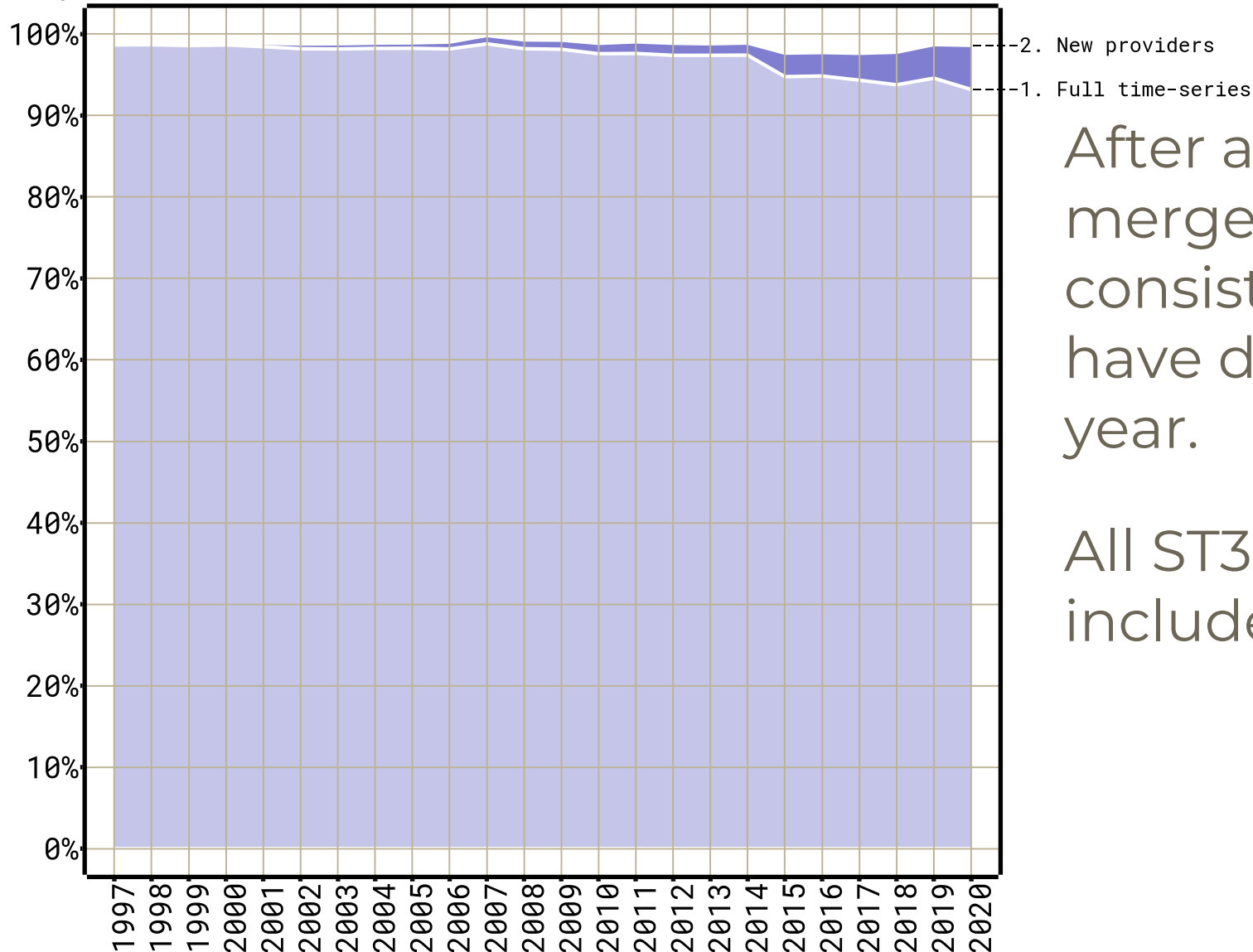


dataHE - Confidential for the Sutton Trust (20230203) NSR1rnc1

School type definition constant. But school system structure has changed over this period. Hard to know from the aggregate data. But is a **Broad based measure** (c.90+% population).

# Entrants: consistent universities

FD/T1a Low participation: new providers or with full time-series  
Proportion of total students



After allowing for mergers we create a consistent subset that have data for every year.

All ST30 and RG are included.

dataHE - Confidential for the Sutton Trust (20230203) sbst00

---

## **(2) Fair Access**

---

# Why is fair access important?

---

Not always agreement about:

- (i) What the 'right' overall level of university entry should be
- (ii) or whether certain universities should obtain 'better' outcomes for their students in terms of opportunities

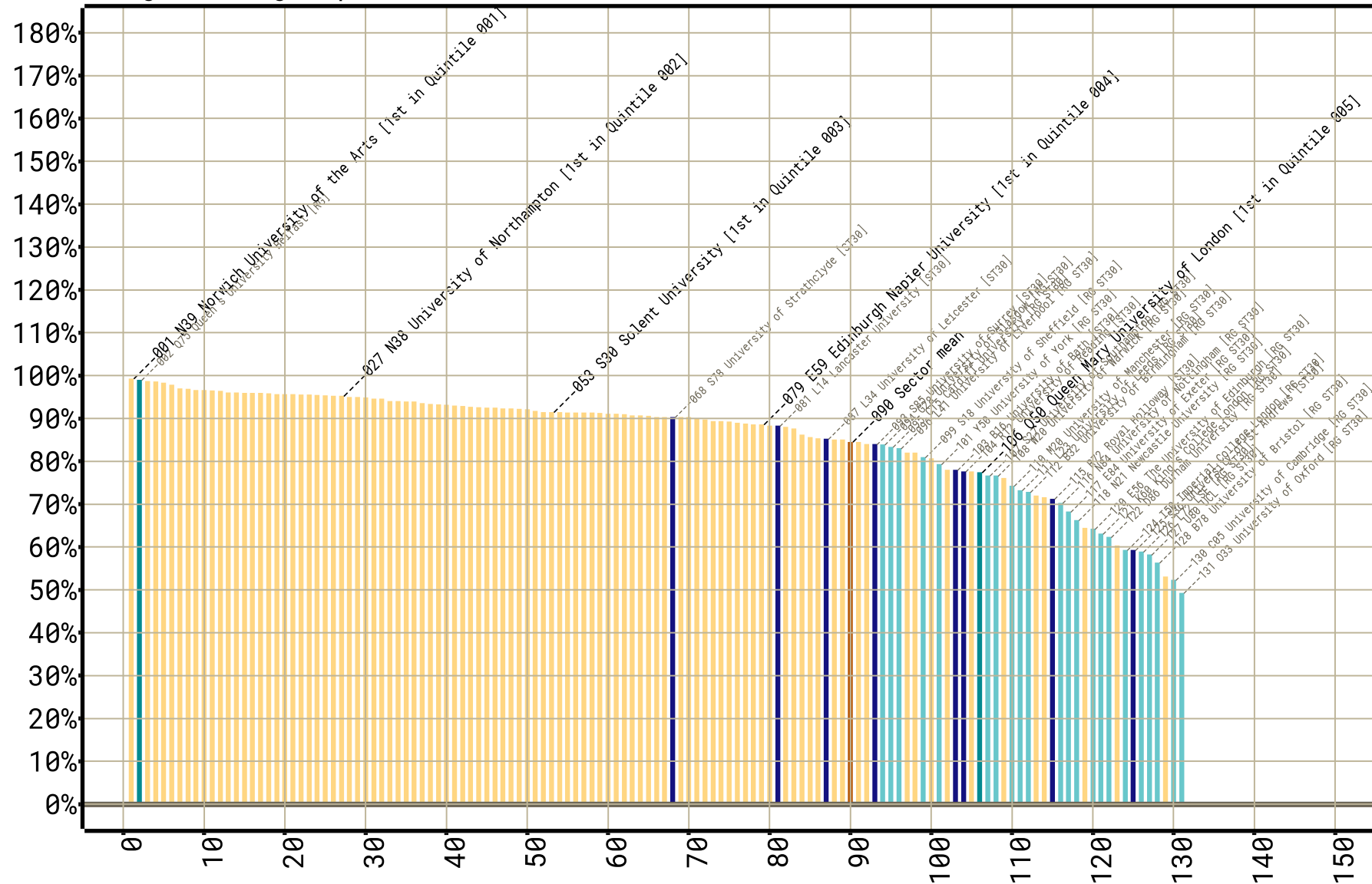
But most agree that the proportions of entrants in certain groups shouldn't vary too much across universities. Or, at least, without good reason (like entry qualifications or subjects).

“Fair access” gets at these issues.

Defining approach of the Trust in early days.

# 1997-1999: State school distribution

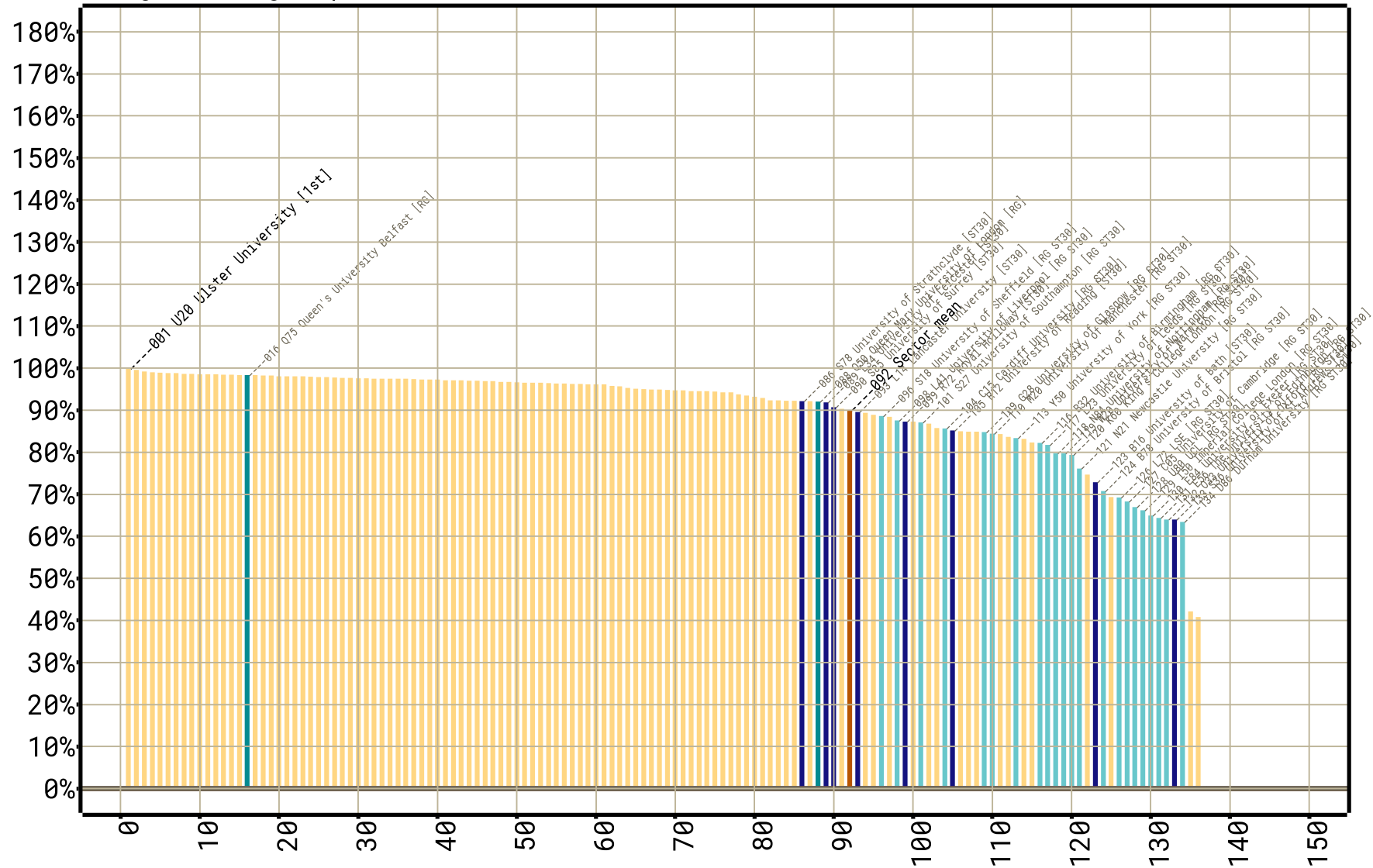
FD/T1a State schools or colleges: 1997-1999 mean  
Percentage from group



dataHE - Confidential for the Sutton Trust (20230203) ndssq111

# 2018-2020: State school distribution

FD/T1a State schools or colleges: 2018-2020 mean  
Percentage from group

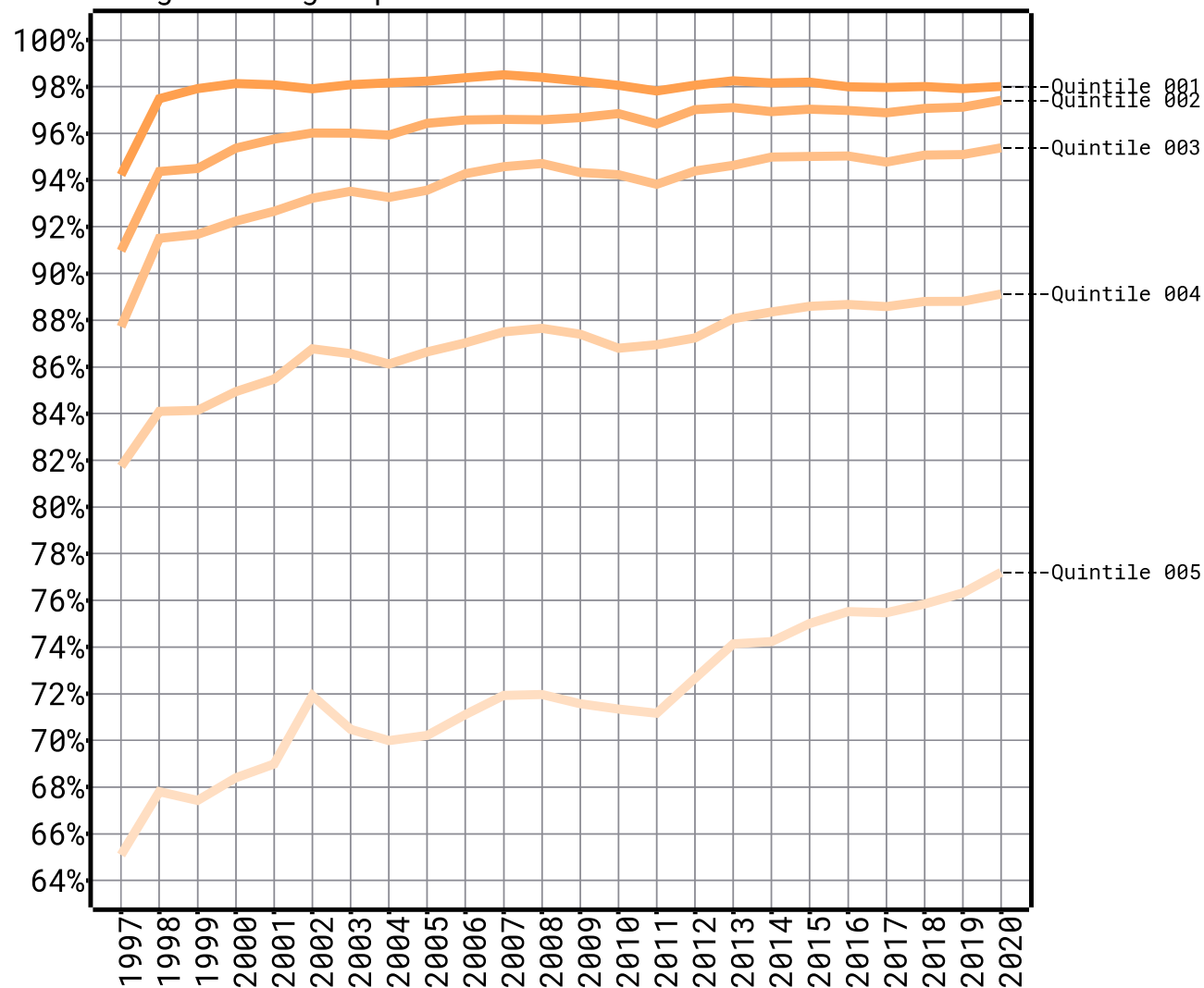


dataHE - Confidential for the Sutton Trust (20230203) ndssqlx1



# Tracking state school proportions

FD/T1a State schools or colleges: provider quintiles  
Percentage from group



dataHE - Confidential for the Sutton Trust (20230203) ntasql11

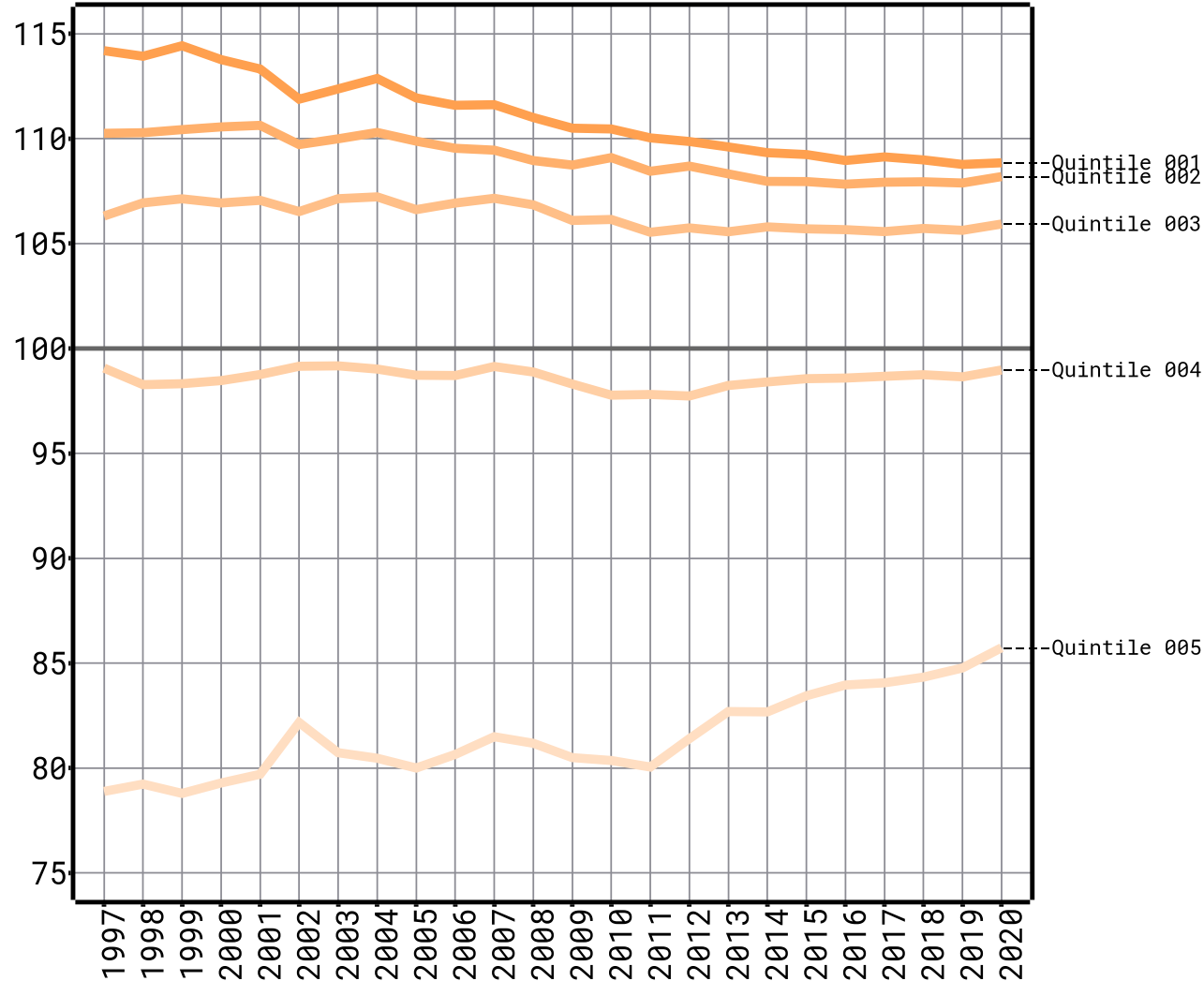
Using the late 1990s groupings to see what has changed.

Proportions of state pupils have increased across all types of universities.

But are already so high at many can not move much further.

# Indexing back to the average

FD/T1a State schools or colleges: provider quintiles  
Percentage from group (wgt avg=100)



dataHE - Confidential for the Sutton Trust (20230203) ntaSQL11

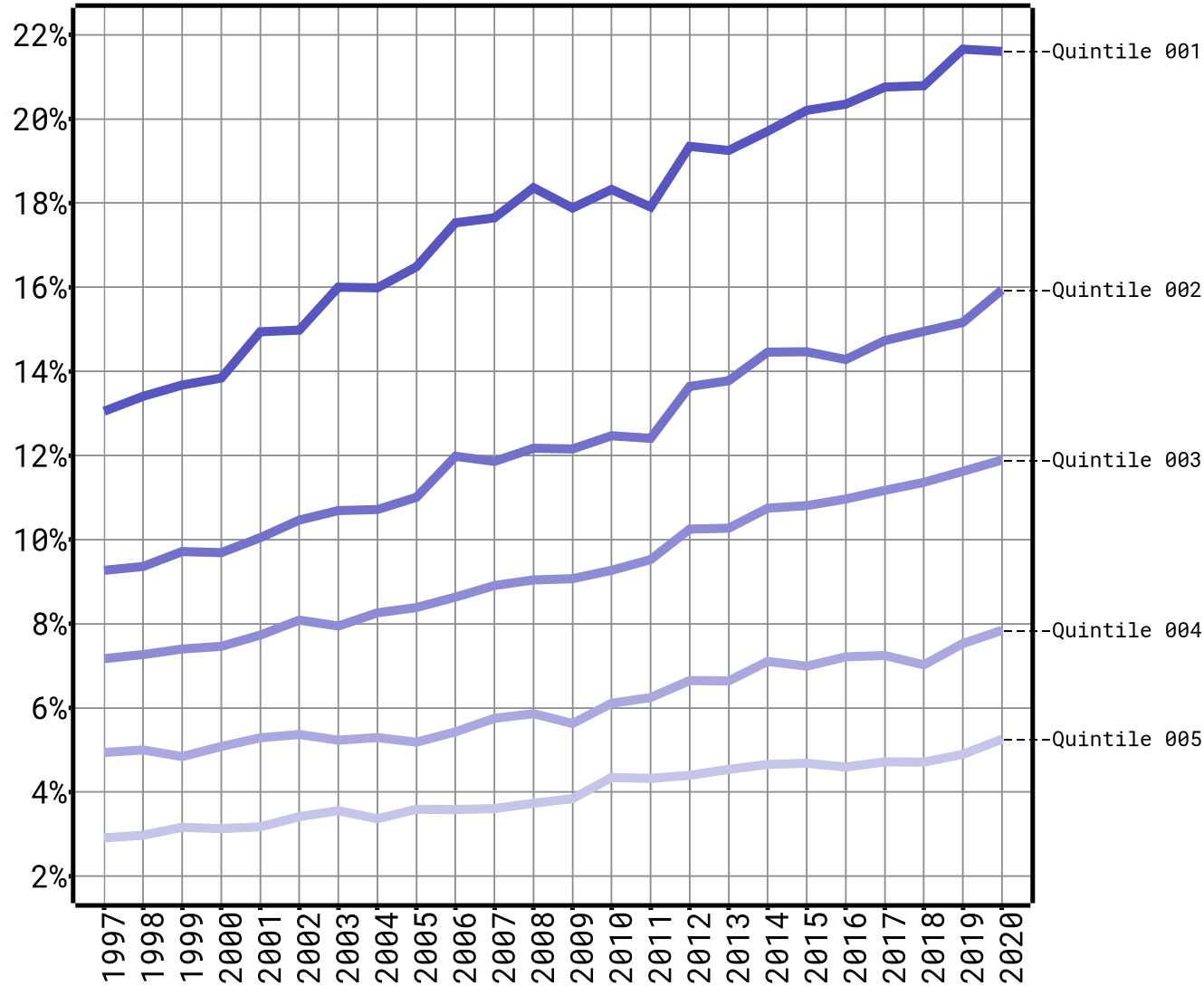
Taking out the effect of increasing state school numbers.

See only a modest relative move in fifth of institutions with lowest proportion of state. From 80% to 85% of sector average

Measure affected by saturation.

# Tracking neighbourhood proportions

FD/T1a Low participation: provider quintiles  
Percentage from group



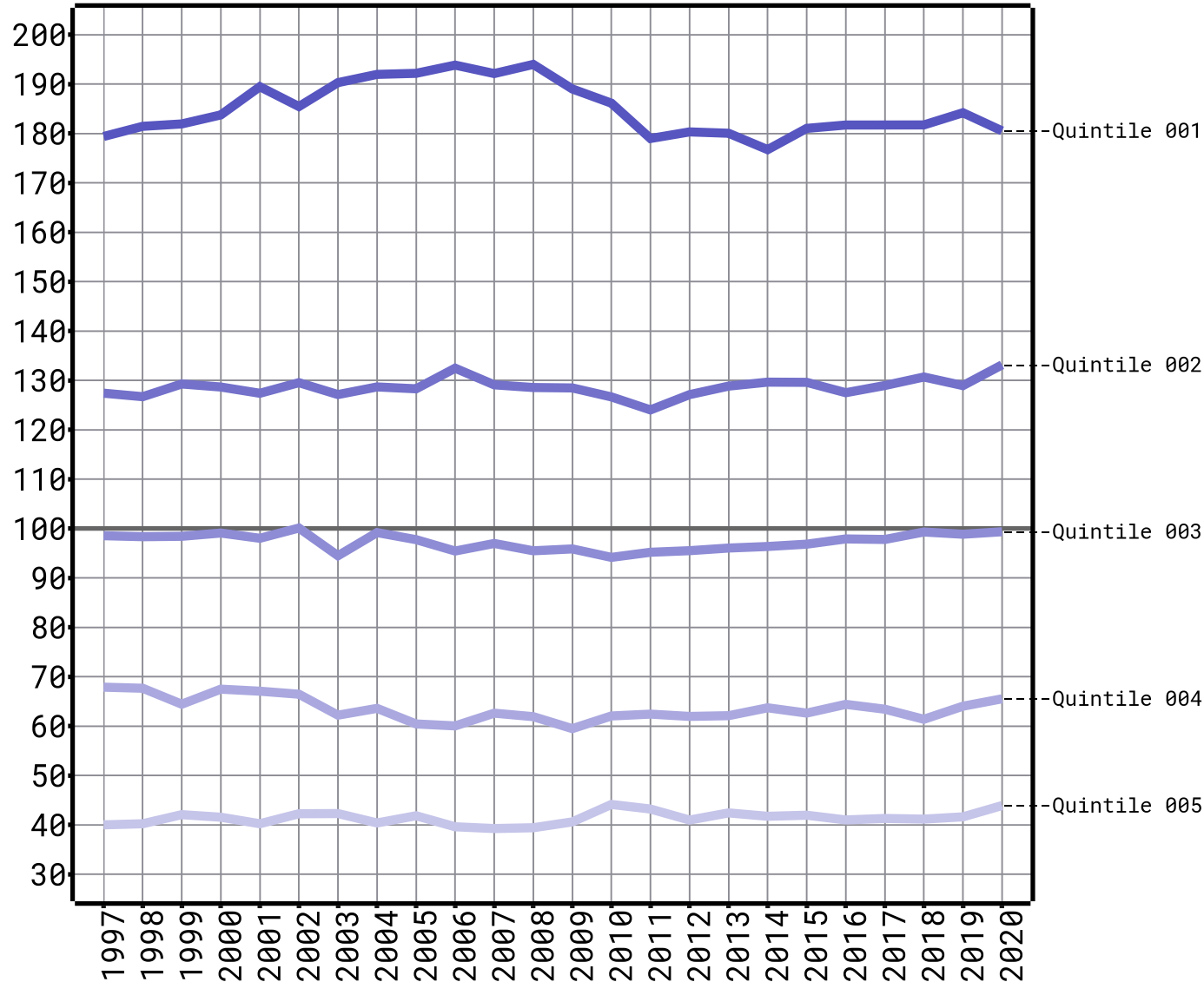
Opposite type of measure. Very narrow.

Increasing everywhere, including those universities with lowest proportions in late 1990s.

dataHE - Confidential for the Sutton Trust (20230203) ntapq111

# But no change in relative distribution

FD/T1a Low participation neighbourhoods: provider quintiles  
Percentage from group (wgt avg=100)



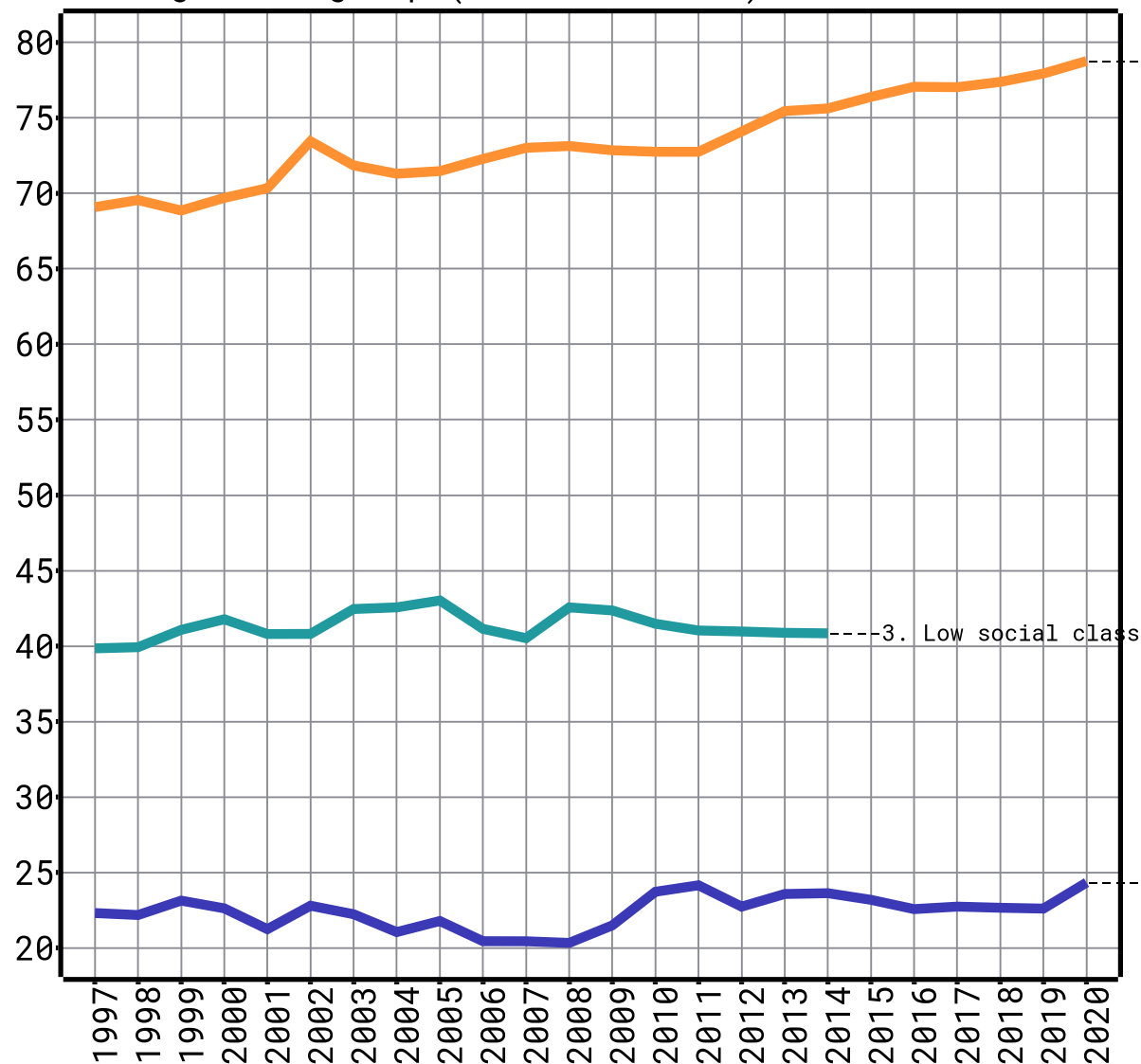
The relative distribution of this under-represented group across universities has not materially changed.

The most 'exclusive' universities were at 40% of sector in 1997, and the same in recent years.

dataHE - Confidential for the Sutton Trust (20230203) ntapq111

# Overall change across the measures

FD/T1a Summary statistic: Quintiles 5 of providers  
Percentage from group (Quintile 1=100)



dataHE - Confidential for the Sutton Trust (20230203) gapq1

2. State schools or colleges

Here indexed to providers with highest representation.

The proportion from state schools has improved. But this is a saturated measure.

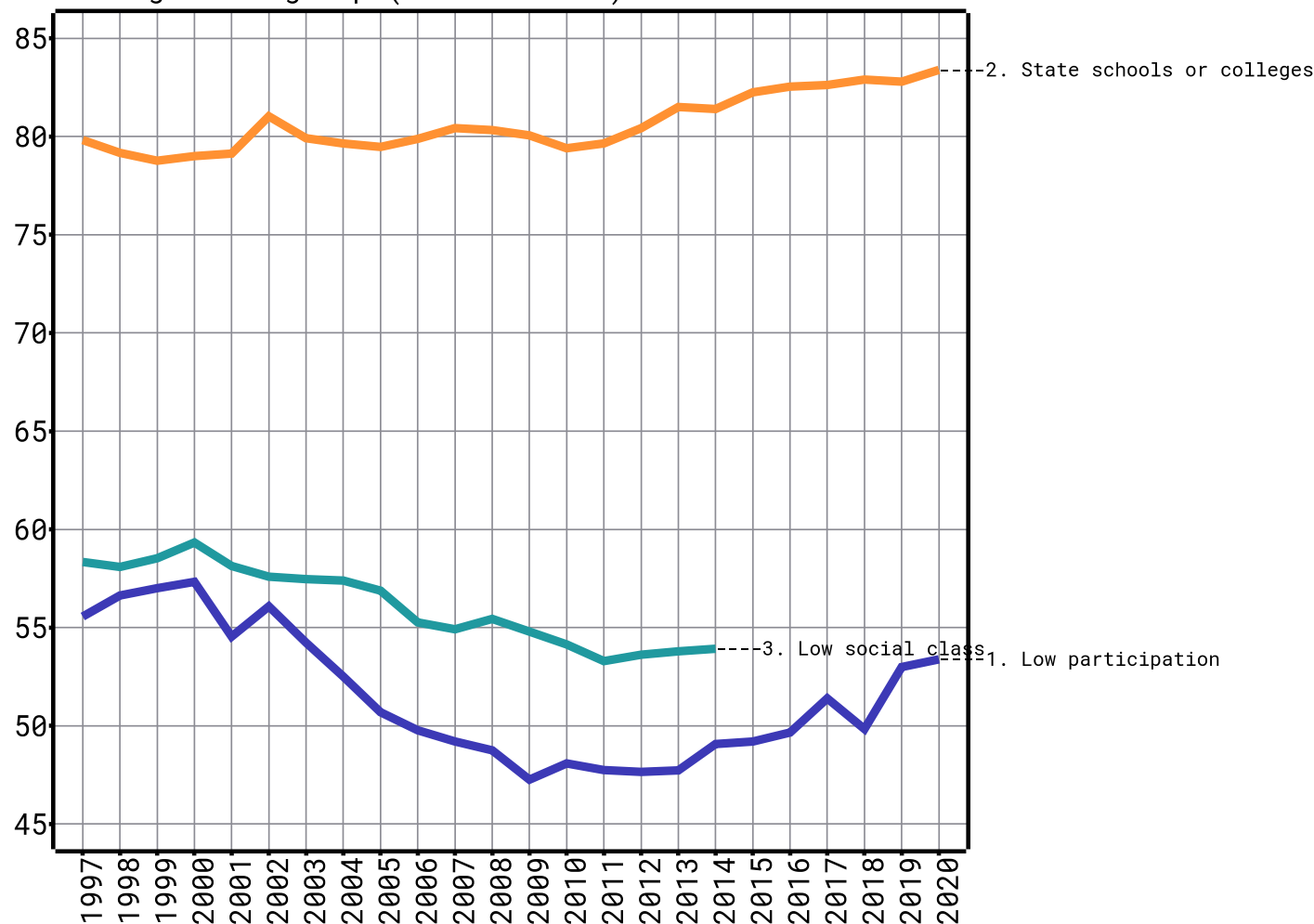
3. Low social class

1. Low participation

Social class and low participation measures show no real change.

# What about the ST30 group?

FD/T1a Summary statistic: Sutton Trust 30  
Percentage from group (non-ST30=100)



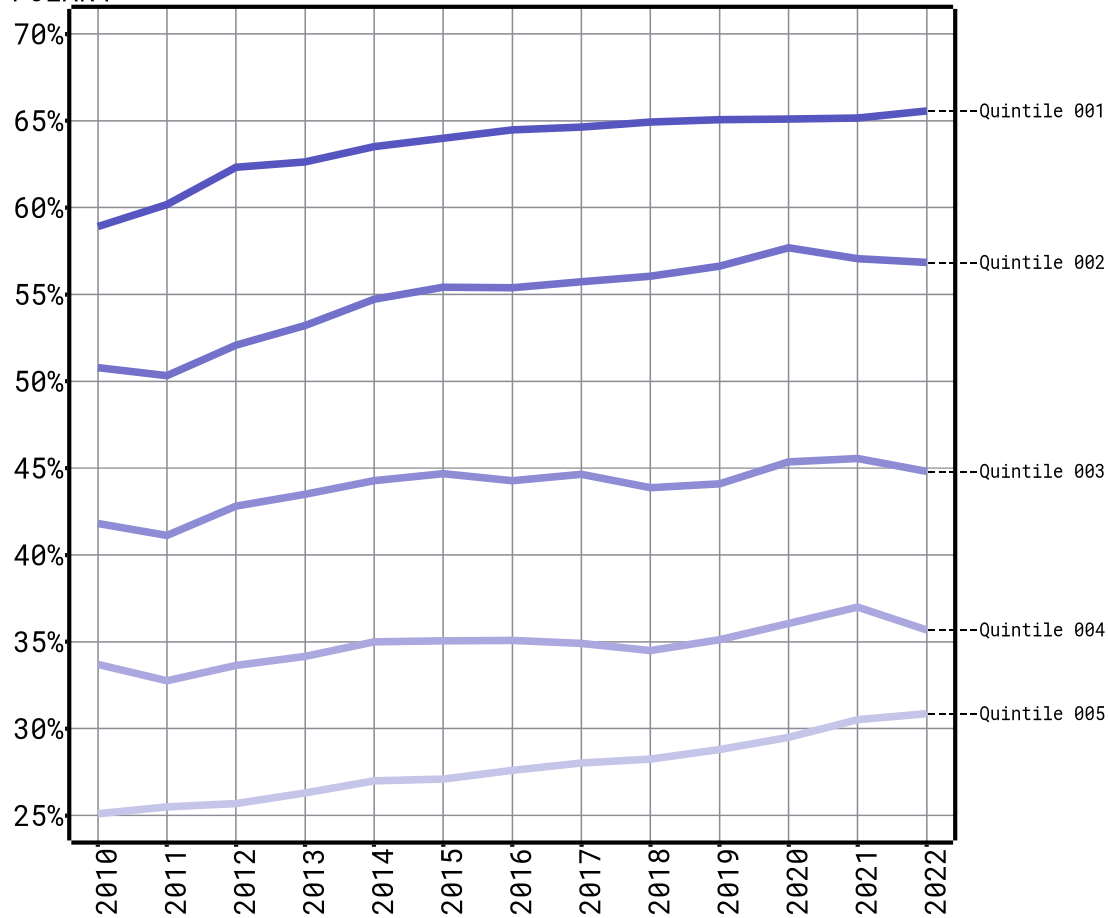
dataHE - Confidential for the Sutton Trust (20230203) gapst1

Rather than use quintiles against the mean, we can index a group against the *rest* of the sector.

For the ST 30, state school has improved but occupation and area background no better than late 1990s.

# Post-pandemic: the UCAS data

Percentage from underrepresented group by provider quintile  
POLAR4



dataHE - Confidential for the Sutton Trust (02/02/23) q1p2

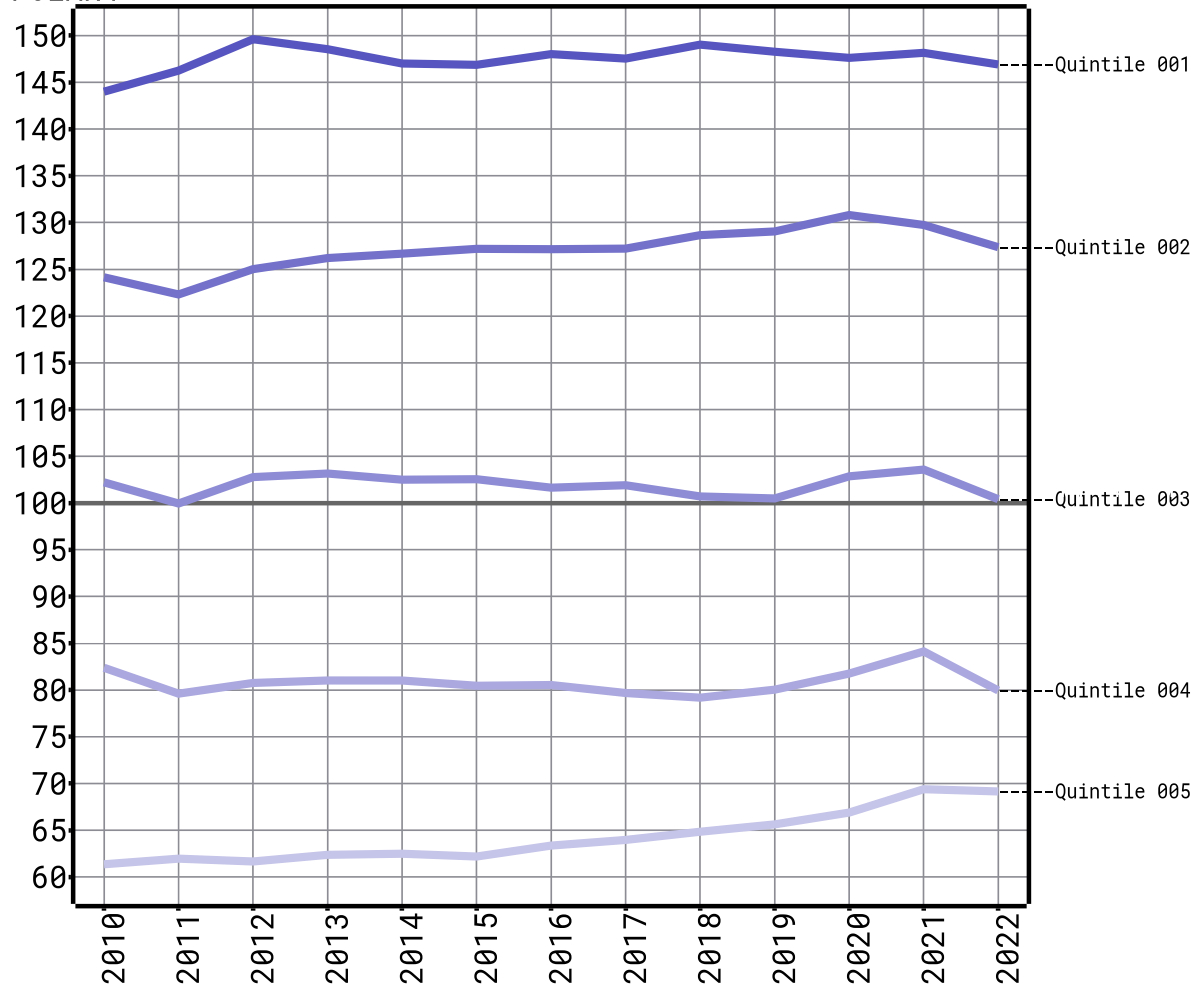
We can create a more up to date distributional picture using UCAS data.

Just age 18, and a wider group (1,2 and 3). Smaller set of universities, but consistent across period.

“PI style” proportion up for the universities where most under-represented – including 2021 and 2022.

# Post-pandemic: the UCAS data

Percentage from underrepresented group by provider quintile (wgt avg = 100)  
POLAR4



dataHE - Confidential for the Sutton Trust (02/02/23) q1m2

Controlling out the changing overall share on this (broader and younger) measure by indexing.

Shows some modest relative improvements in relative representation in 2021, held in 2022.



# Fair access: The “missing 3,000”

---

---

## **The Missing 3000**

**State school students  
under-represented  
at leading universities**

August 2004

---

---

## **Preface**

The analysis for this paper has been prepared by the Analytical Services Group at the Higher Education Funding Council for England (HEFCE). The Sutton Trust is extremely grateful for their co-operation and excellent work. The views expressed in this paper are those of the Sutton Trust.

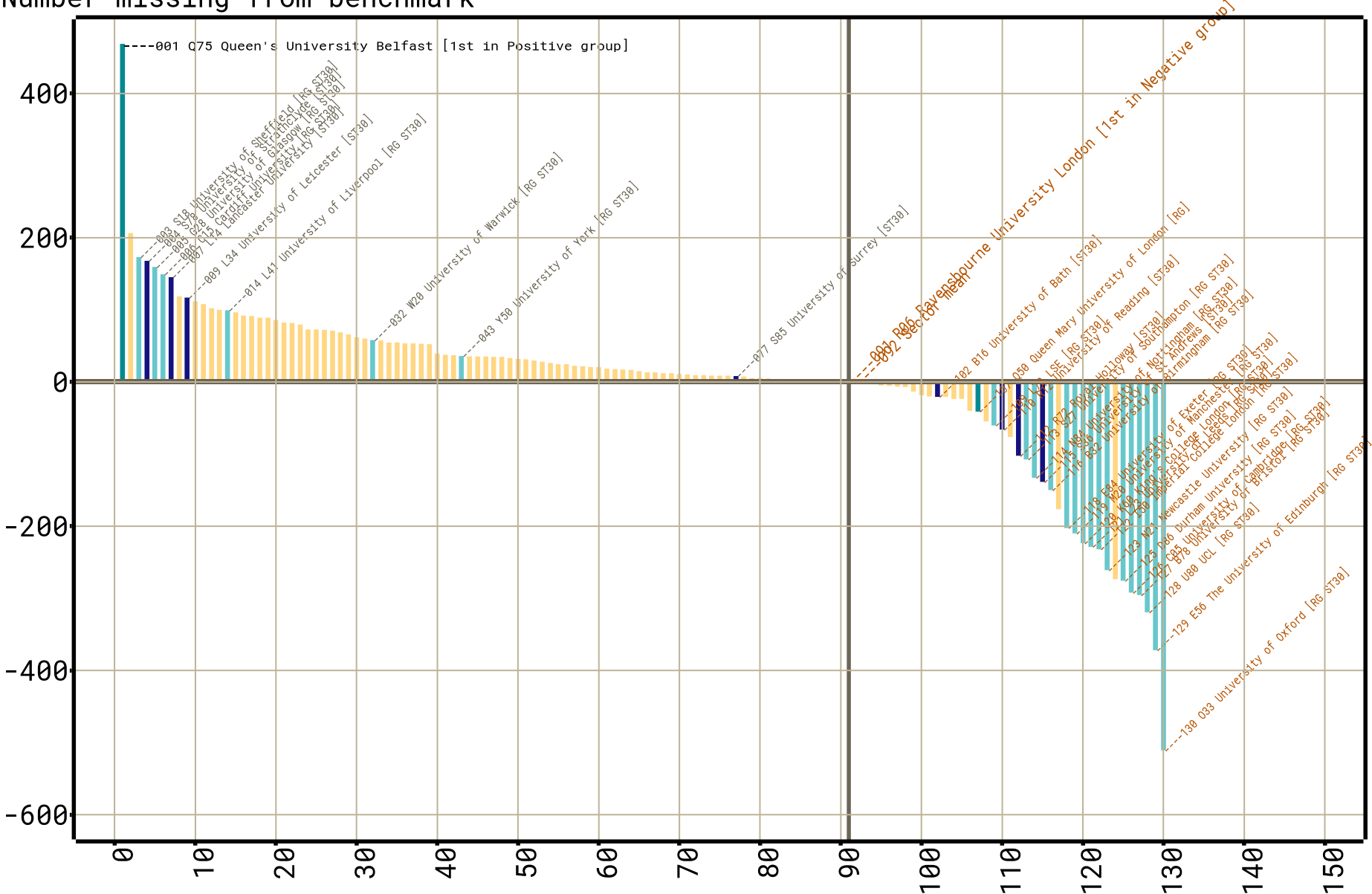
The missing 3,000 might well be the most influential statistic in fair access.

The Trust used it skilfully to galvanise policy and action.

Was based on ST13, but we use the ST30 (gives very similar value)

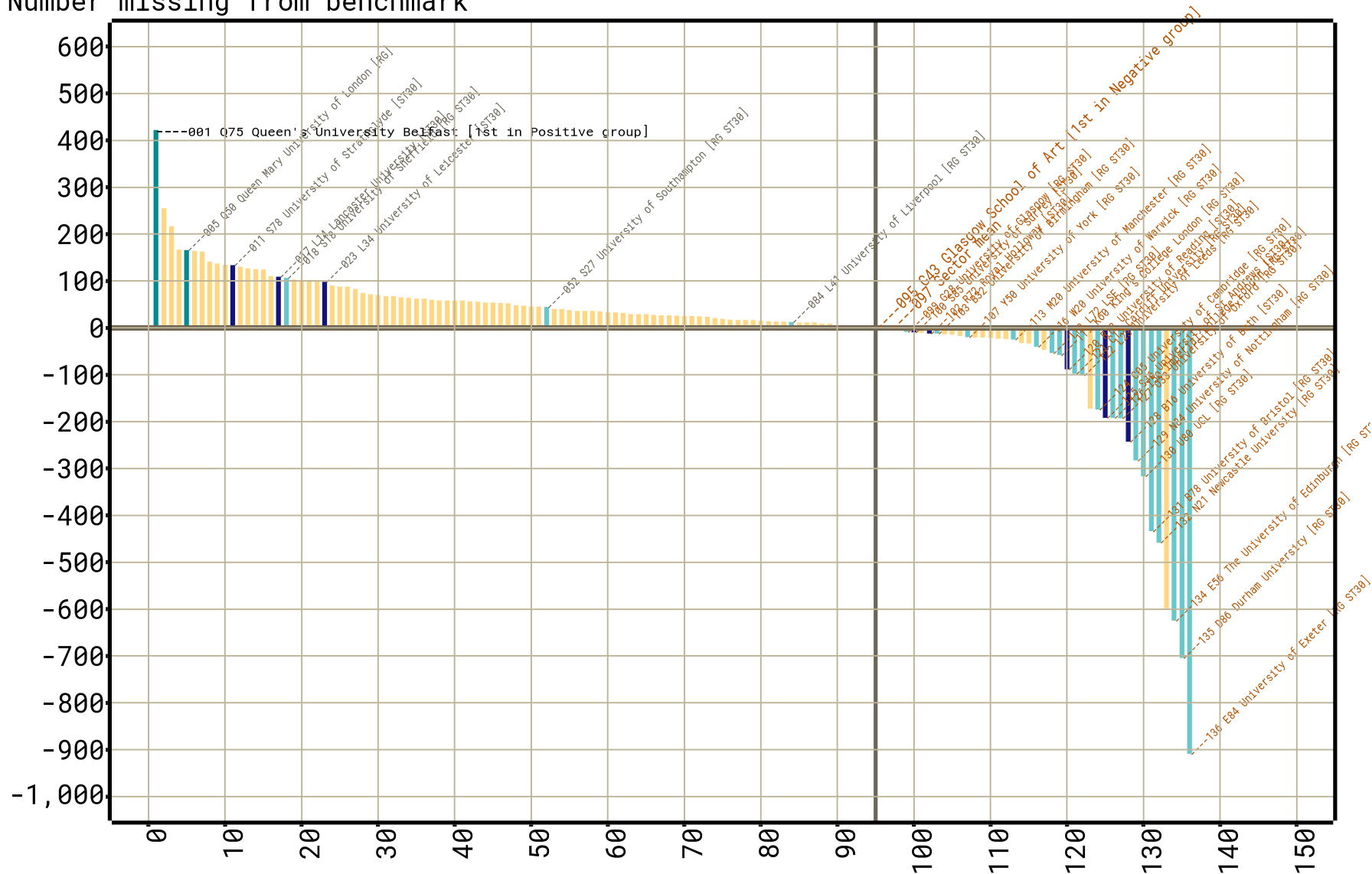
# Conditional fair access: 'benchmarks'

FD/T1a State schools or colleges: 1997  
Number missing from benchmark



# 'Missing' state school 2020

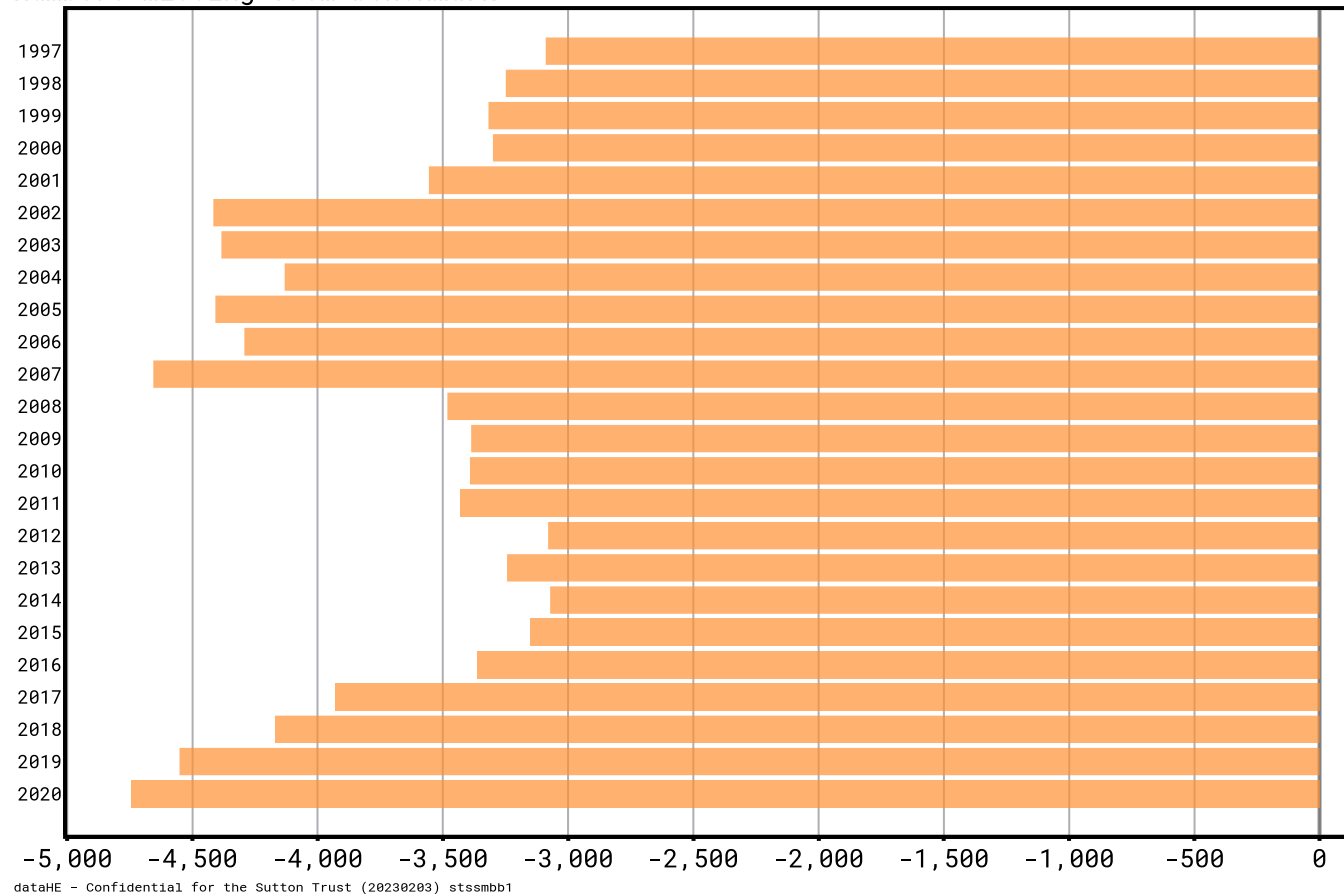
FD/T1a State schools or colleges: 2020  
 Number missing from benchmark



dataHE - Confidential for the Sutton Trust (20230203) ndssmb02

# The “missing 3,000” through time

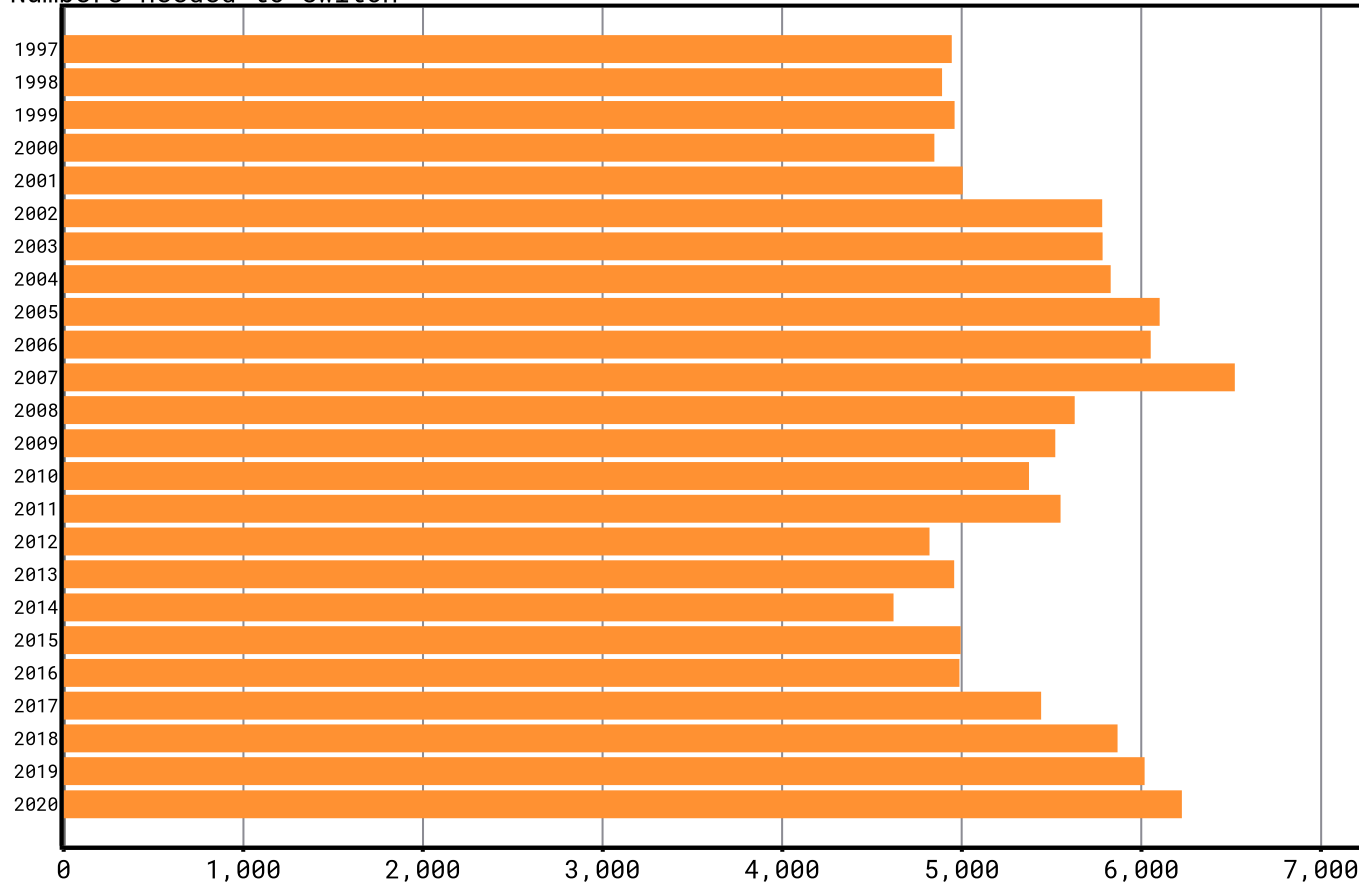
FD/T1a State schools or colleges: Sutton Trust 30  
Numbers missing from benchmark



Around 3,000 state school entrants a year were ‘missing’ (relative to the benchmark) from ST30 universities in the late 1990s. This rose to over 4,000 in the early 2000s before falling back to around 3,000 2012-2015. Now over 4,500.

# 'Switching' number for 'fair access'

FD/T1a State schools or colleges  
Numbers needed to switch



dataHE analysis for the Sutton Trust // 20221116\_ssmbb2

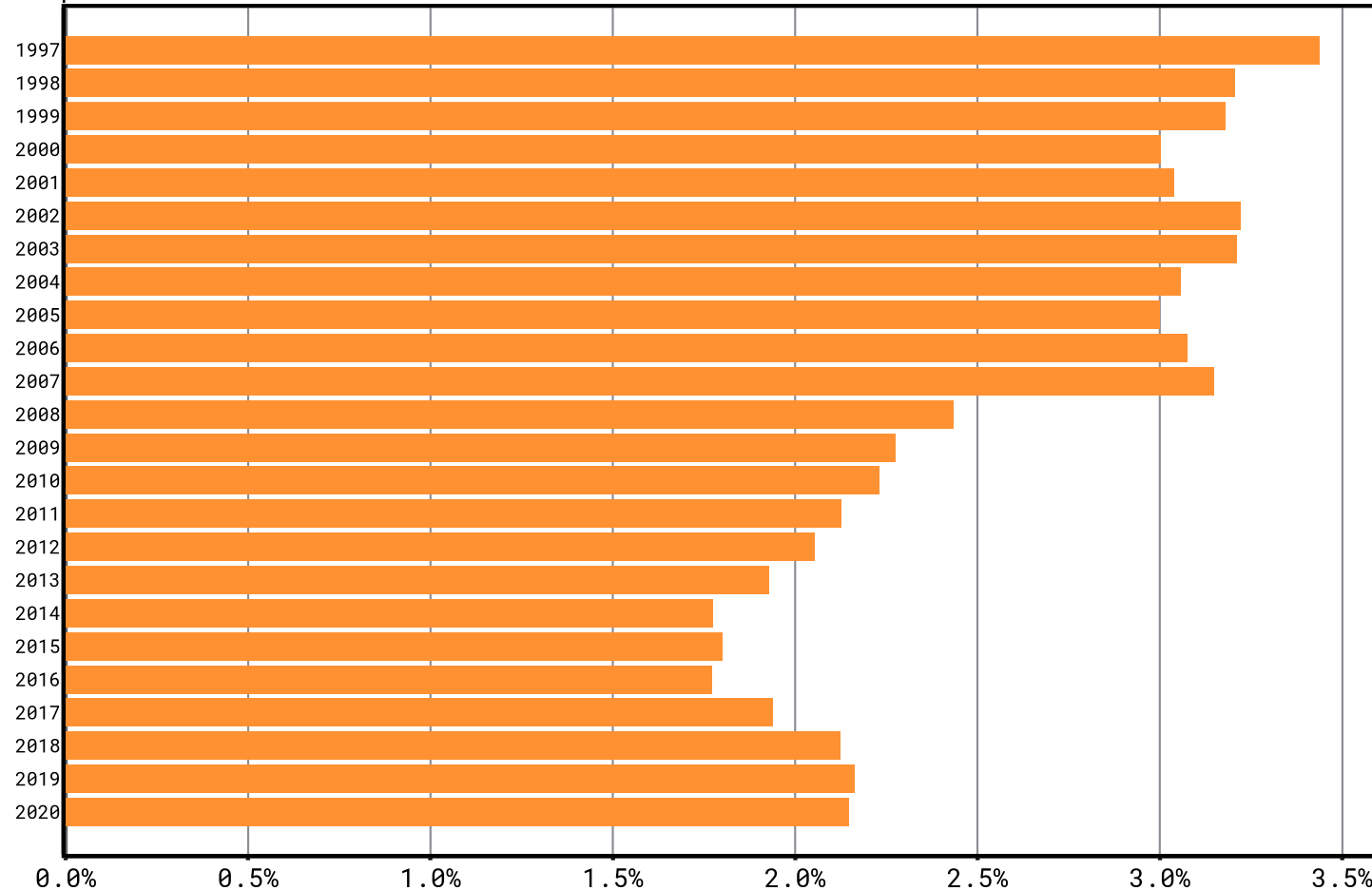
A more generalised statistic is the 'switching' number: How many need to move for everyone to be on their benchmark?

Was around 5,000 in 1997.

Over 6,000 in 2020.

# 'Switching' as a proportion of entrants

FD/T1a State schools or colleges  
Proportion needed to switch



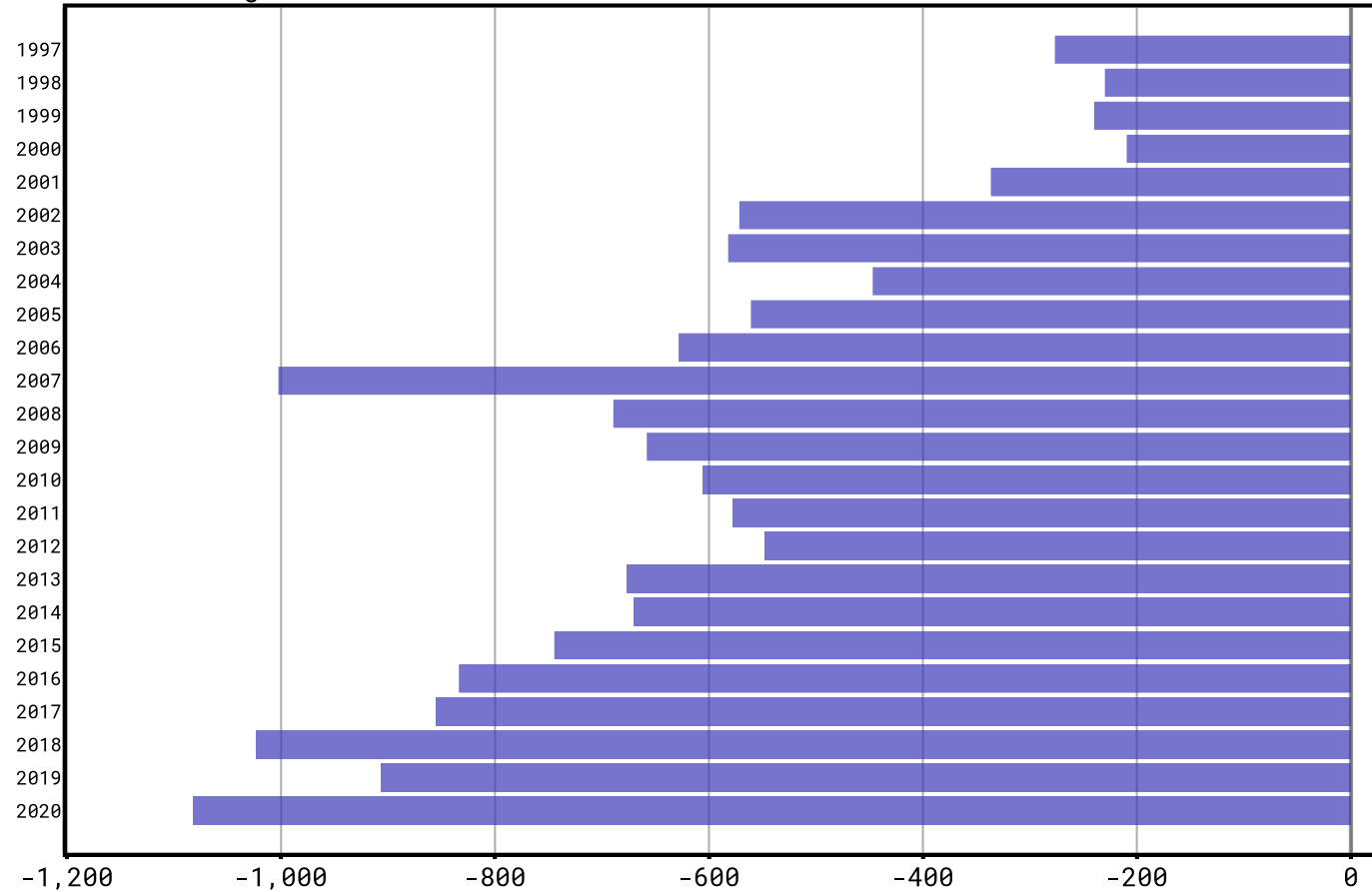
dataHE analysis for the Sutton Trust // 20221116\_ssmbb4

But intakes have got larger. And state proportions have grown.

As a % of state school entrants, the number of switchers needed has fallen by a third.

# Missing low participation at ST30

FD/T1a Low participation: Sutton Trust 30  
Numbers missing from benchmark

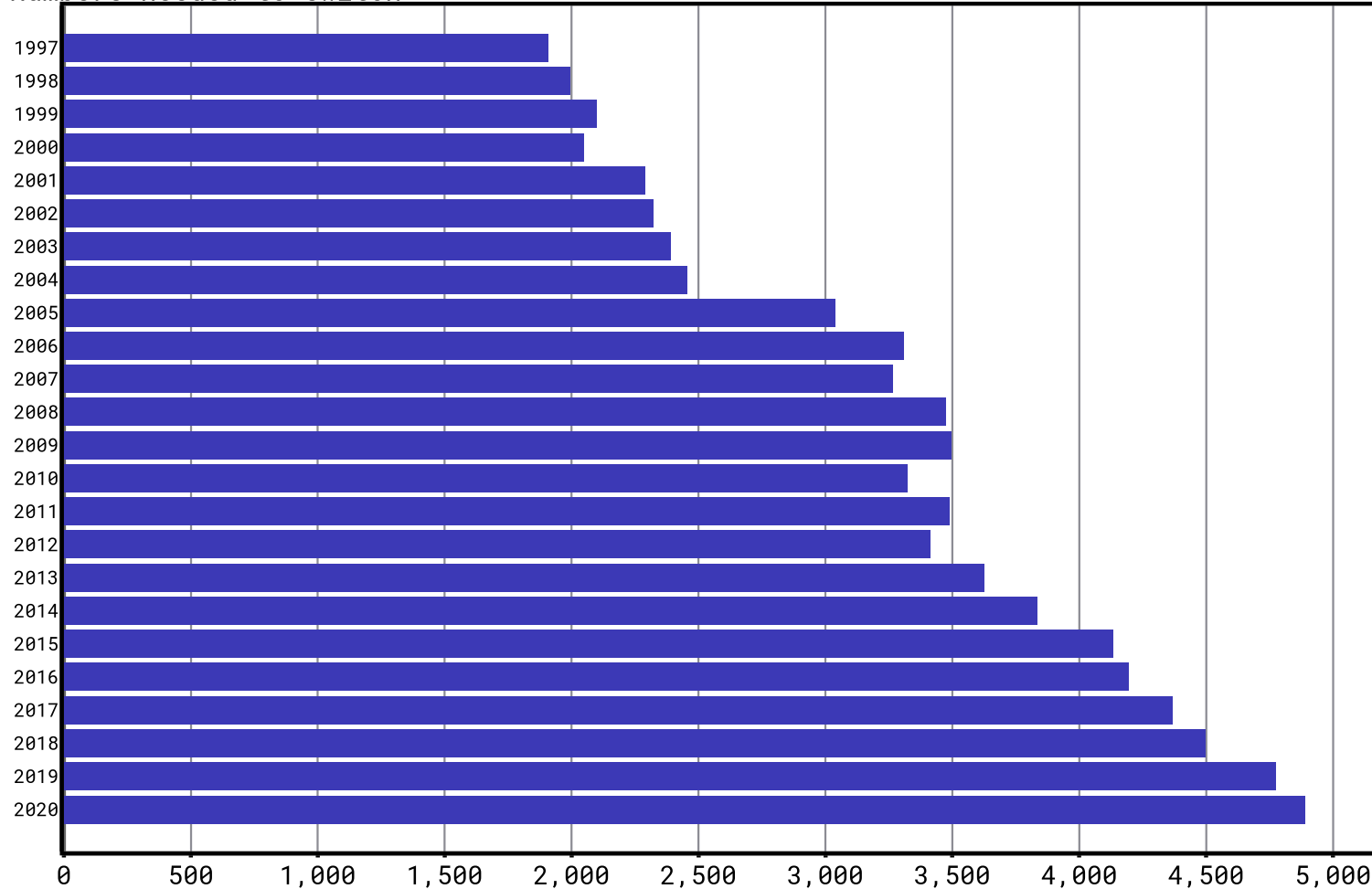


dataHE - Confidential for the Sutton Trust (20230203) st1pmbb1

Around 1,000 POLAR4 Q1 'missing' from ST30 in 2020.  
Equivalent figure for late 1990s was around 200.

# PI low participation 'switchers' increase

FD/T1a Low participation  
Numbers needed to switch



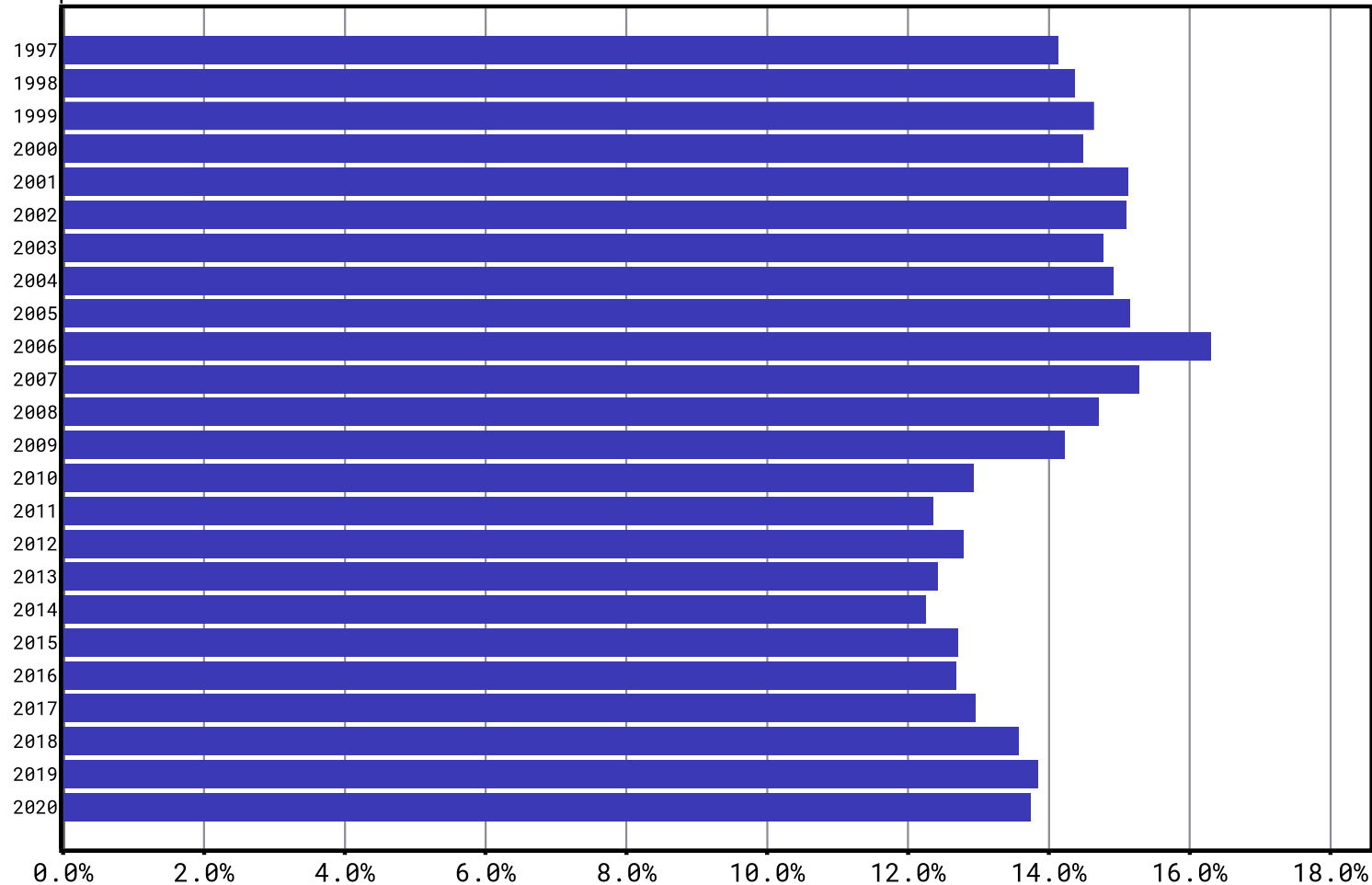
dataHE analysis for the Sutton Trust // 20221116\_1pmbb2

Across the sector the number of switchers needed for 'fair access' has increased from around 2,000 to almost 5,000 in 2020.



# PI low participation 'switchers'

FD/T1a Low participation  
Proportion needed to switch



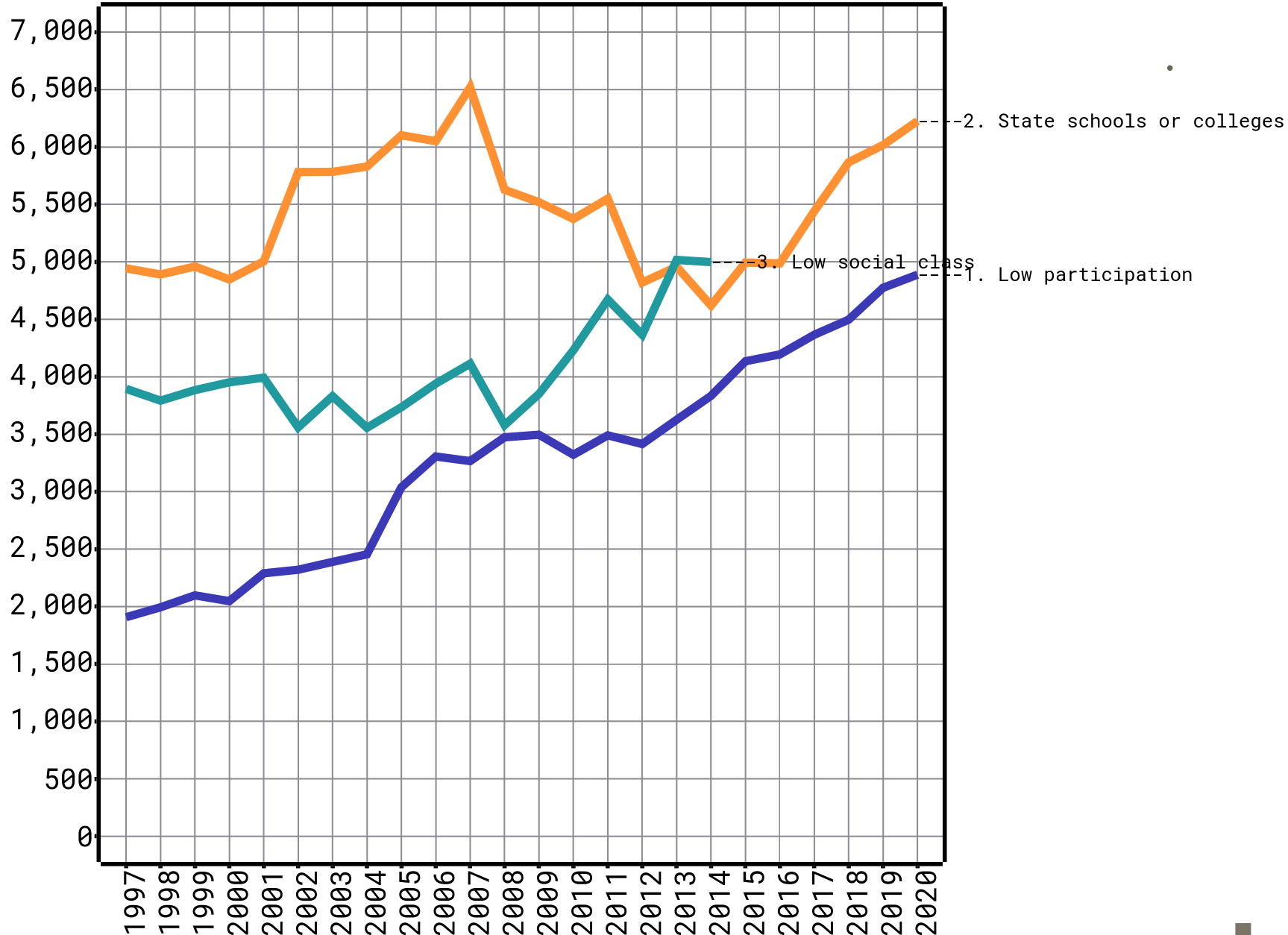
dataHE analysis for the Sutton Trust // 20221116\_1pmbb4

A lot of the growth in switchers comes from growth in entry.

As a proportion of Q1 entrants has neither got much worse or much better.

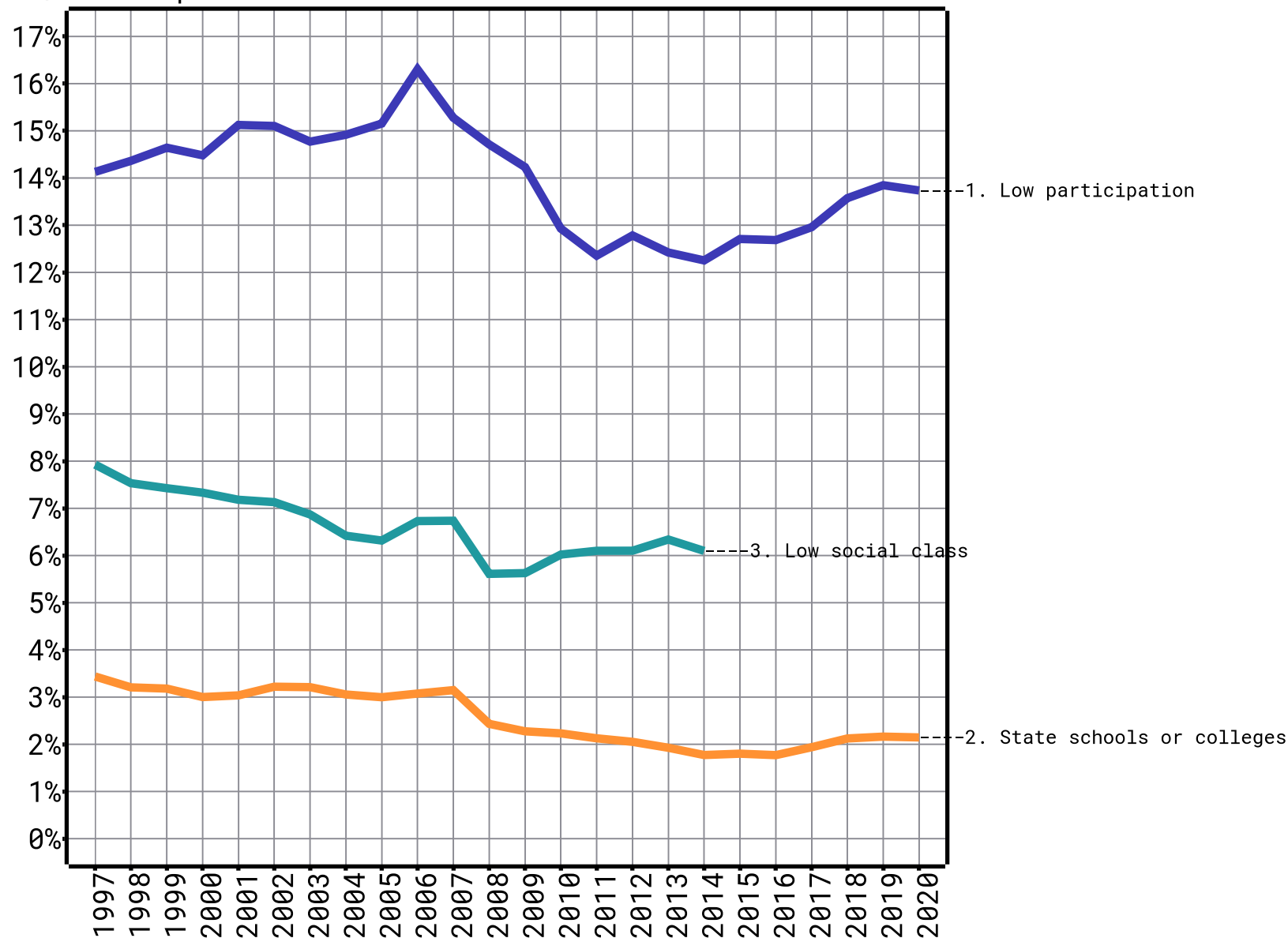
# Switchers needed by dimension

FD/T1a Numbers needed to switch



# Switchers needed by dimension (%)

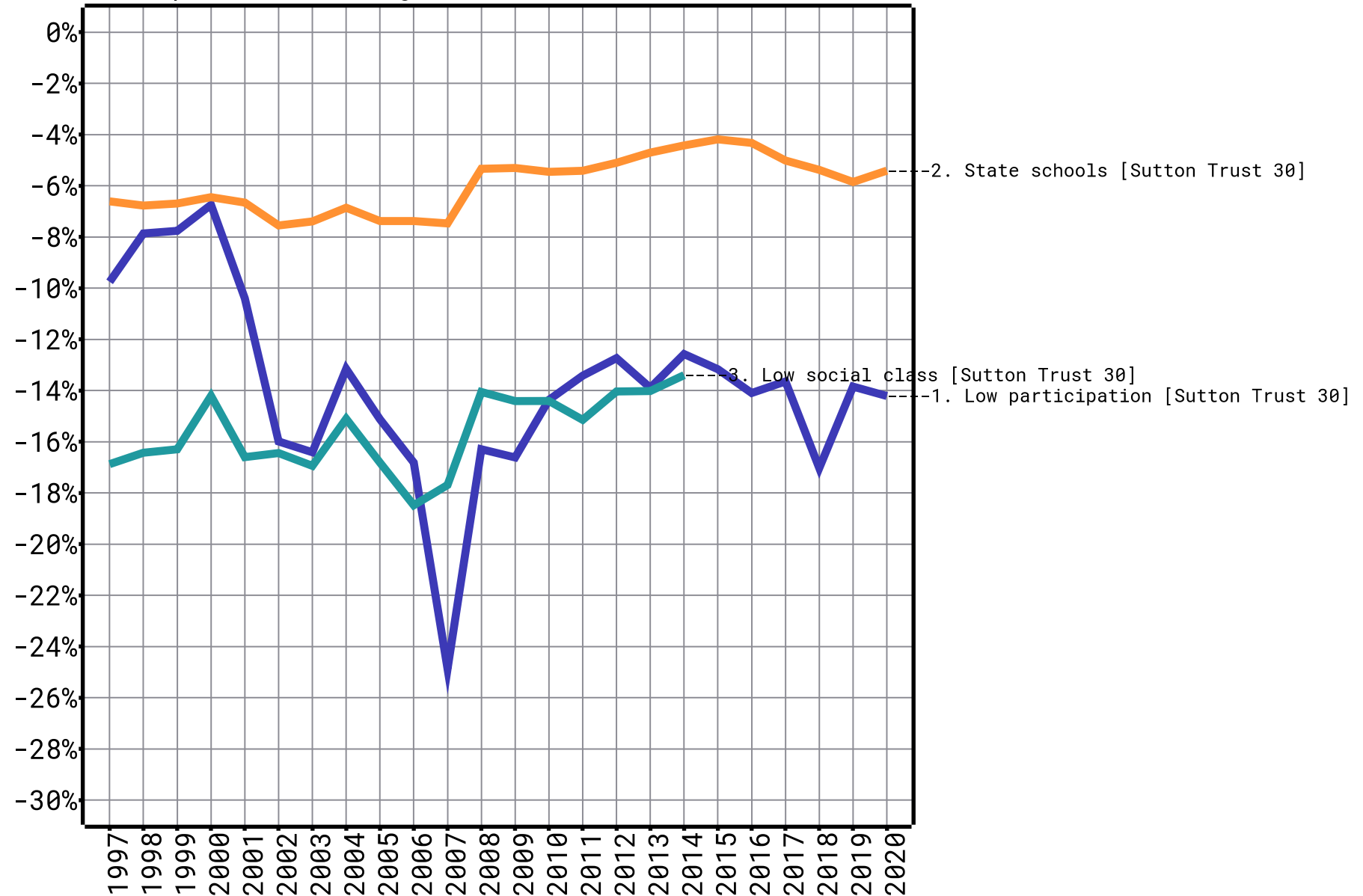
FD/T1a Proportion needed to switch



dataHE - Confidential for the Sutton Trust (20230203) 1pmb10p

# ST30: % of group missing little changed

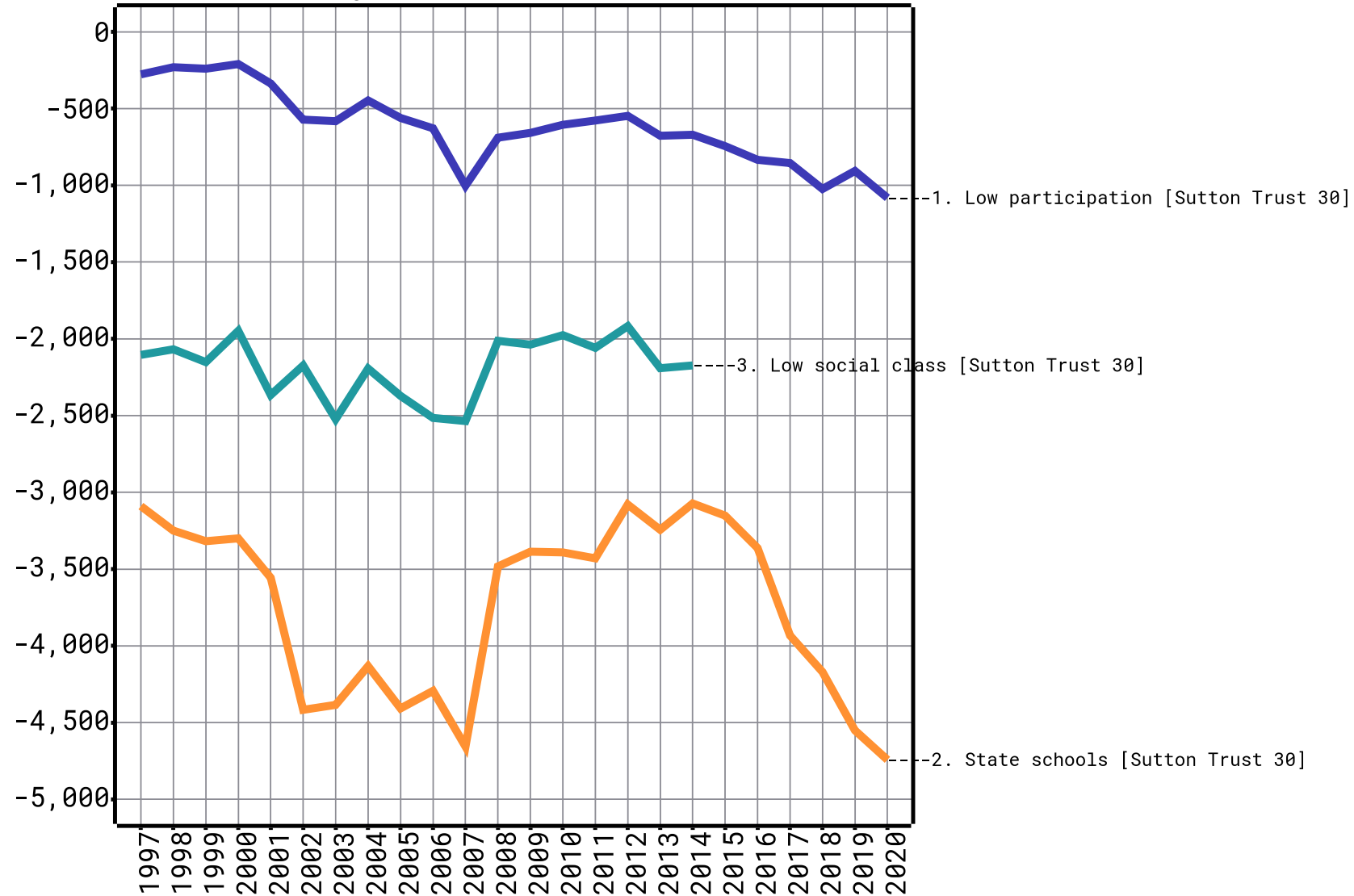
FD/T1a Proportion missing from benchmark



dataHE - Confidential for the Sutton Trust (20230203) 1pmb12p

# ST30: but numbers affected growing

FD/T1a Numbers missing from benchmark



dataHE - Confidential for the Sutton Trust (20230203) 1pmb12

---

## (3) Equality

---

# Sector-level equality

---

Sector level equality focus on the big life-outcome divide of going to university or not. It tracks the overall equality of the UK sector by different dimensions. We define “under-represented” categories by entry rates in the mid-2000s.

## (1) **Application, success and entry rates:**

Different ways of measuring equality

## (2) **How entry rates have changed**

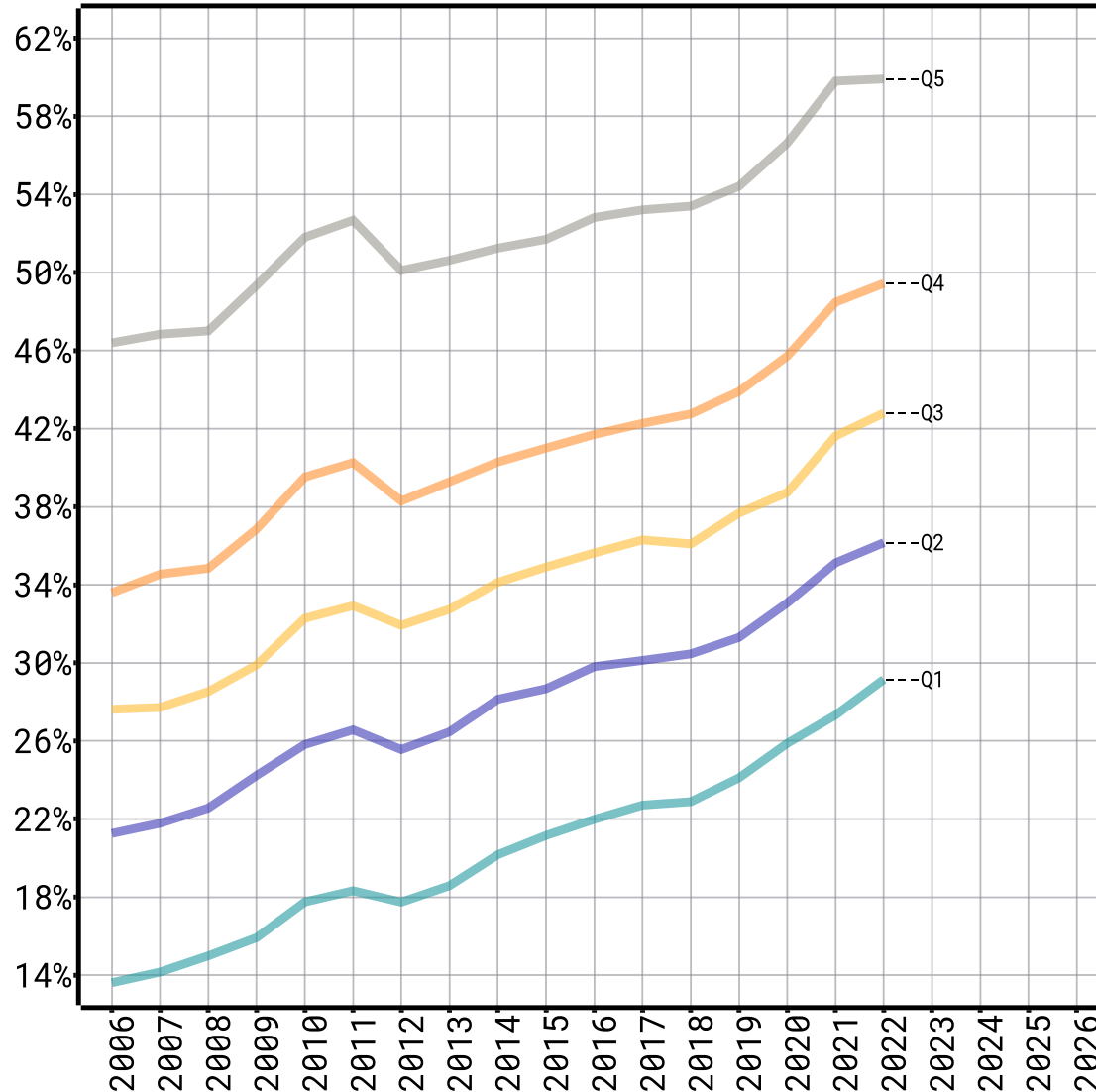
Overall, and relative to others – isolated indices

## (3) **How to scale the equality differences:**

What it says about future priorities

# POLAR4 application rates

2006-2022 AR by POLAR4



dataHE - Confidential for the Sutton Trust (20221118) arPOL

Application rate is the start of the process.

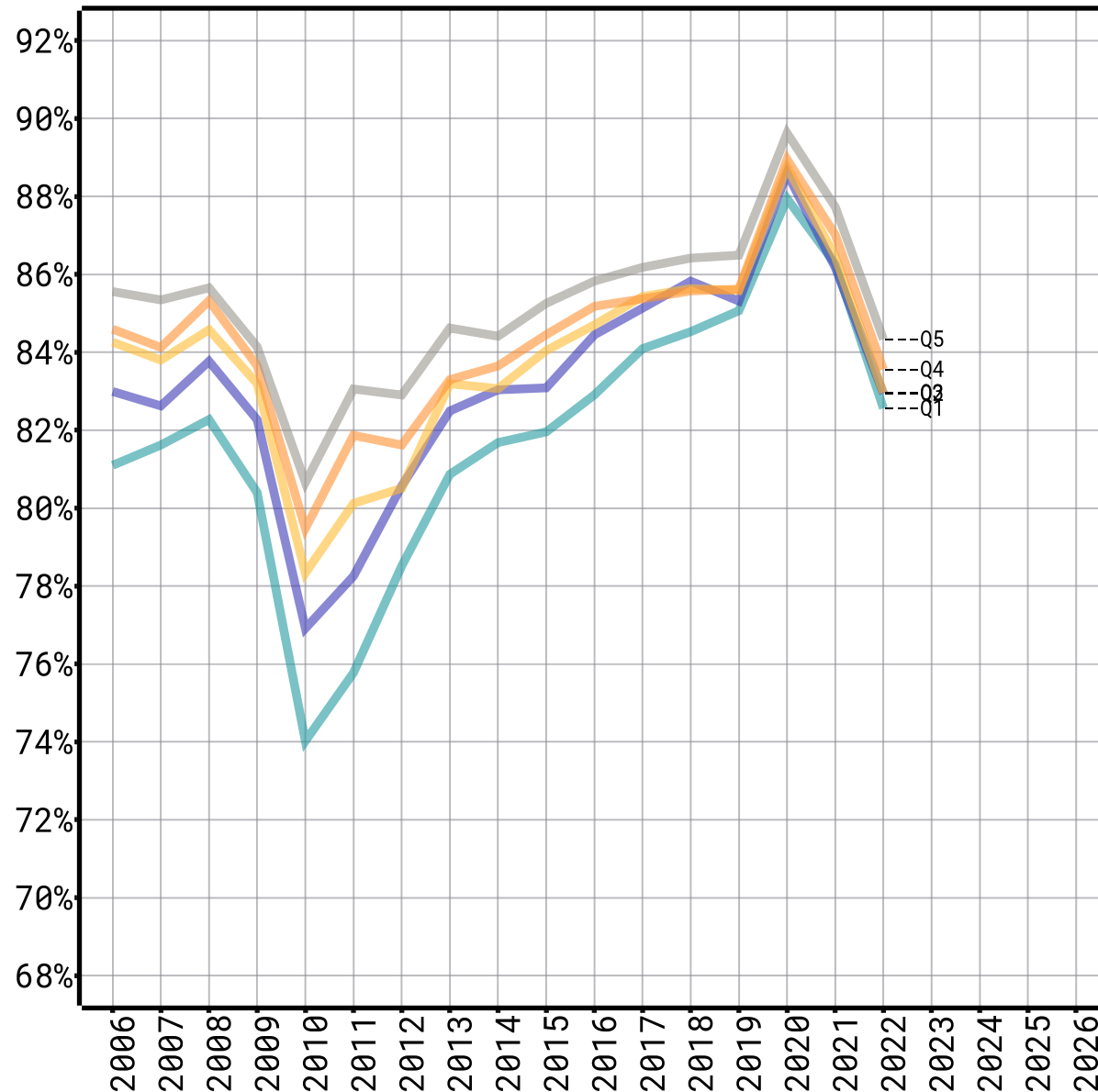
Records equality of demand/aspiration.

These show large differences



# POLAR4 success rates

2006-2022 SR by POLAR4



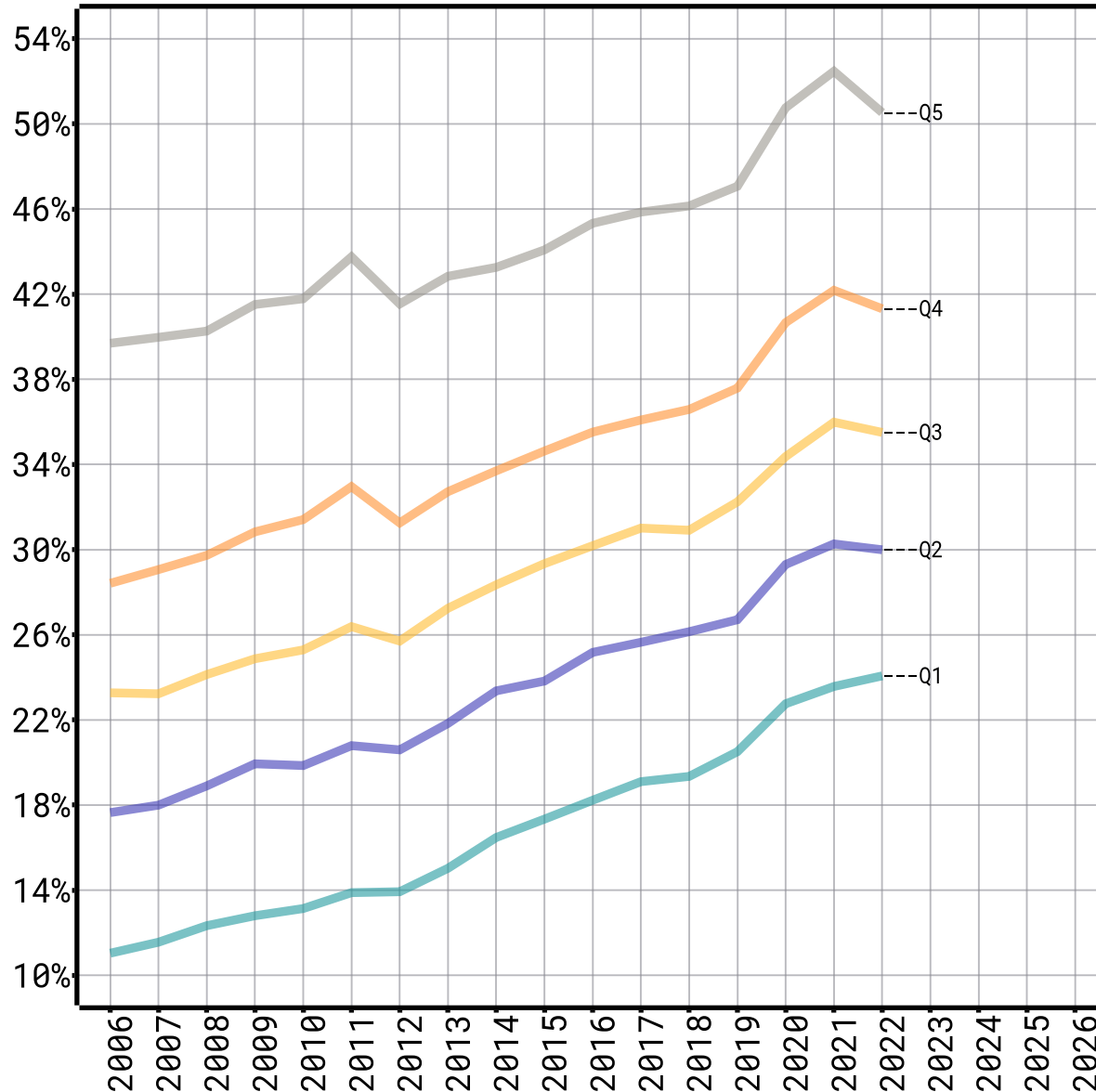
dataHE - Confidential for the Sutton Trust (20221118) srPOL

The proportion of applicants who get placed is recorded through the success rate.

This will reflect lots of things (choices, grades, intention, etc.) – but if the admission process was very unfair you would expect to see a signal here.

# POLAR4 entry rates

2006-2022 ER by POLAR4



dataHE - Confidential for the Sutton Trust (20221118) erPOL

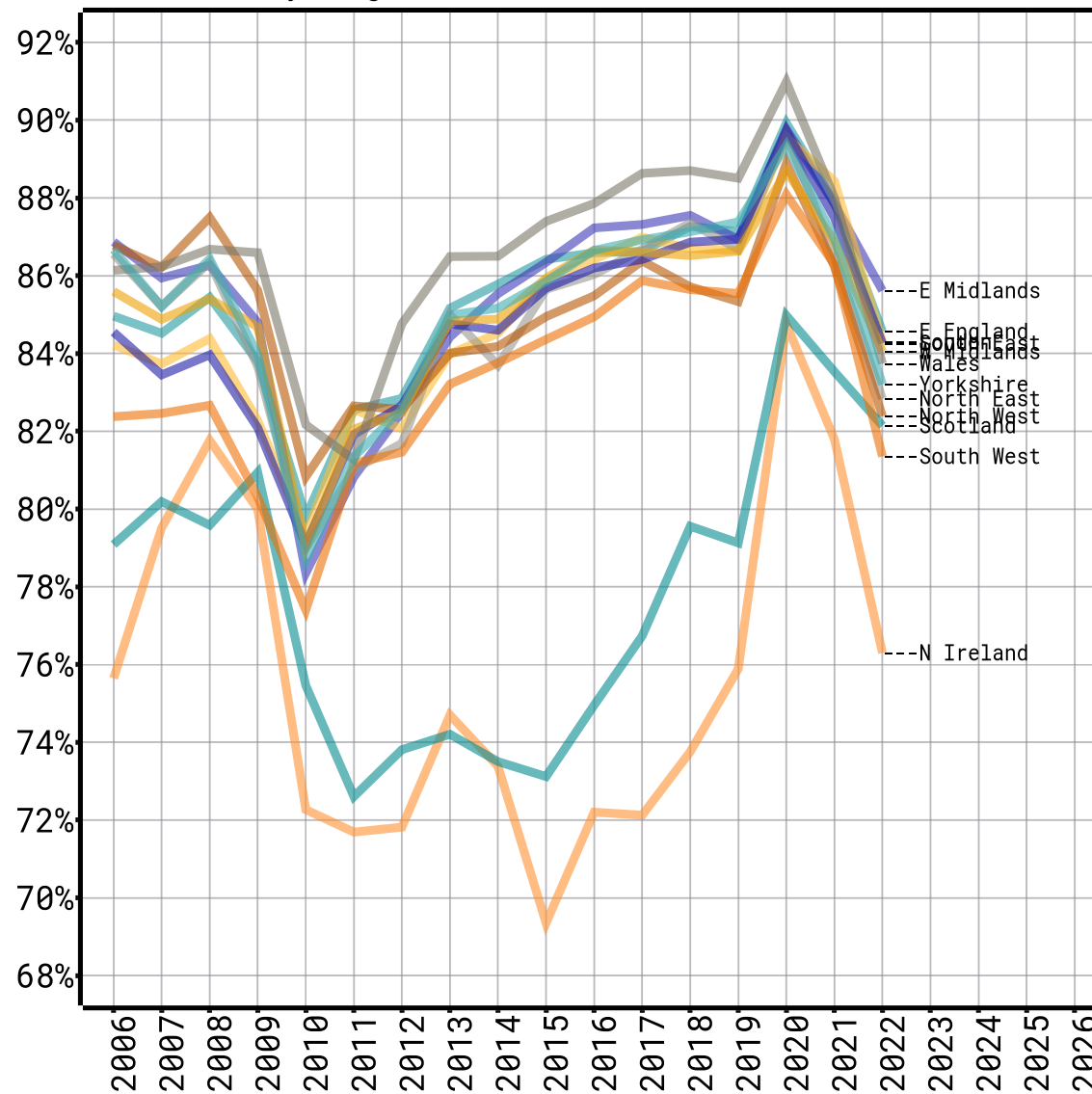
The entry rate is the proportion of the population who get a place.

As the outcome it is the key measure and we define under-represented groups as those below average in the early part of the period.

Q1-Q3 for POLAR.

# Success rates generally unimportant

2006-2022 SR by Region



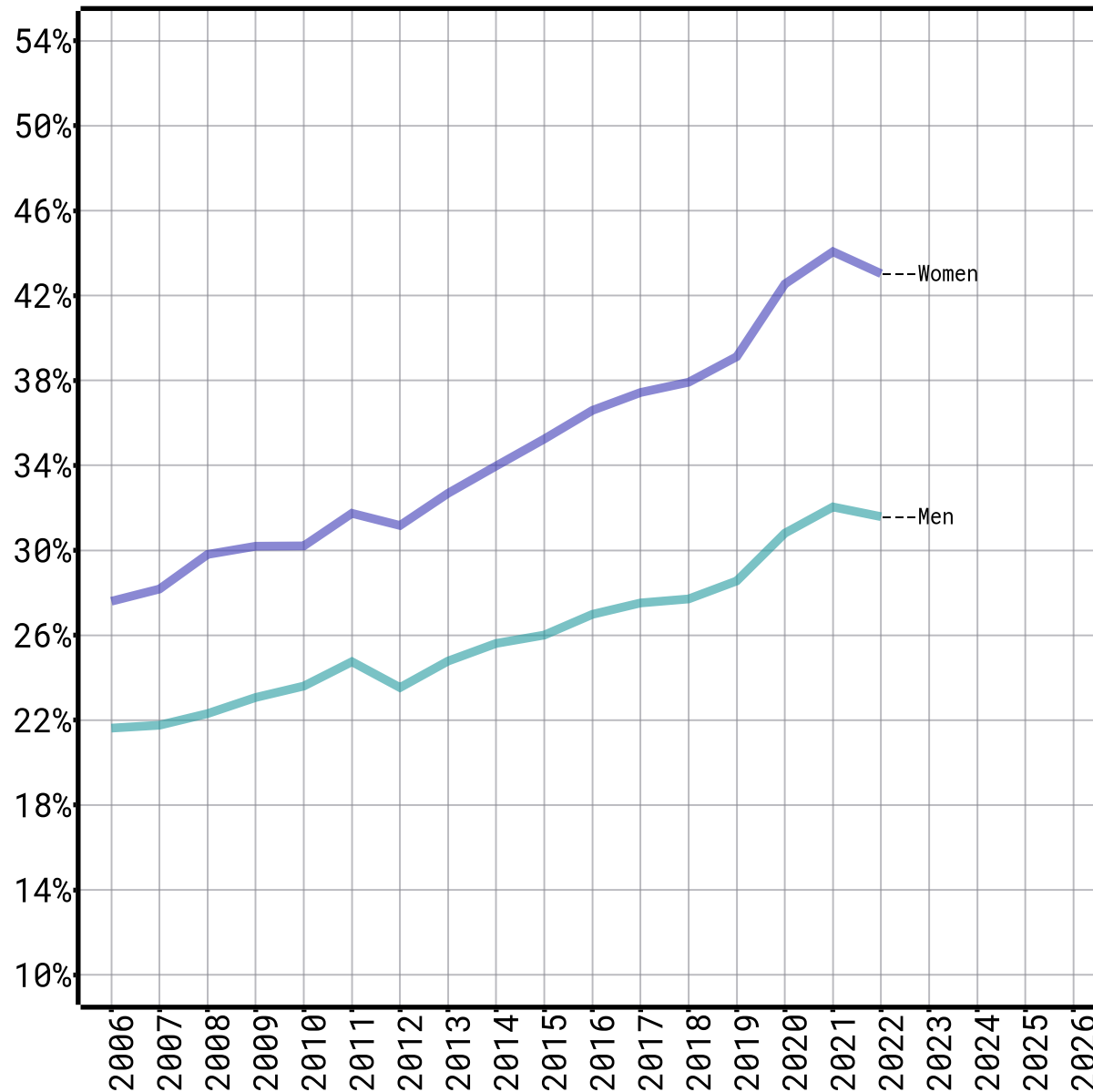
With one or two exceptions (notably region/country) the “success rates” do not vary much and therefore do not drive equality outcomes.

So we concentrate on entry rates.

dataHE - Confidential for the Sutton Trust (20221118) srReg

# Entry rates by sex

2006-2022 ER by Sex

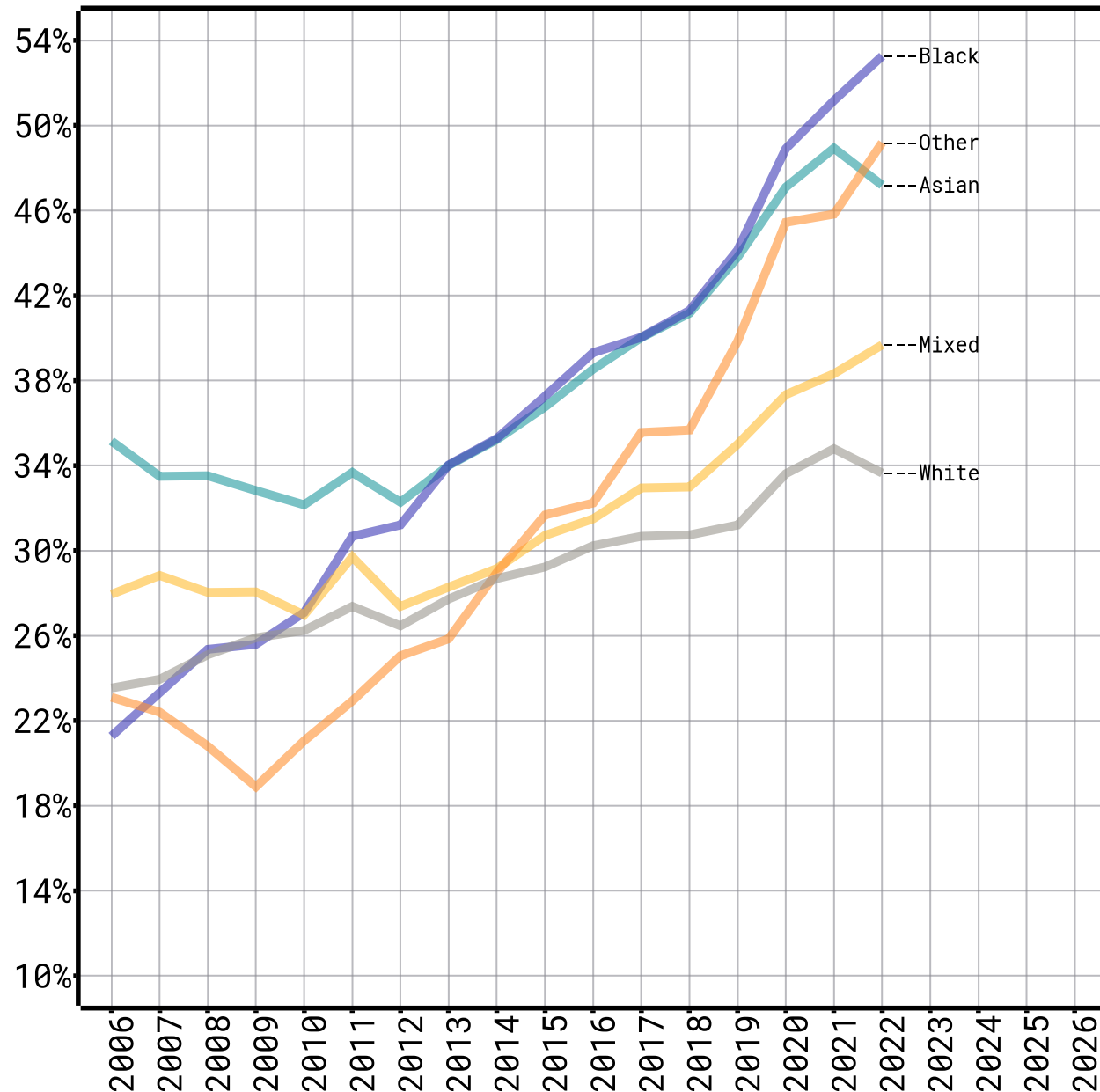


dataHE - Confidential for the Sutton Trust (20221118) erSex

Men are set as the under-represented group on the sex dimension.

# Entry rates by ethnic group

2006-2022 ER by Ethnicity

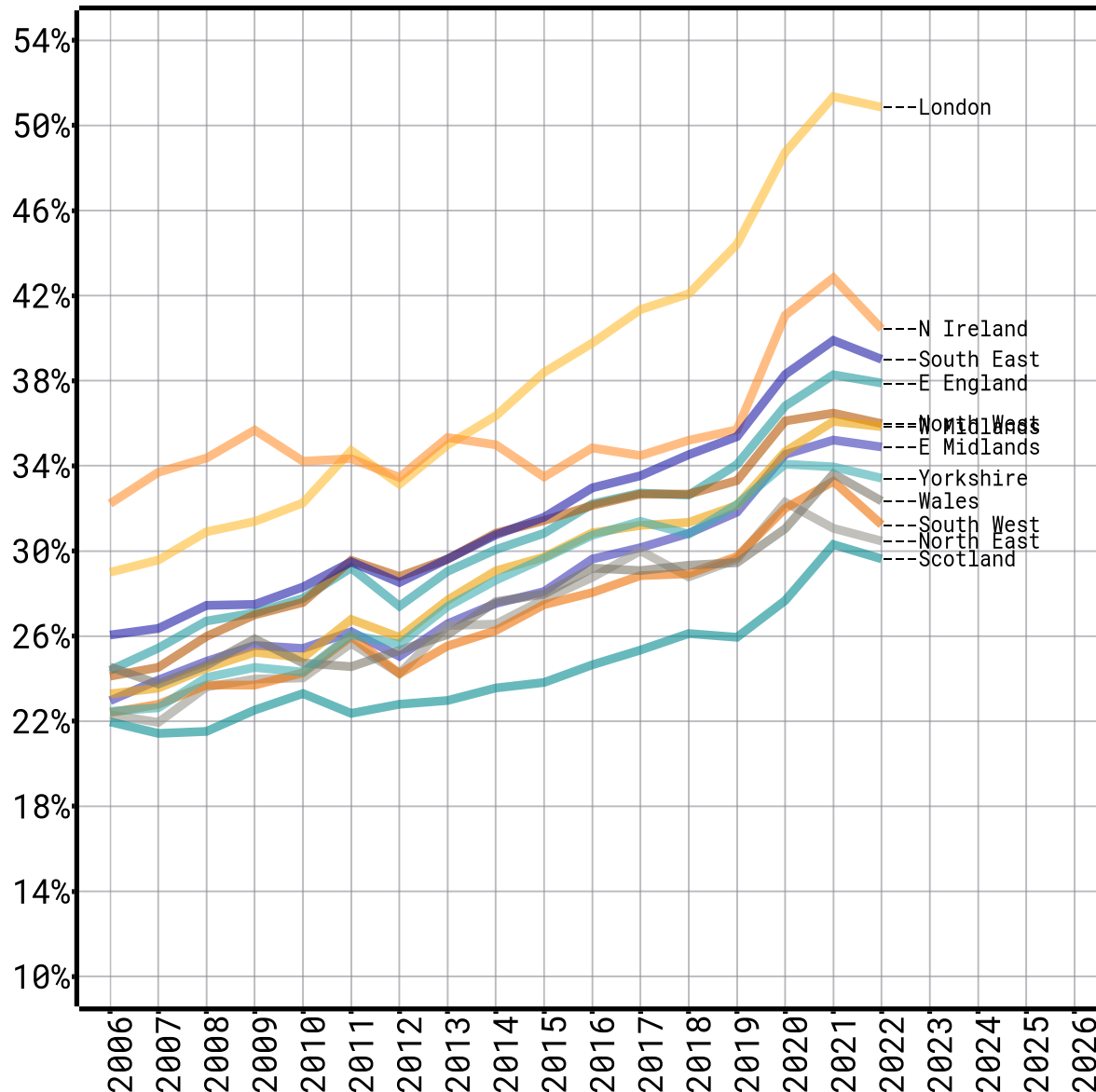


UCAS self-declared summary groups combined with (ultimately) census and APS estimates.

For ethnic group White and Other are taken as under-represented.

# Entry rates by region

2006-2022 ER by Region



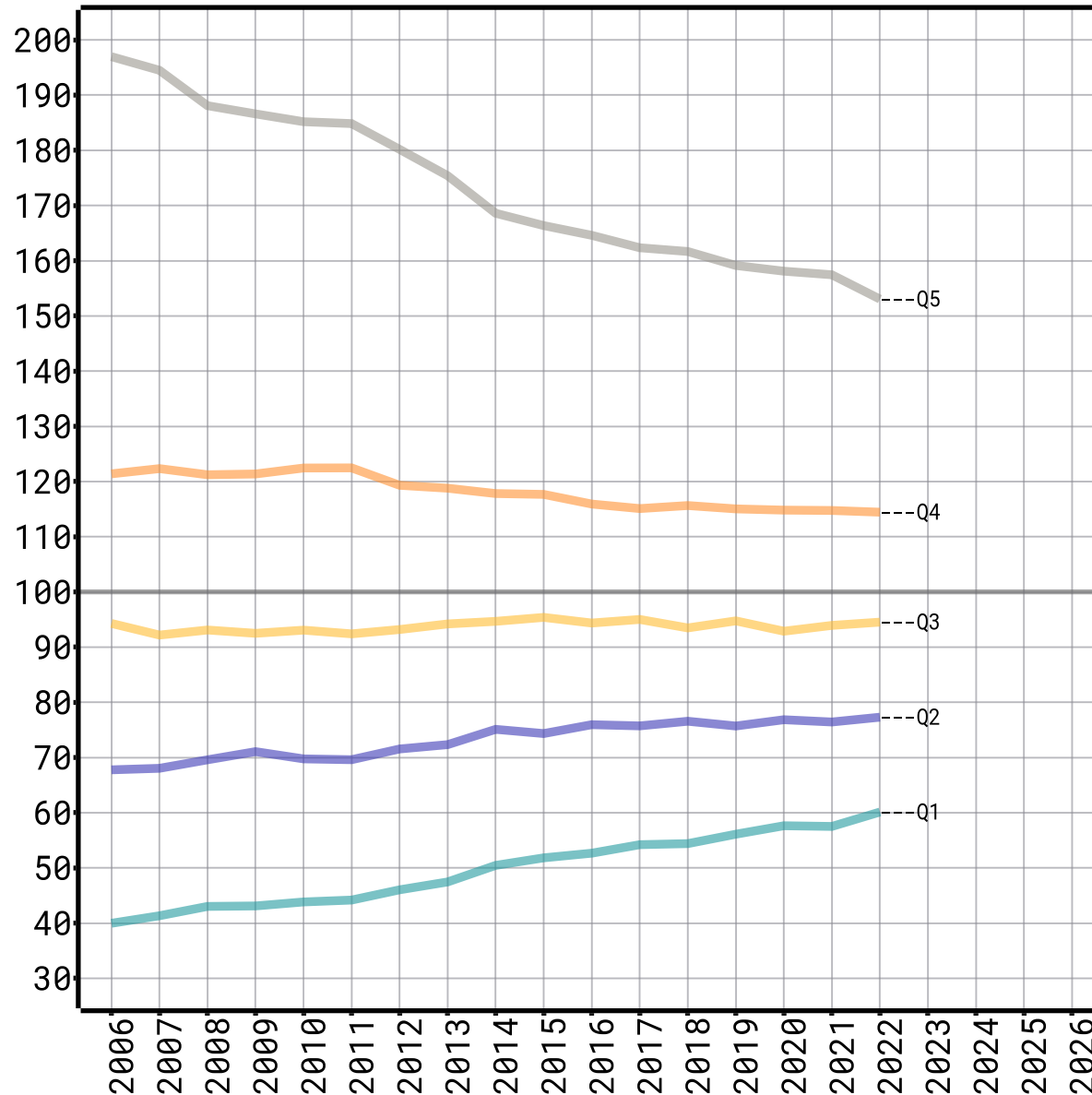
dataHE - Confidential for the Sutton Trust (20221118) erReg

We use the declared region of domicile (not postcode based).

Scotland has large below degree sector in HE colleges not in UCAS data. But a good account of chances of entering university-provided higher education.

# Entry rates indexed to others: POLAR

2006-2022 ER isolated dimension indexed: POLAR4



dataHE - Confidential for the Sutton Trust (20221118) ert2

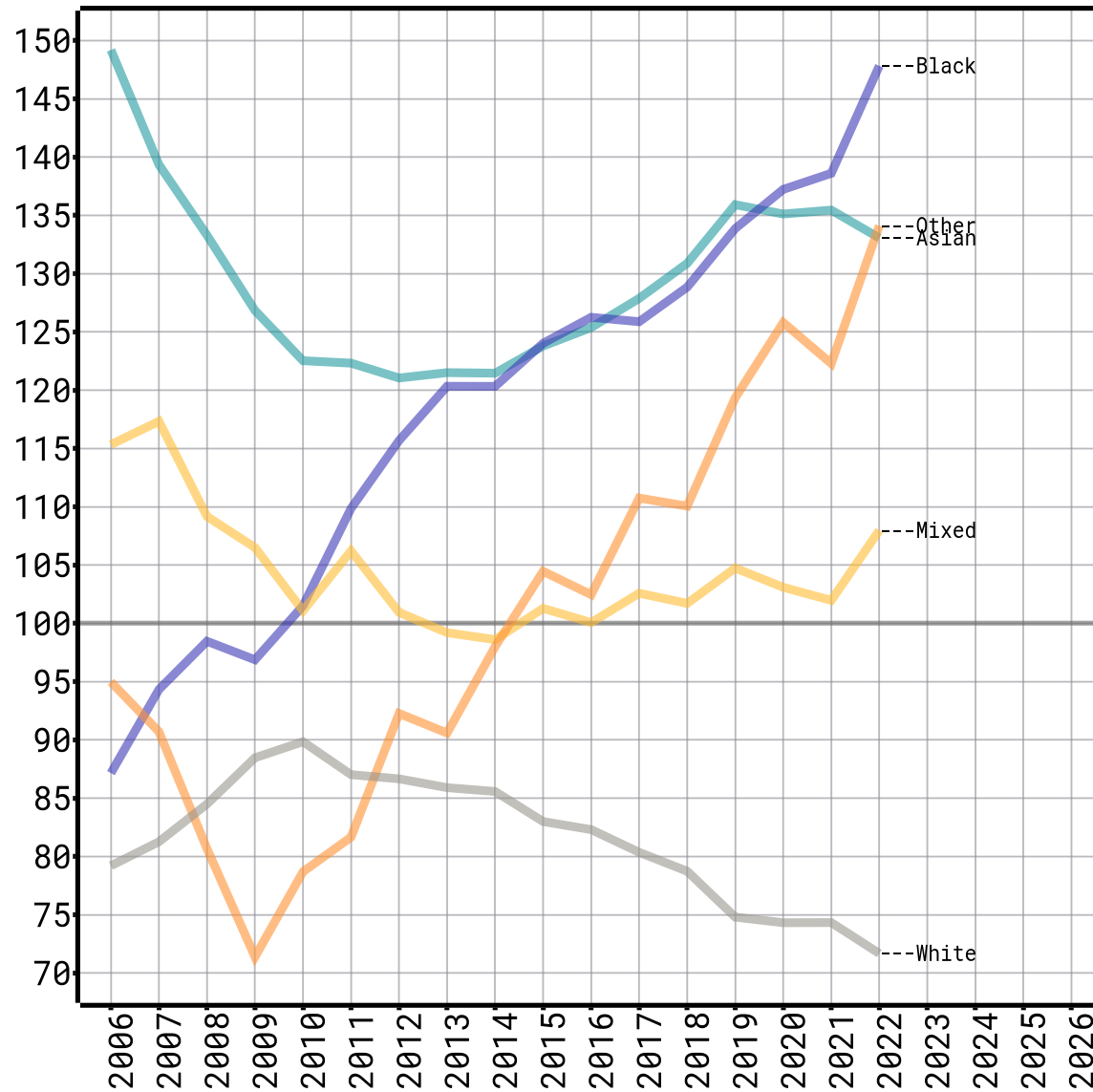
We prefer isolated indexed rates. These are the entry rate of the group relative to the entry rate of all *other* groups.

This reveals true trends in equality, removing effects of changes in overall entry rates *and* the effect of group on average.

POLAR clearly improving.

# Isolated index for ethnic groups

2006-2022 ER isolated dimension indexed: Ethnicity



dataHE - Confidential for the Sutton Trust (20221118) ert1

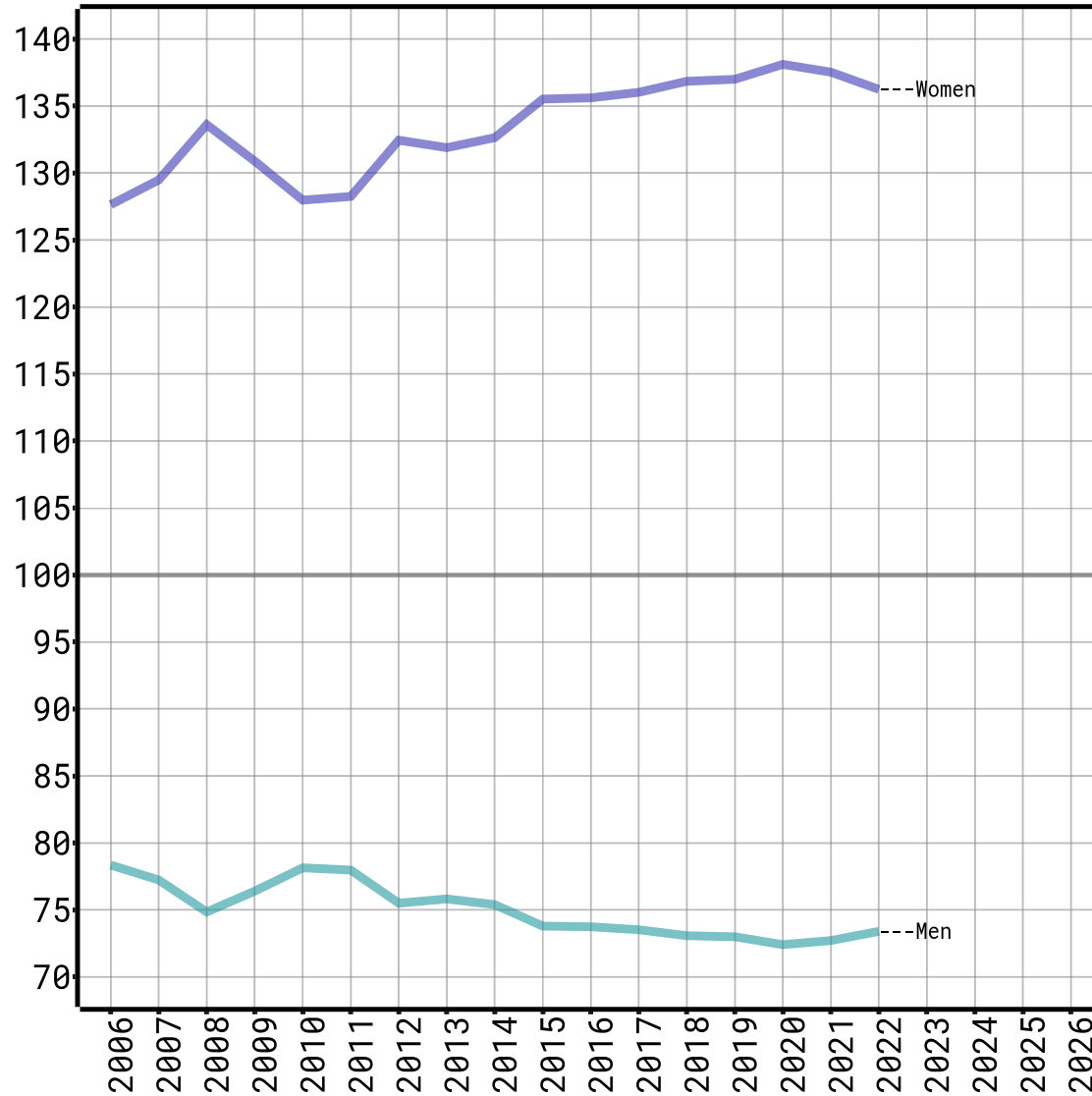
Ethnic groups are very uneven in population size so the isolated approach is particularly useful.

Young people in the White ethnic group now have less than 75% of the entry rate of other ethnic groups.



# Isolated index for groups by sex

2006-2022 ER isolated dimension indexed: Sex



dataHE - Confidential for the Sutton Trust (20221118) ert4

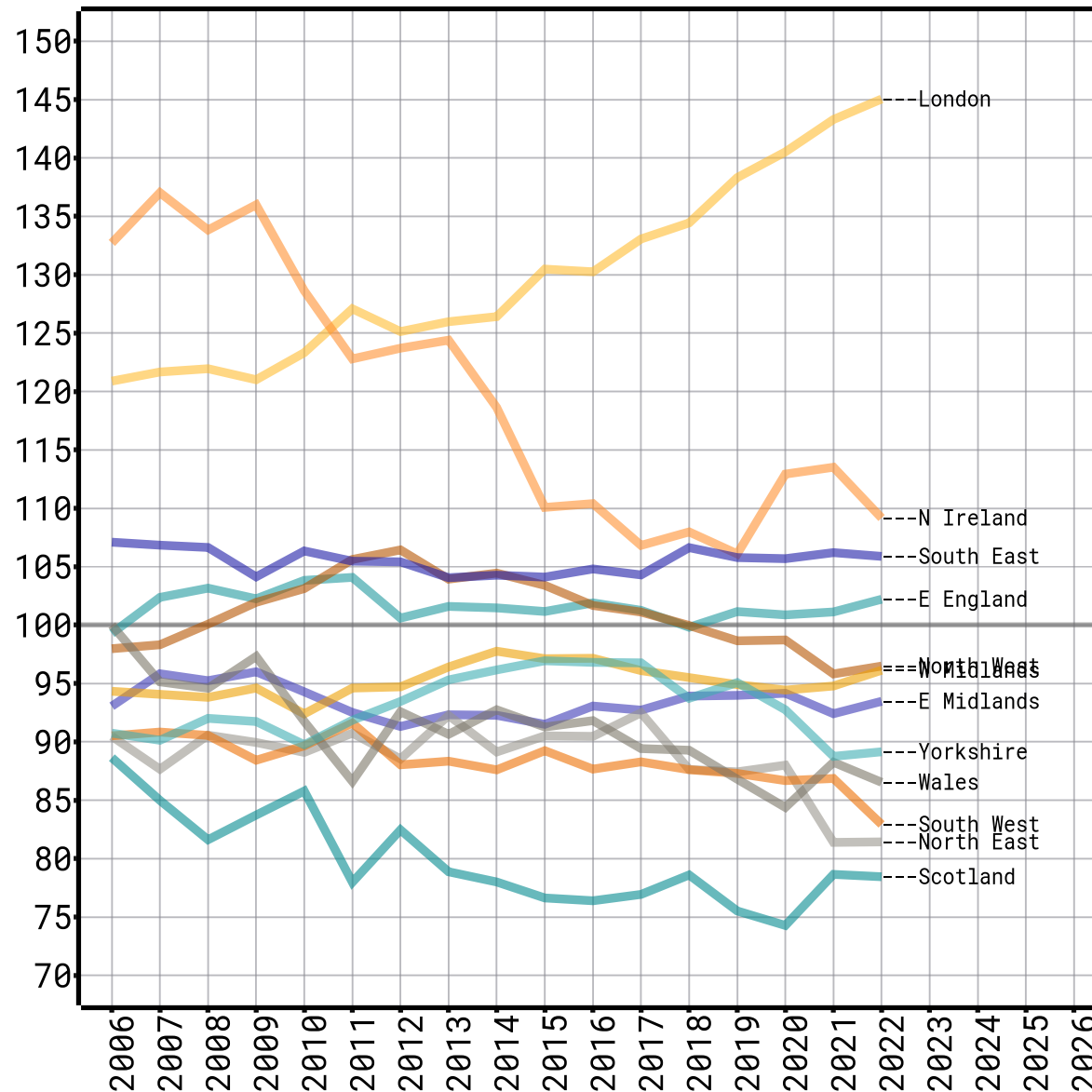
Men have entry rates less than 75% of that of women.

Women are over a third more likely to enter than men.

These differences are wider than in the 2000's but have stabilised in recent years.

# Isolated index for regions

2006-2022 ER isolated dimension indexed: Region



dataHE - Confidential for the Sutton Trust (20221118) ert3

Young people in London are now 45% more likely to enter higher education than peers in other regions, a gap that is growing steadily.

Young people in Scotland, the North East and the South West are around 20% less likely to enter (university) higher education than people in other regions.

# Scaling the gaps: numbers

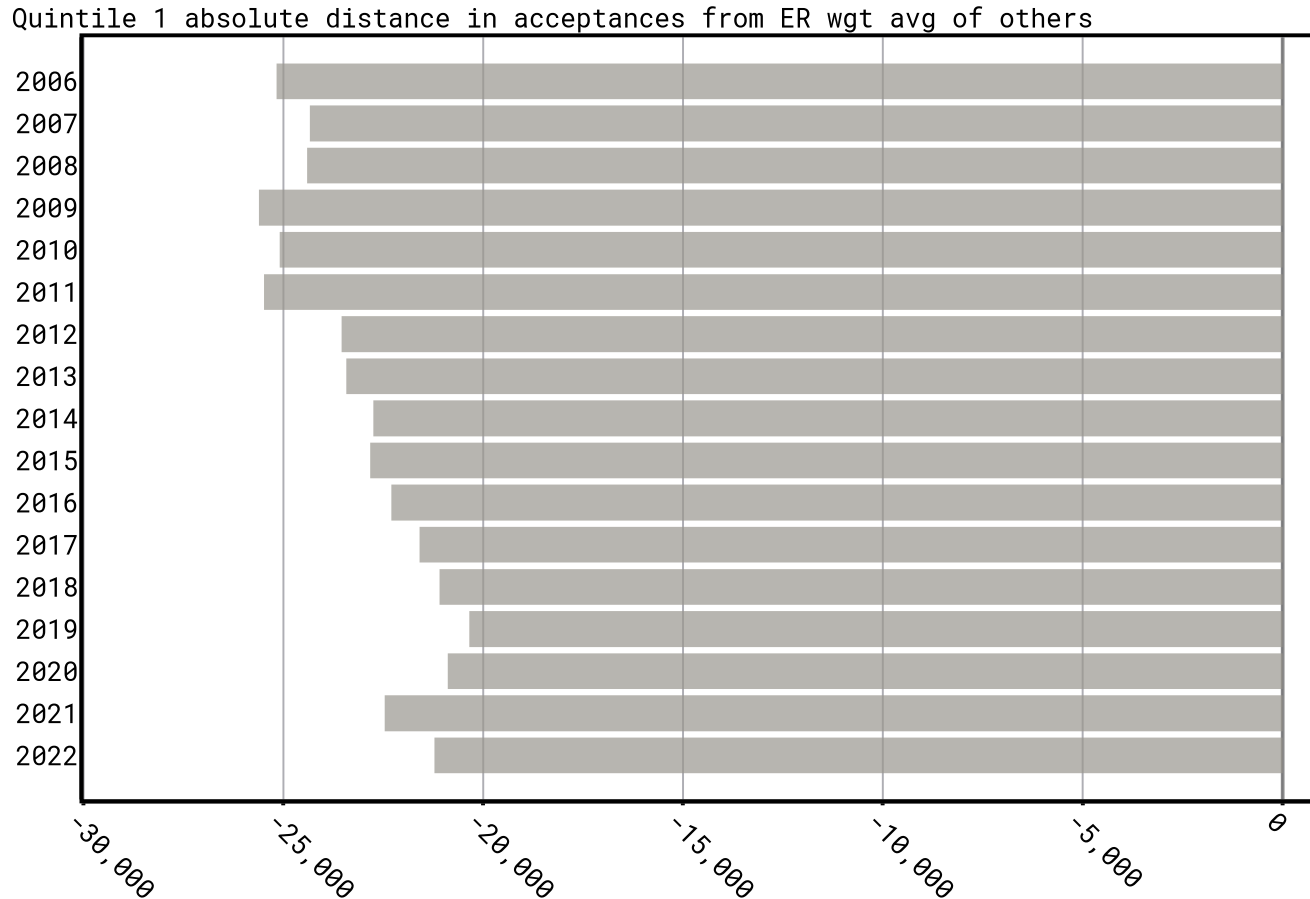
Distance in abs acceptances from ER wgt avg of others: POLAR4

|      |          |          |          |          |          |
|------|----------|----------|----------|----------|----------|
| 2006 | -25,169  | -12,756  | -2,146   | 7,807    | 33,293   |
| 2007 | -24,333  | -12,630  | -2,975   | 8,180    | 32,917   |
| 2008 | -24,404  | -12,557  | -2,761   | 8,233    | 32,788   |
| 2009 | -25,606  | -12,498  | -3,181   | 8,685    | 33,921   |
| 2010 | -25,087  | -13,170  | -2,941   | 9,062    | 33,448   |
| 2011 | -25,481  | -13,457  | -3,281   | 9,352    | 34,327   |
| 2012 | -23,545  | -12,119  | -2,842   | 7,900    | 32,183   |
| 2013 | -23,425  | -12,125  | -2,495   | 7,978    | 31,736   |
| 2014 | -22,748  | -11,118  | -2,339   | 7,773    | 29,954   |
| 2015 | -22,826  | -12,015  | -2,128   | 8,053    | 30,543   |
| 2016 | -22,301  | -11,316  | -2,644   | 7,434    | 30,693   |
| 2017 | -21,594  | -11,473  | -2,360   | 7,157    | 30,197   |
| 2018 | -21,092  | -10,831  | -3,042   | 7,304    | 29,645   |
| 2019 | -20,348  | -11,317  | -2,466   | 7,123    | 29,041   |
| 2020 | -20,884  | -11,435  | -3,591   | 7,480    | 30,589   |
| 2021 | -22,462  | -12,502  | -3,235   | 7,943    | 32,620   |
| 2022 | -21,222  | -12,097  | -2,948   | 7,843    | 30,634   |
|      | $\rho_1$ | $\rho_2$ | $\rho_3$ | $\rho_4$ | $\rho_5$ |

dataHE - Confidential for the Sutton Trust (20221118) oerhm2

One measure of the equality deficit is how many extra entrants are needed to equal the rates of the *other* groups on that dimension.

# Shortfall in entrants: POLAR Q1

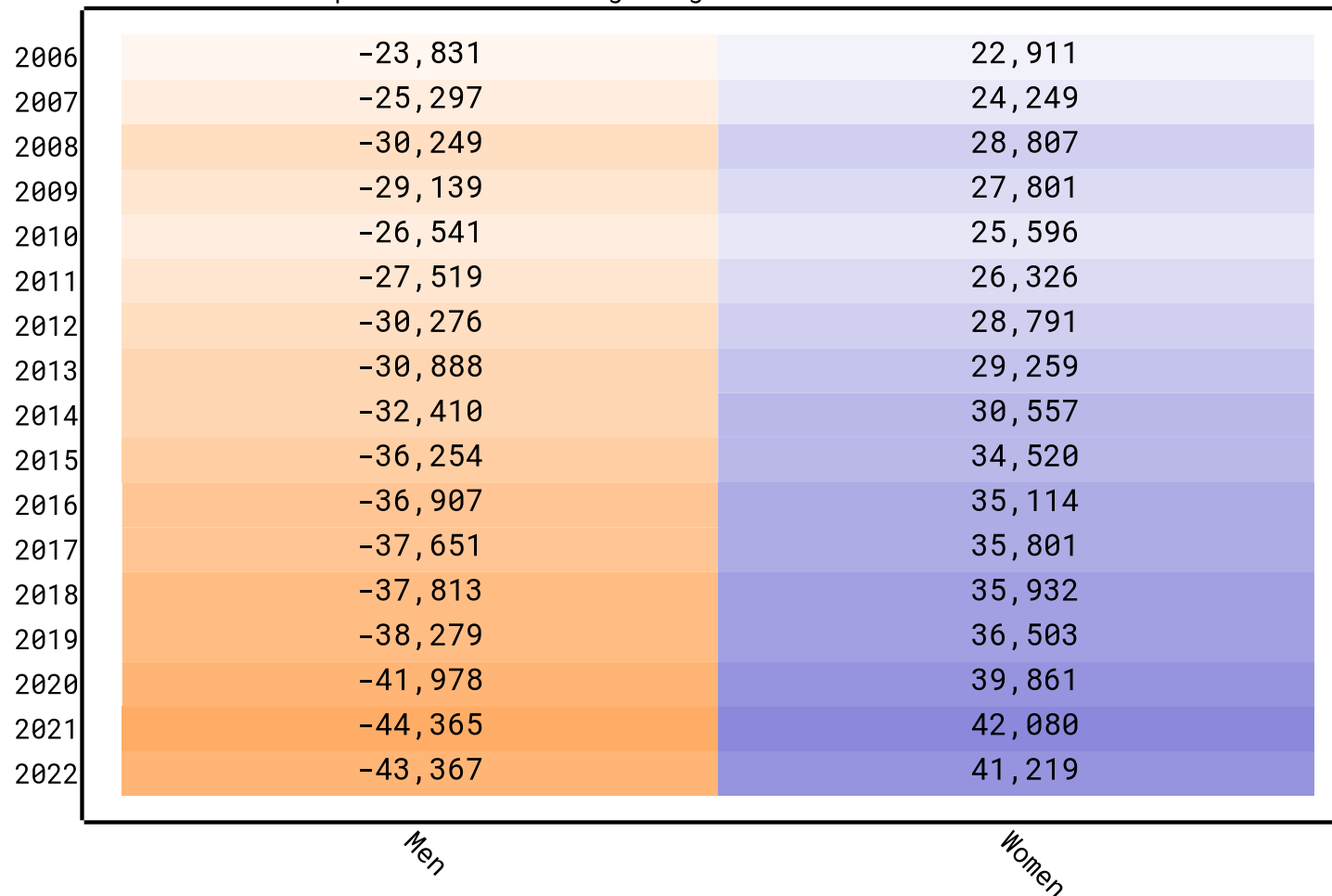


dataHE - Confidential for the Sutton Trust (20221118) boer2\_5

For POLAR Q1 around 20,000 extra entrants are needed to match the entry rate of the rest of the population. This is lower than in the mid-2000s.

# Shortfall in entrants: men

Distance in abs acceptances from ER wgt avg of others: Sex

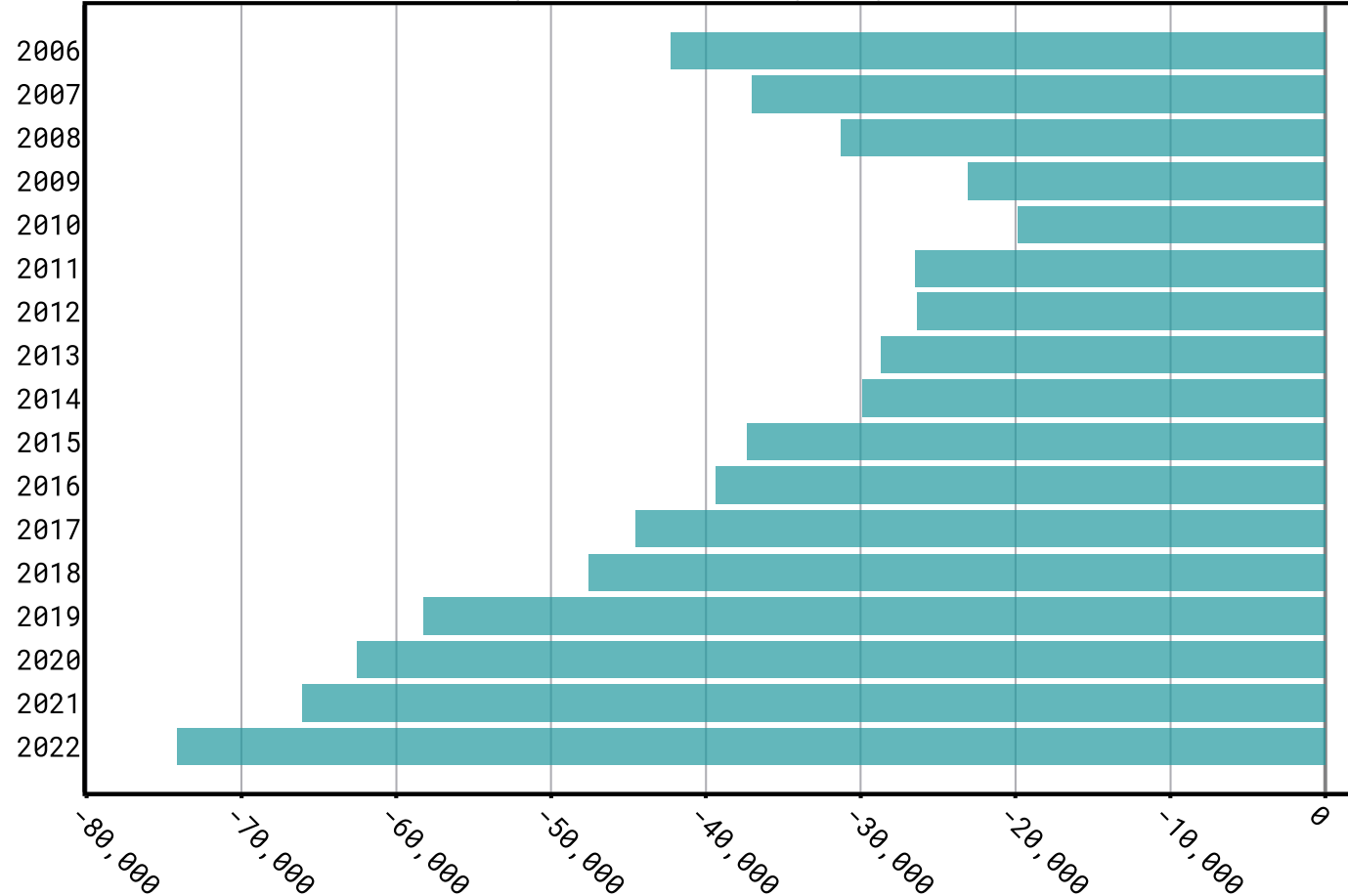


dataHE - Confidential for the Sutton Trust (20221118) oerhm4

Around 44,000 men are now missing in the young entry population – twice the numbers in 2006.

# Shortfall in entrants: White ethnic group

White absolute distance in acceptances from ER wgt avg of others

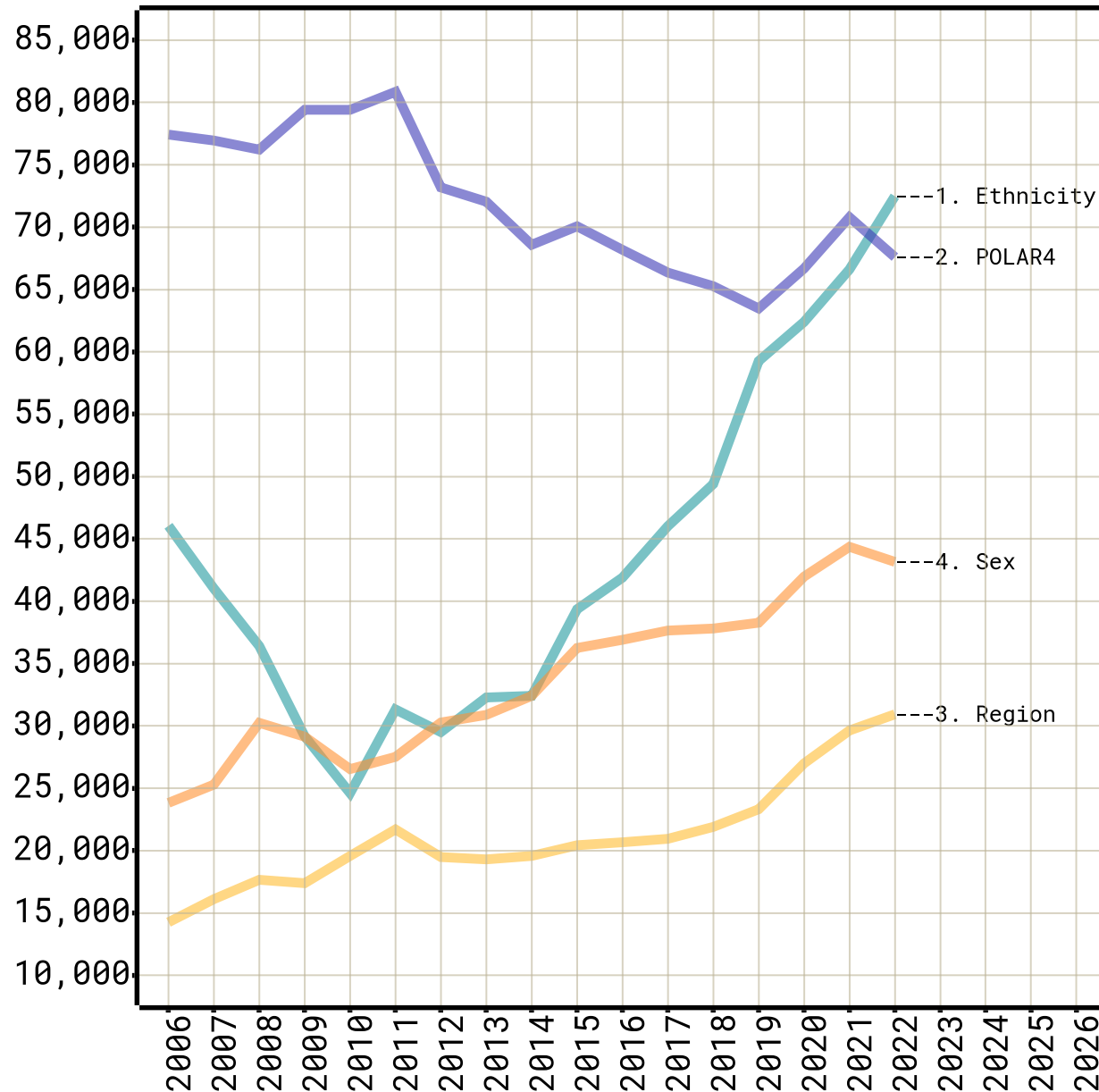


dataHE - Confidential for the Sutton Trust (20221118) boer1\_1

The large population share, low entry rate, and high and rapidly growing comparator rates means the shortfall in young entrants from the White ethnic group is large – now over 70,000 a year, around twice the values at the start of the period

# Numerical equality gaps by dimension

ER total missing under-represented students by equality category



Missing students between under-/over-represented groups by dimension.

- 1. Ethnicity
- 2. POLAR4
- 3. Region
- 4. Sex

Missing by neighbourhood – the key problem in 2006 – has improved. But others, notably ethnic group, worse.

---

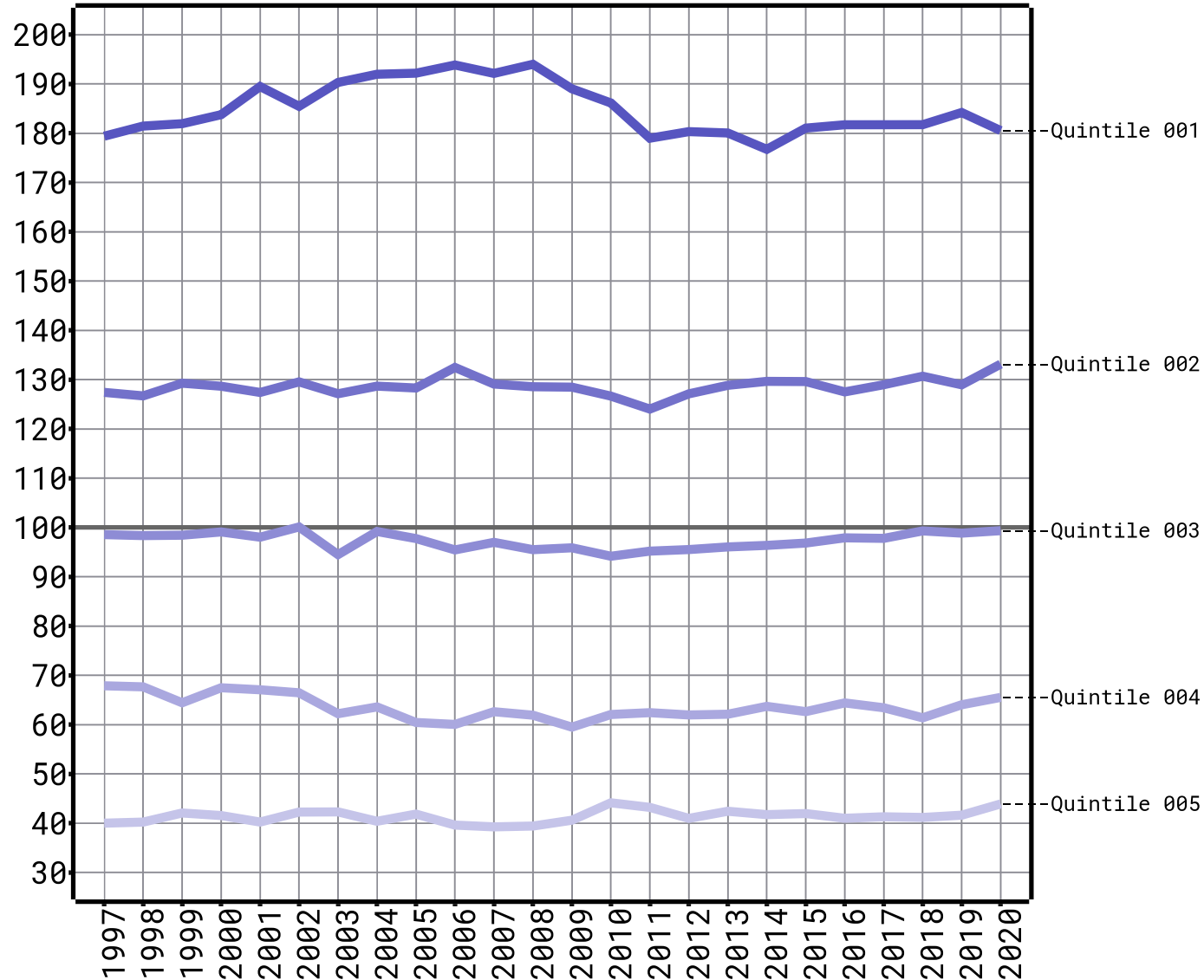
## (4) Thoughts

---



# Fair access is hard

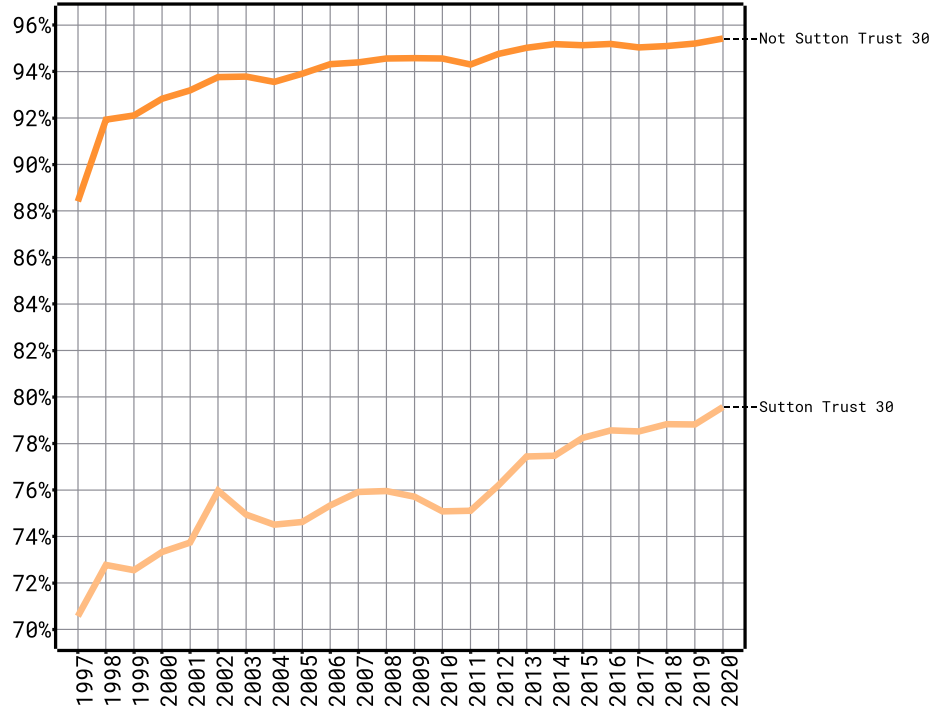
FD/T1a Low participation neighbourhoods: provider quintiles  
Percentage from group (wgt avg=100)



dataHE - Confidential for the Sutton Trust (20230203) ntapq111

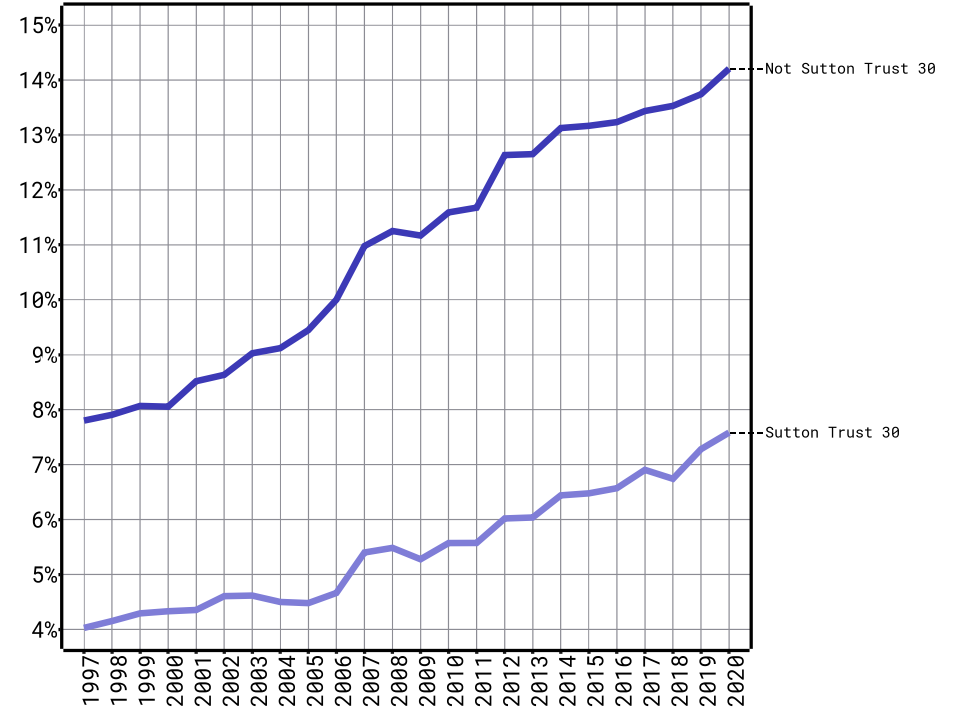
# Raw proportions increasing where focus

FD/T1a State schools or colleges: Sutton Trust 30/non-ST30  
Percentage from disadvantaged group



dataHE - Confidential for the Sutton Trust (20230201) ntasst01

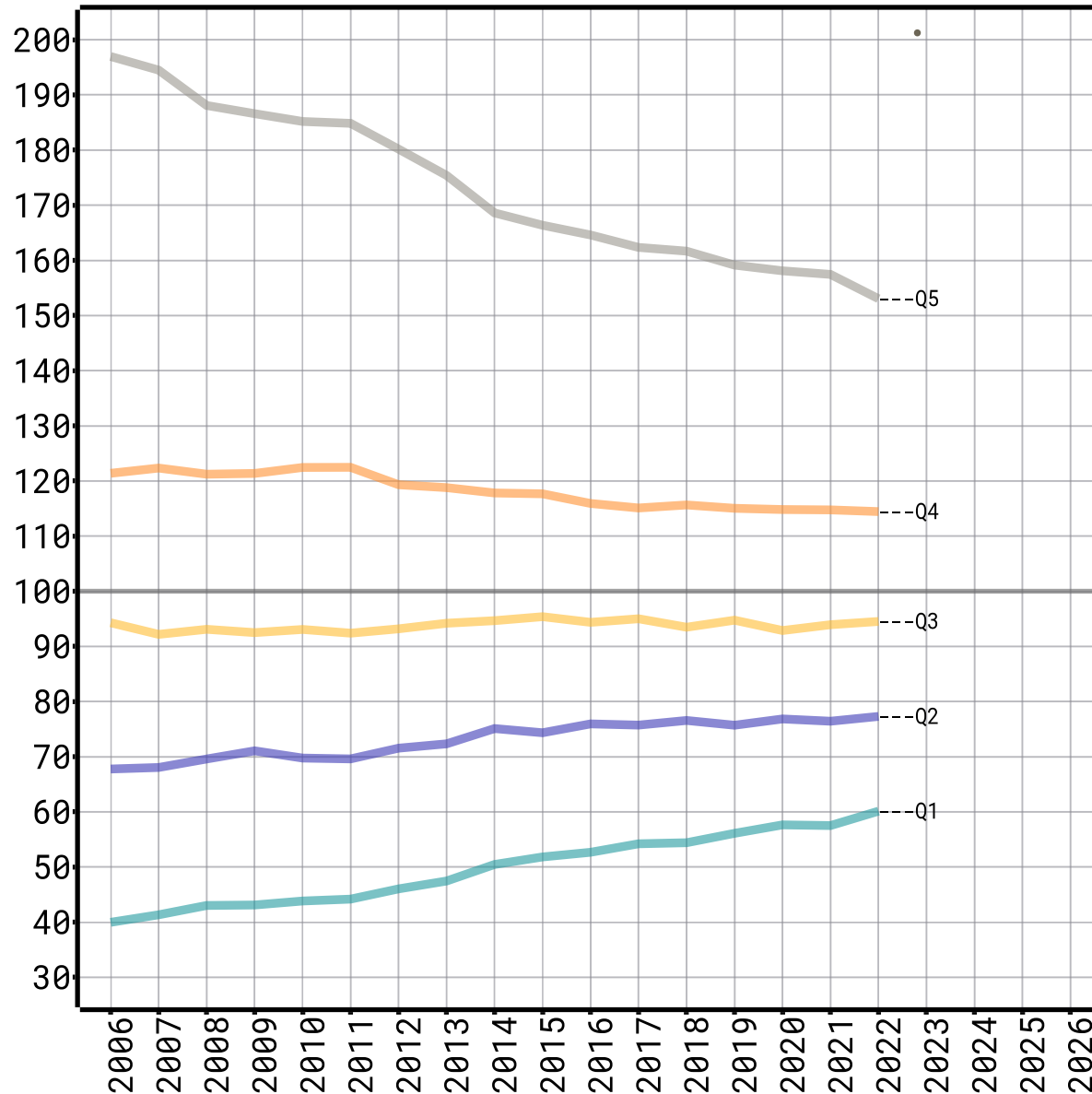
FD/T1a Low participation: Sutton Trust 30/non-ST30  
Percentage from disadvantaged group



dataHE - Confidential for the Sutton Trust (20230201) ntapst01

# Areas driven by national focus?

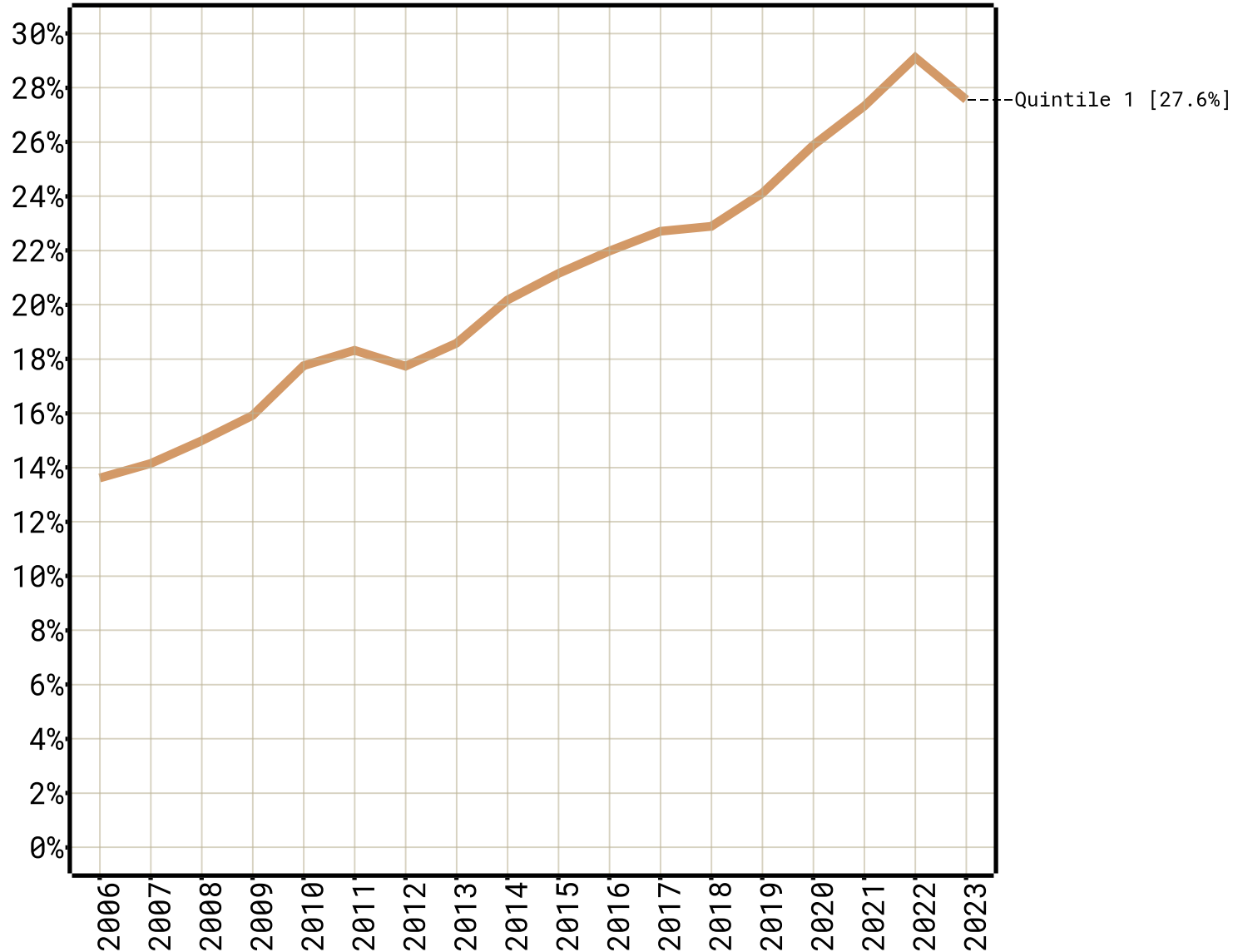
2006-2022 ER isolated dimension indexed: POLAR4



dataHE - Confidential for the Sutton Trust (20221118) ert2

# Application rate is key to outcomes

18 year old UCAS application rates  
MLX, 2023 estimated from January deadline, UK. POLAR4Q1

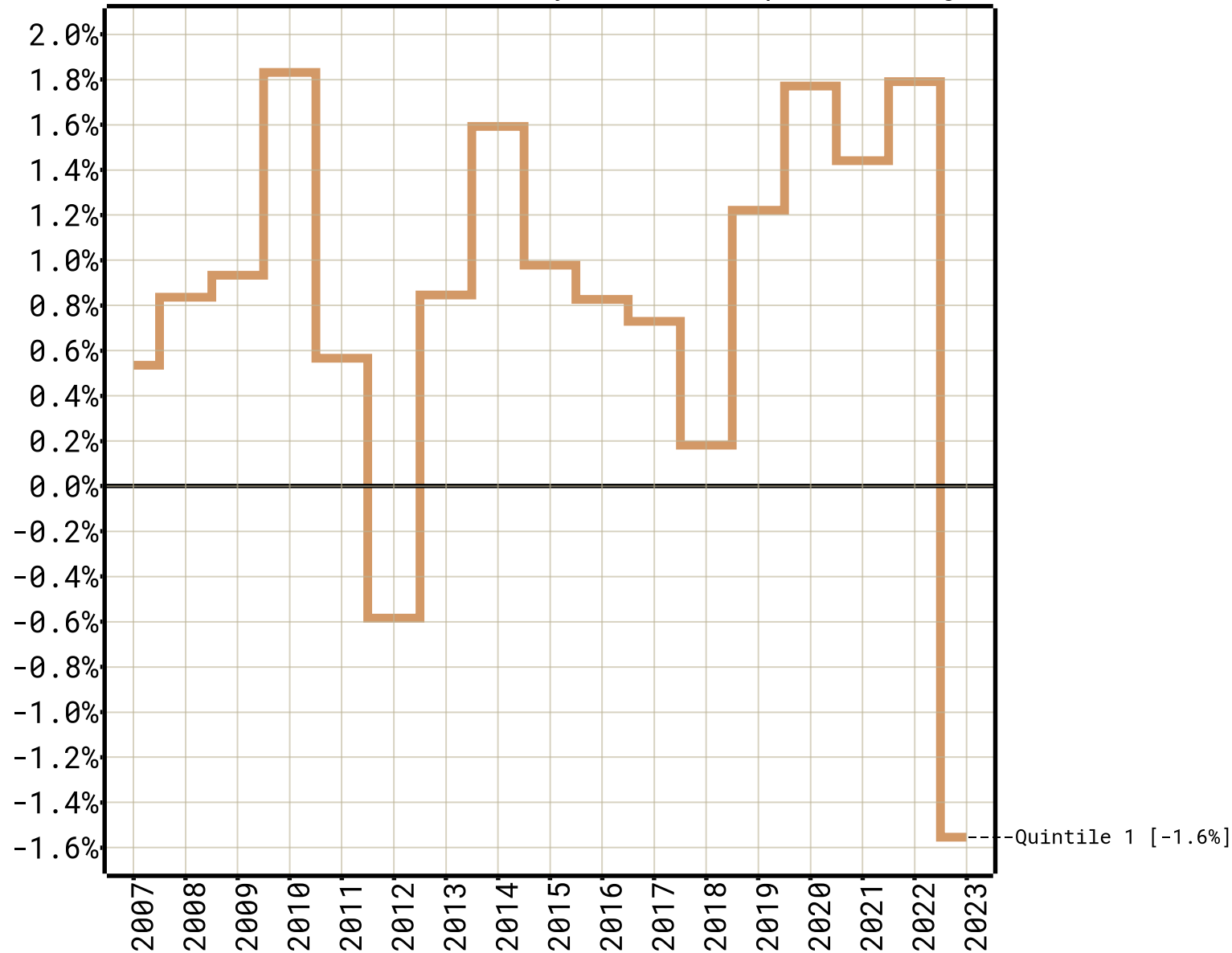


dataHE analysis ofucas.com and ONS.gov.uk data : 03UKQ1L

# So this is a worry

18 year old UCAS application rates

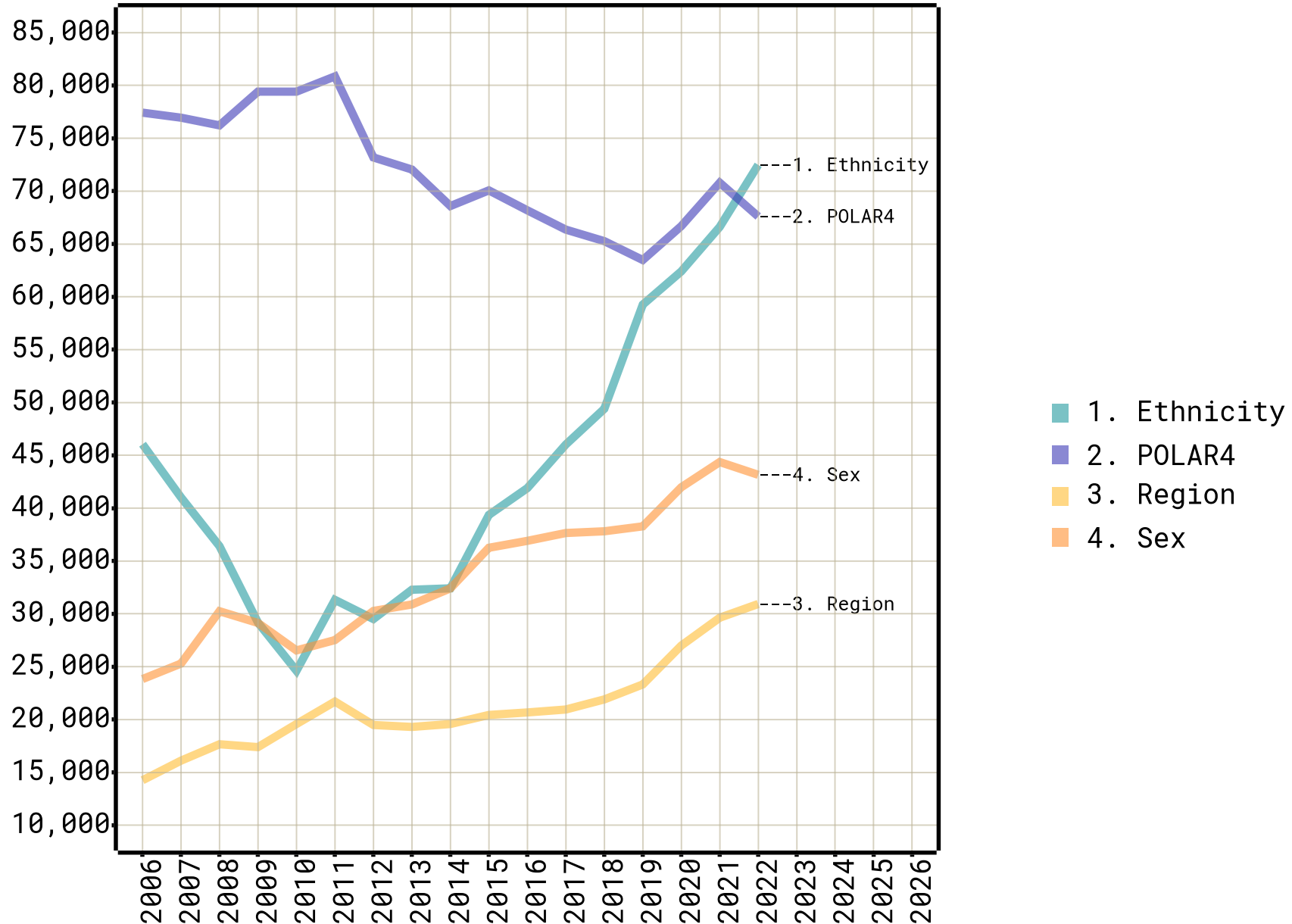
MLX, 2023 estimated from January deadline, point change, POLAR4Q1



dataHE analysis of ucas.com and ONS.gov.uk data : 03UKQ1

# And equality problems grow elsewhere

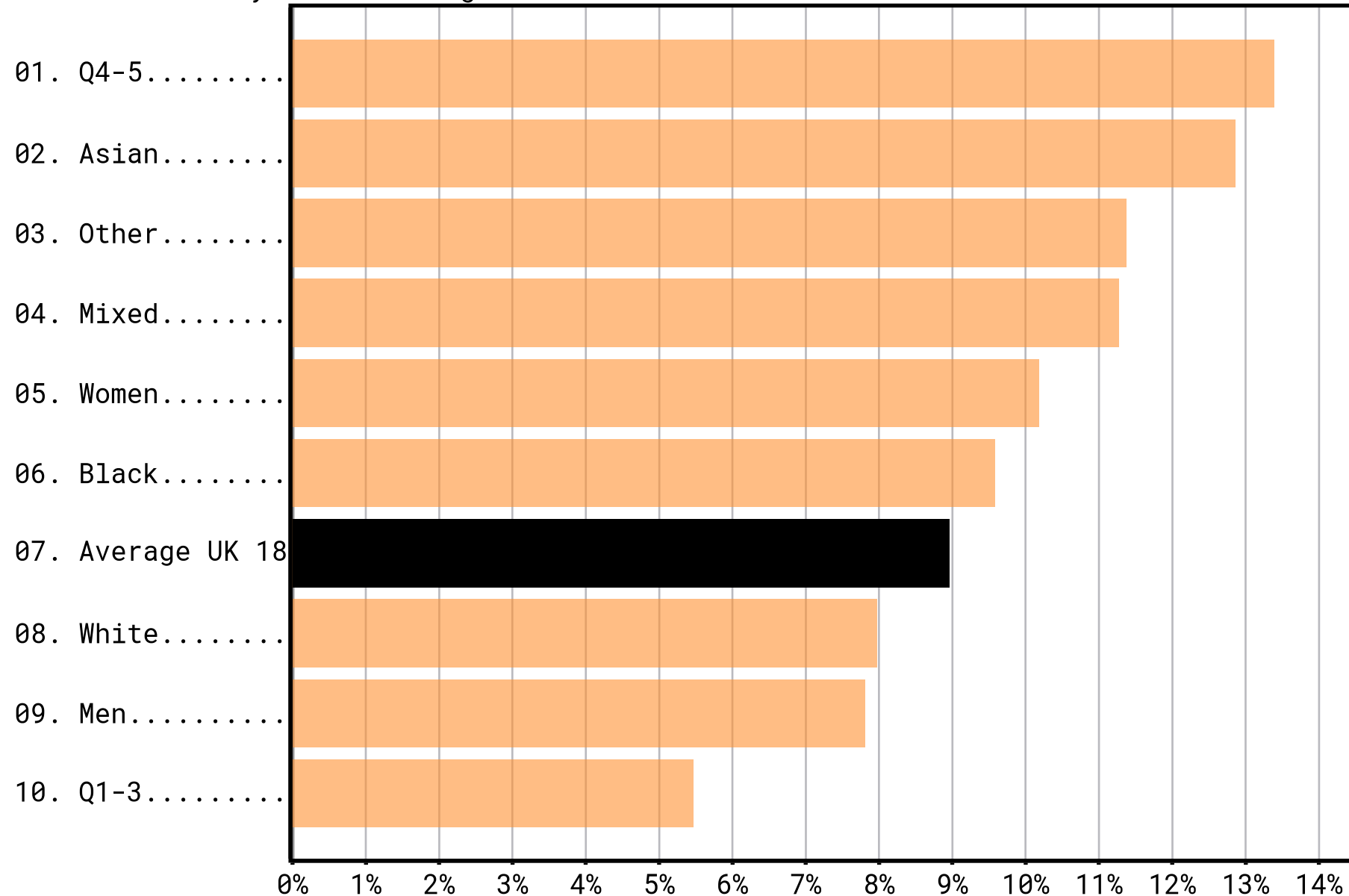
ER total missing under-represented students by equality category



dataHE - Confidential for the Sutton Trust (01/02/23) m11ER

# Not just sector equality: English RG

2022 UK 18 entry rates to English RG



dataHE analysis ofucas.com and ONS.gov.uk data : rgeq1

data