Institute for Fiscal Studies

Widening participation in higher education: analysis using linked administrative data

Haroon Chowdry, Claire Crawford, Lorraine Dearden, Alissa Goodman & Anna Vignoles

Background and motivation

- Rapid expansion of HE in the UK
  - 39% of 17-30 year olds participated full-time in 2008-09 compared to 12% in 1979
- But “widening participation” to students from lower socio-economic backgrounds still of huge policy concern, particularly following the introduction and subsequent increase of tuition fees

- Research questions:
  - How strong is the relationship between HE participation and socio-economic status (including participation at high status universities)?
  - How much of this relationship can be explained by prior achievement?
    - To what extent are barriers at the point of entry (e.g. credit constraints) a problem?

Previous literature

- Most studies compare socio-economic gaps in participation after controlling for other relevant factors (including prior attainment) and use significant results as evidence of the extent of credit constraints (or other barriers arising at the point of entry)
- Mixed results:
  - Cunha et al (2006), Carneiro & Heckman (2002) and Bekhradnia (2003) find little evidence of credit constraints driving HE participation
  - On the other hand, Belley & Lochner (2007), Gayle et al (2002) and Dearden et al (2004) find some evidence for credit constraints

Our contribution

- We use a unique national administrative dataset containing the population of students in England at age 16, who are followed into higher education at any UK university at age 18/19 or 19/20
  - Includes both participants and non-participants
  - Contains very detailed information on prior achievement, measured at ages 11, 14, 16 and 18, allowing us to:
    - Consider when educational inequalities emerge and how these affect HE participation; in particular academic trajectories during secondary school
    - Better allow for unobservable factors that may influence HE participation, assuming that such unobserved factors are also likely to affect earlier achievement as well

Methodology

- Two binary outcomes:
  - HE participation
  - HE participation in a “high status” institution
- Use a linear probability model:

$$HE_{is} = \alpha + \beta_1 SEG_i + \beta_2 X_i + \beta_3 PA_i + \mu_i + \eta_s$$

  - HE is our binary outcome of interest
  - SEG represents a series of dummy variables based on our deprivation index
  - X is a vector of other individual characteristics (e.g. gender, ethnicity, FSM)
  - PA represents a series of variables reflecting the individual’s prior achievement, from age 11 to 18
  - μ is an individual error term and η are school fixed effects
- Models are estimated sequentially

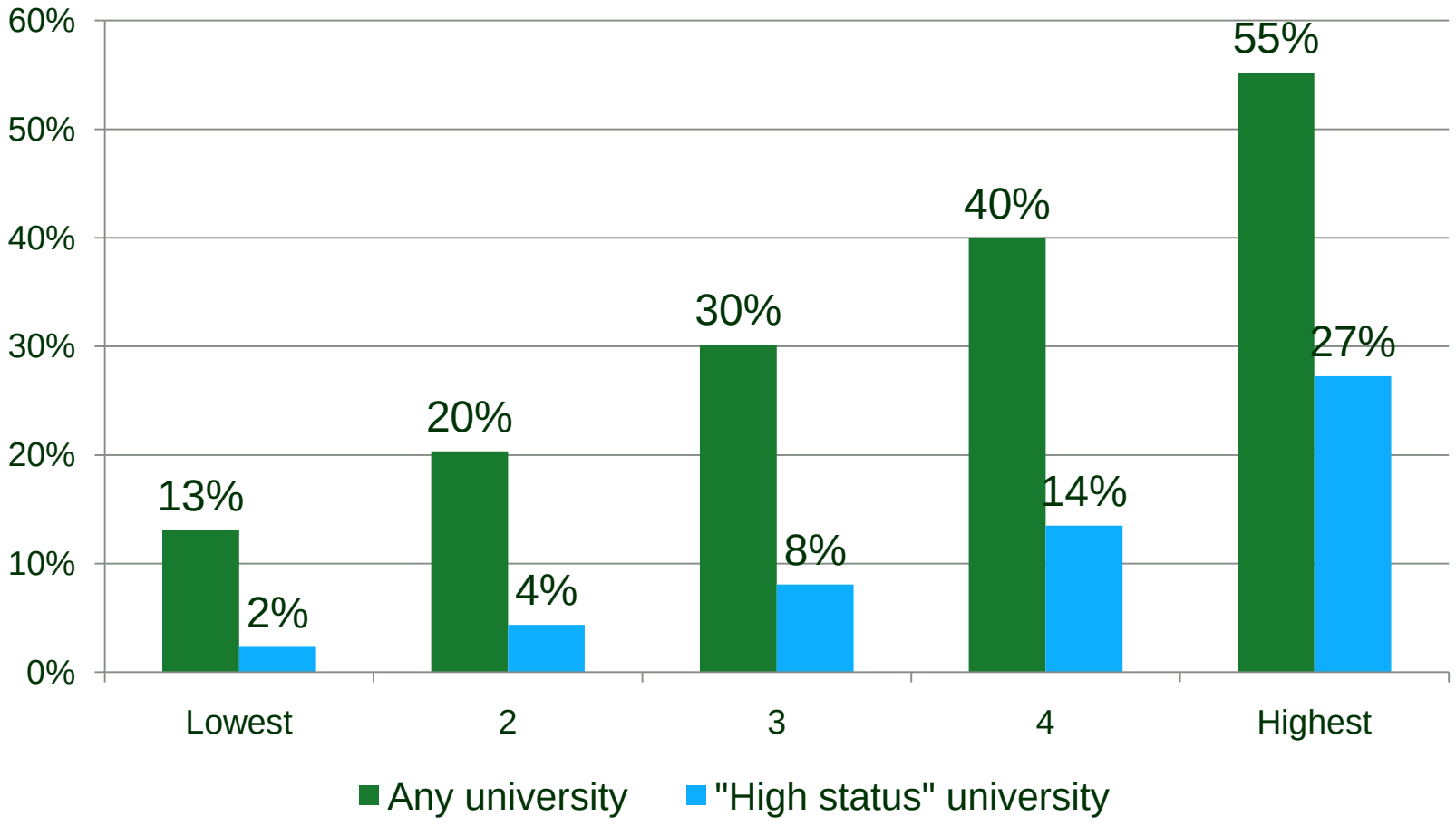
Data

- Linked individual-level administrative data on the population of students in England from schools, FE and HE institutions
- Consider two cohorts: sat GCSEs in 2001-02 or 2002-03
- For state school students:
  - Key Stage 2, 3, 4 and 5 results, plus gender, month of birth, FSM, ethnicity, EAL, SEN, home postcode, school identifiers
- For private school students:
  - Key Stage 4 and 5 results, plus gender, month of birth, school identifiers
- "High status" institutions defined as:
  - Russell Group institutions (20 in total)
  - Plus any UK university with a 2001 average RAE score higher than the lowest amongst the Russell Group (an extra 21 institutions)
  - 35% of participants attend such institutions

Constructing our deprivation index

- Combine (using PCA):
  - Eligibility for free school meals at age 16
  - Geographic information based on home postcode at age 16:
    - IMD score
    - ACORN type
    - OA-level socio-economic status, highest educational qualification and housing tenure
- Do not observe this information for private school students
  - Instead assume they come from higher SES families than most state school pupils and hence place them at the top of the distribution
- Create quintiles and include the four lowest in our model, so all estimates are relative to those in the least deprived quintile

Link between deprivation and HE participation



Institute for Fiscal Studies

Gradients in HE participation for females

	No controls	Individual and school controls	Plus age 11 test results	Plus age 14 test results	Plus age 16 exam results	Plus age 18 exam results
2nd quintile	-0.142** [0.004]	-0.110** [0.003]	-0.085** [0.002]	-0.068** [0.002]	-0.047** [0.002]	-0.024** [0.002]
Middle quintile	-0.251** [0.005]	-0.202** [0.003]	-0.153** [0.003]	-0.121** [0.002]	-0.080** [0.002]	-0.038** [0.002]
4th quintile	-0.362** [0.005]	-0.294** [0.003]	-0.217** [0.003]	-0.168** [0.003]	-0.104** [0.002]	-0.048** [0.002]
Most deprived quintile	-0.446** [0.005]	-0.358** [0.003]	-0.25688 [0.003]	-0.193** [0.003]	-0.113** [0.002]	-0.053** [0.002]
Observations	572,881					

Notes: all specifications include a cohort dummy. Robust standard errors are clustered at school level and are shown in square brackets. ** indicates estimates are significant at the 1% level; * at the 5% level.

Gradients in “high status” participation amongst female HE participants

	No controls	Individual and school controls	Plus age 11 test results	Plus age 14 test results	Plus age 16 exam results	Plus age 18 exam results
2nd quintile	-0.161** [0.006]	-0.051** [0.003]	-0.043** [0.003]	-0.038** [0.003]	-0.034** [0.003]	-0.022** [0.003]
Middle quintile	-0.231** [0.007]	-0.092** [0.004]	-0.075** [0.004]	-0.065** [0.003]	-0.054** [0.003]	-0.032** [0.003]
4th quintile	-0.291** [0.007]	-0.129** [0.004]	0.104** [0.004]	-0.088** [0.004]	-0.070** [0.004]	-0.043** [0.003]
Most deprived quintile	-0.324** [0.007]	-0.152** [0.005]	-0.116** [0.004]	-0.094** [0.004]	-0.069** [0.004]	-0.043** [0.004]
Observations	204,387					

Notes: all specifications include a cohort dummy. Robust standard errors are clustered at school level and are shown in square brackets. ** indicates estimates are significant at the 1% level; * at the 5% level.

Summary of results

- Large and significant raw gradients:
  - Difference between most and least deprived:
    - Over 40 ppts for HE participation
    - Over 30 ppts for “high status” HE participation
  - Slightly steeper for females than males
- These gaps are reduced by around 90% once we control for everything up to attainment at Key Stage 5
- Gap between top two quintiles is of similar magnitude to gap between 2nd and bottom quintiles
  - Suggests much of the gap is driven by high participation at the top
- Some evidence that schools play more of a role in driving participation at high status institutions than HE participation per se

Conclusions

- Students from poorer backgrounds much less likely to go to university than students from richer backgrounds
- Gap substantially reduced – but not entirely eliminated – by the inclusion of controls for prior attainment
- Suggests that interventions aimed at widening participation may be most effective if they are focused on improving attainment in secondary school amongst students from poorer backgrounds