# University Selectivity and the Relative Returns to Higher Education: Evidence from the UK

Ian Walker
(Lancaster University and IZA)
and
**Yu Zhu** (University of Dundee)

Labour Force and Annual Population Surveys
User Conference
20 November 2017

# Contributions of this paper

- Follow-up on Walker & Zhu 2013 BIS RP on graduate life-time earnings in the UK (had no controls for differences in ability across subjects using QLFS 1993-2010)
- First UK study combining nationally representative survey data (QLFS) with mean standardized A-level tariff scores by UG entry year, HEI and subject (JACS) from HESA (Higher Edu Stat Agency)
  - Wages rather than earnings
  - All graduates working as employees, not just Student Loan borrowers
  - Full life-cycle (at least till age 41 for the merged data)
  - Controls for UG degree subjects, degree class, HEI (type) as well as ***college selectivity***
  - Able to estimate ***"value-added"*** by HEI and/or subjects using Inverse Probability Weighted Regression Adjustment (IPWRA)

# Returns to college selectivity:

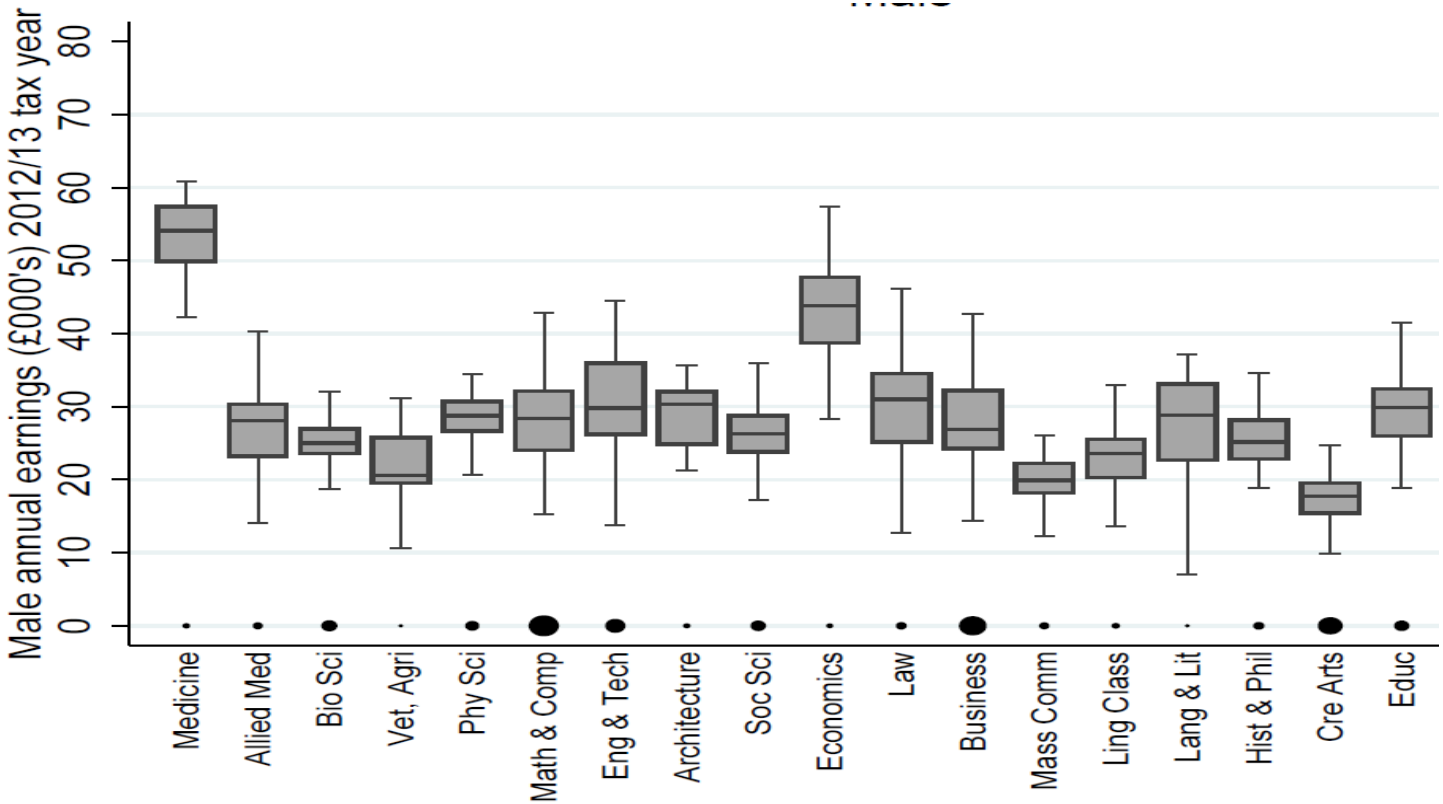
- Loury & Garman (1995, JoLE): omitting college performance overstates the effect of college selectivity for Whites and understates it for Blacks.
  - However, black students with below median SAT scores of the college they attend have lower probability of graduation.
- Brewer et al. (1999, JHR): significant return to attending elite private HEI even after accounting for selection (on tuition fees & financial aid)
- Dale & Krueger (2002 QJE): adjusting for unobservable student characteristics by comparing students who applied, and were accepted and rejected by the same set of colleges, find little evidence of returns to attending more prestigious colleges for students with same ability
- Chen et al. (2012 EL): substantial returns to MBA program selectivity using Dale & Krueger method
- Dale & Krueger (2014 JHR): after partially adjusting for unobserved student ability by controlling for the average SAT score of the colleges that students applied to, estimates of the effects of college characteristics fall substantially and are generally indistinguishable from zero, except for students from disadvantaged background.

# UK Studies on HE selectivity

- Chevalier & Conlon (2003, IZA DP): attending RG leads to 6% wage premium, but PSM imprecise because of thin common support
- Hussain, McNally and Telhaj (2009, CEE DP): 6% earnings difference for 1 SD in HEI quality using Graduate Cohort Studies (early career)
- Broecke (2012, EcEdRev) follows the Dale & Krueger (2002) approach and finds that one standard deviation in selectivity in the UK leads to a 7% increase in earnings
  - However, based on one single cohort of graduates 3.5 years after graduation
  - Response rate was only 26%.
- Walker & Zhu (2013, BIS RP) using QLFS :
  - using BHPS, find that wage differences across HEI types are smaller & become statistically insignificant after including basic family background controls.
  - main analysis distinguishes between subjects but not HEI types due to data limitation
- Britton, Dearden, Shephard & Vignoles (2016, IFS WP): use tax & Student Loan admin data on graduate earnings up to 10 years into labour market:
  - substantial premiums for Medicine, Economics, Law, Maths and Business;
  - Students from higher income families have median earnings 25% (10%) more than those from lower income families, before (after) accounting for HEIs attended and subjects chosen.

# Britton et al. (IFS WP 2016)

- HMRC tax records for very young workers
  - Relies on match to SLC records

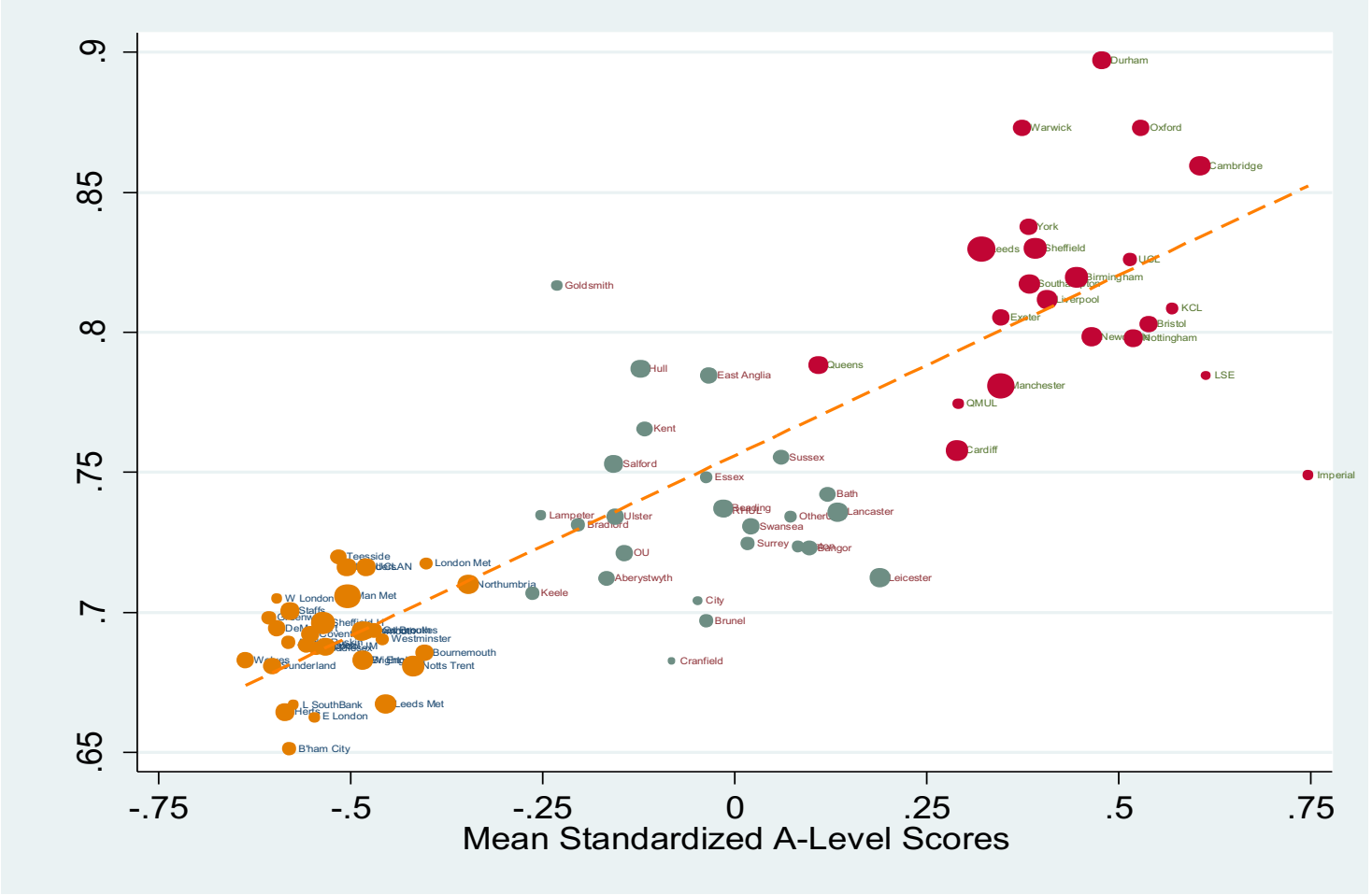


# This paper

- Differences by major*HEI - **relative** returns (ref HEI: Manchester Metropolitan University, or MMU)
- 2012+ Secure Lab LFS contains major and institution (HEI):
  - 21k graduates (post 92 - so still quite young)
  - No self-employed
  - Drop cells < 15
- Micro HESA data contains A-level scores for all students
  - Standardized within each entry cohort, before detrending to allow for subject-specific grade inflation
  - Can now merge detrended standardzied "A-level" scores into micro LFS by cell*cohort
- Lots of variation in selectivity within cells (except top end)
  - Lots of variation across HEIs
  - But not much across subjects | HEI
- OLS controlling for **course** selectivity, as well as age, race, birth decades, region of work

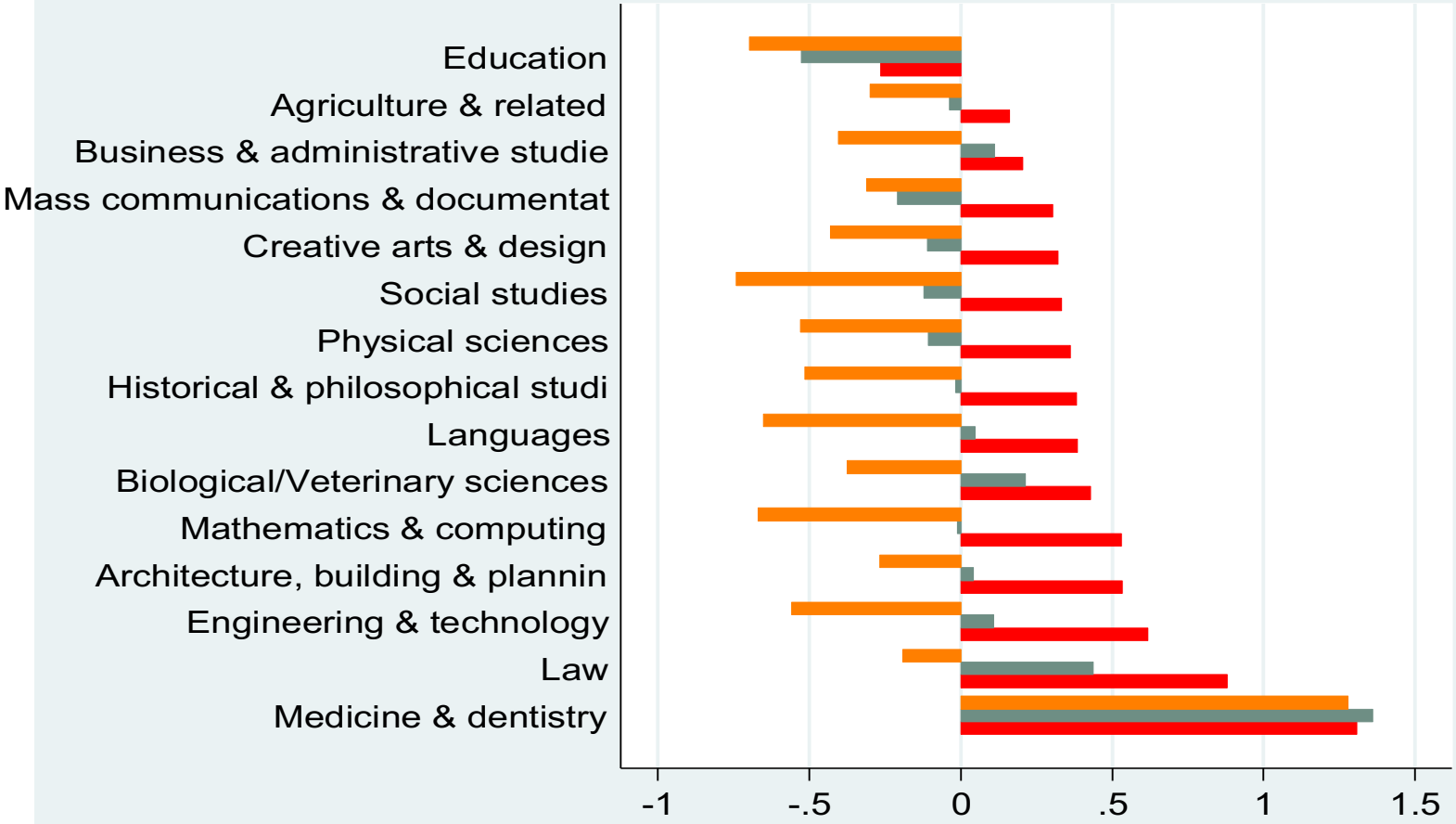
# Variance in ability high for more selective HEIs

## Y axis is SD of tariff scores



# Ability across subject by type of HEI

## Mean tariff scores (standardised)



# Estimates

- Graduates only
  - Relative returns
- Group HEIs
  - RG and Old vs New
- Majors
  - vs Modern languages
- Similar ranking to WZ 2013
  - No course selectivity controls

***Table 4:*** ***Wage equations without PG and degree class controls, various specifications, full graduate sample***

| | Men | | | Women | | |
|---|---|---|---|---|---|---|
| | Baseline | +Subject | +Family | Baseline | +Subject | +Family |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Old university | 0.058*** | 0.065*** | 0.065*** | 0.046*** | 0.046*** | 0.042*** |
| | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) |
| RG university | 0.099*** | 0.091*** | 0.088*** | 0.119*** | 0.106*** | 0.108*** |
| | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) |
| Medicine & dentistry | | 0.426*** | 0.412*** | | 0.448*** | 0.443*** |
| | | (0.038) | (0.037) | | (0.029) | (0.029) |
| Biological/Veterinary sciences | | 0.054* | 0.051* | | 0.081*** | 0.076*** |
| | | (0.029) | (0.029) | | (0.021) | (0.021) |
| Agriculture & related | | 0.012 | 0.004 | | 0.009 | 0.002 |
| | | (0.051) | (0.050) | | (0.040) | (0.040) |
| Physical sciences | | 0.102*** | 0.100*** | | 0.071*** | 0.063** |
| | | (0.029) | (0.028) | | (0.025) | (0.025) |
| Mathematics & computing | | 0.172*** | 0.175*** | | 0.191*** | 0.190*** |
| | | (0.028) | (0.028) | | (0.029) | (0.029) |
| Engineering & technology | | 0.225*** | 0.221*** | | 0.149*** | 0.131*** |
| | | (0.027) | (0.027) | | (0.039) | (0.038) |
| Architecture, building & planning | | 0.142*** | 0.141*** | | 0.059 | 0.059 |
| | | (0.032) | (0.031) | | (0.041) | (0.041) |
| Social studies | | 0.078*** | 0.077*** | | 0.065*** | 0.063*** |
| | | (0.029) | (0.028) | | (0.020) | (0.020) |
| Law | | 0.171*** | 0.168*** | | 0.145*** | 0.144*** |
| | | (0.034) | (0.034) | | (0.028) | (0.027) |
| Business & administrative studies | | 0.208*** | 0.201*** | | 0.152*** | 0.143*** |
| | | (0.028) | (0.028) | | (0.021) | (0.021) |
| Mass communications & documentation | | -0.057 | -0.053 | | 0.021 | 0.013 |
| | | (0.039) | (0.039) | | (0.030) | (0.030) |
| Historical & philosophical studies | | -0.050 | -0.048 | | 0.002 | -0.005 |
| | | (0.032) | (0.032) | | (0.026) | (0.025) |
| Creative arts & design | | -0.045 | -0.046 | | -0.088*** | -0.098*** |
| | | (0.033) | (0.033) | | (0.024) | (0.024) |
| Education | | 0.052* | 0.037 | | 0.123*** | 0.122*** |
| | | (0.030) | (0.030) | | (0.019) | (0.019) |
| Combined subjects | | 0.100*** | 0.102*** | | 0.060*** | 0.059*** |
| | | (0.028) | (0.027) | | (0.019) | (0.019) |
| Constant | -0.272*** | -0.271*** | 0.021 | -0.092 | -0.084 | 0.100 |
| | (0.099) | (0.101) | (0.110) | (0.089) | (0.088) | (0.099) |
| Observations | 10137 | 10137 | 10137 | 10460 | 10460 | 10460 |
| $R^2$ | 0.276 | 0.308 | 0.322 | 0.210 | 0.240 | 0.254 |

Note: Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Full graduate sample (see sample

# Estimated HEI FEs relative to MMU (all subj)

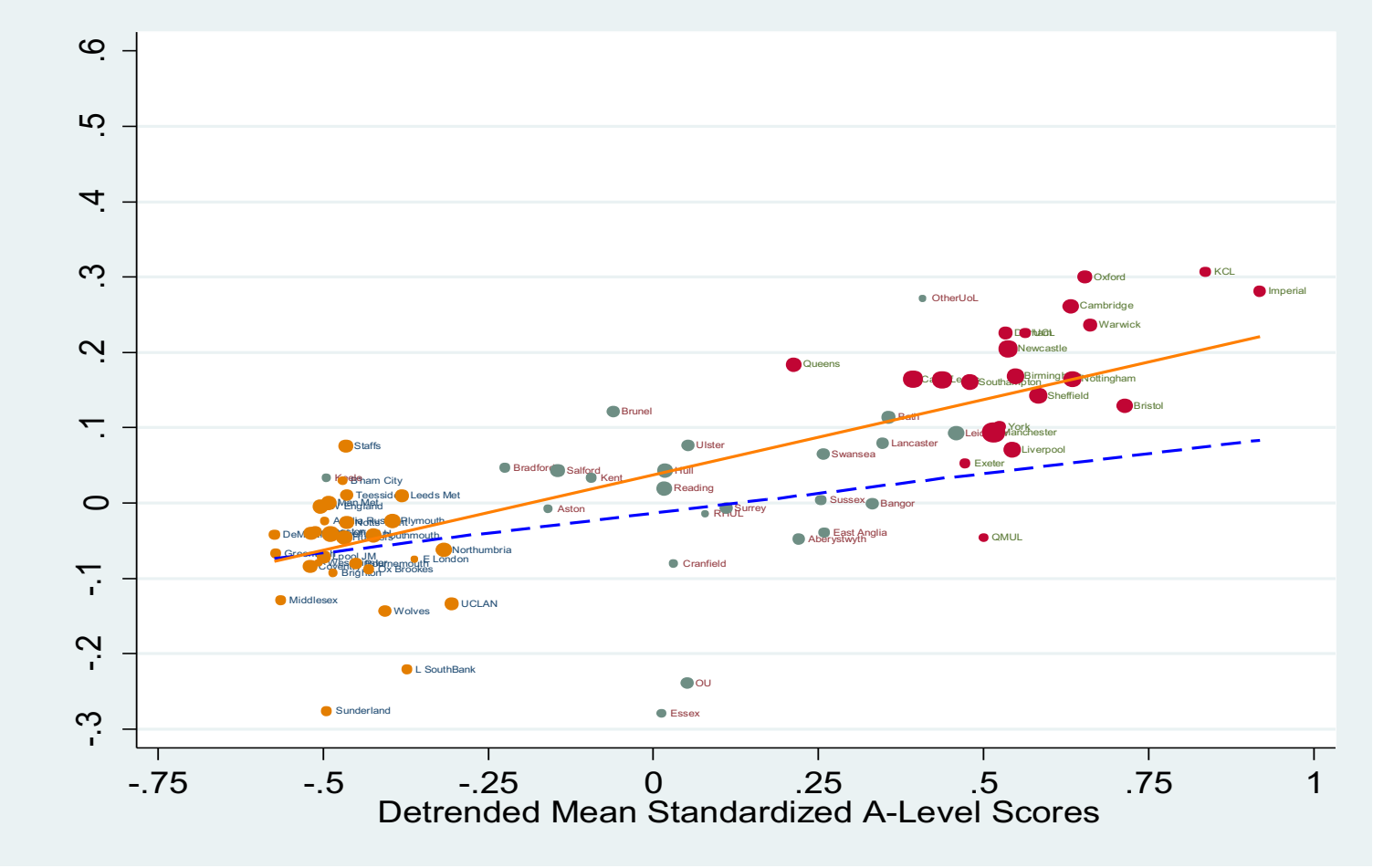
## no selection controls (slope 0.190 (0.013), $R^2$=0.695)

# Estimated HEI FEs relative to MMU (all subj)

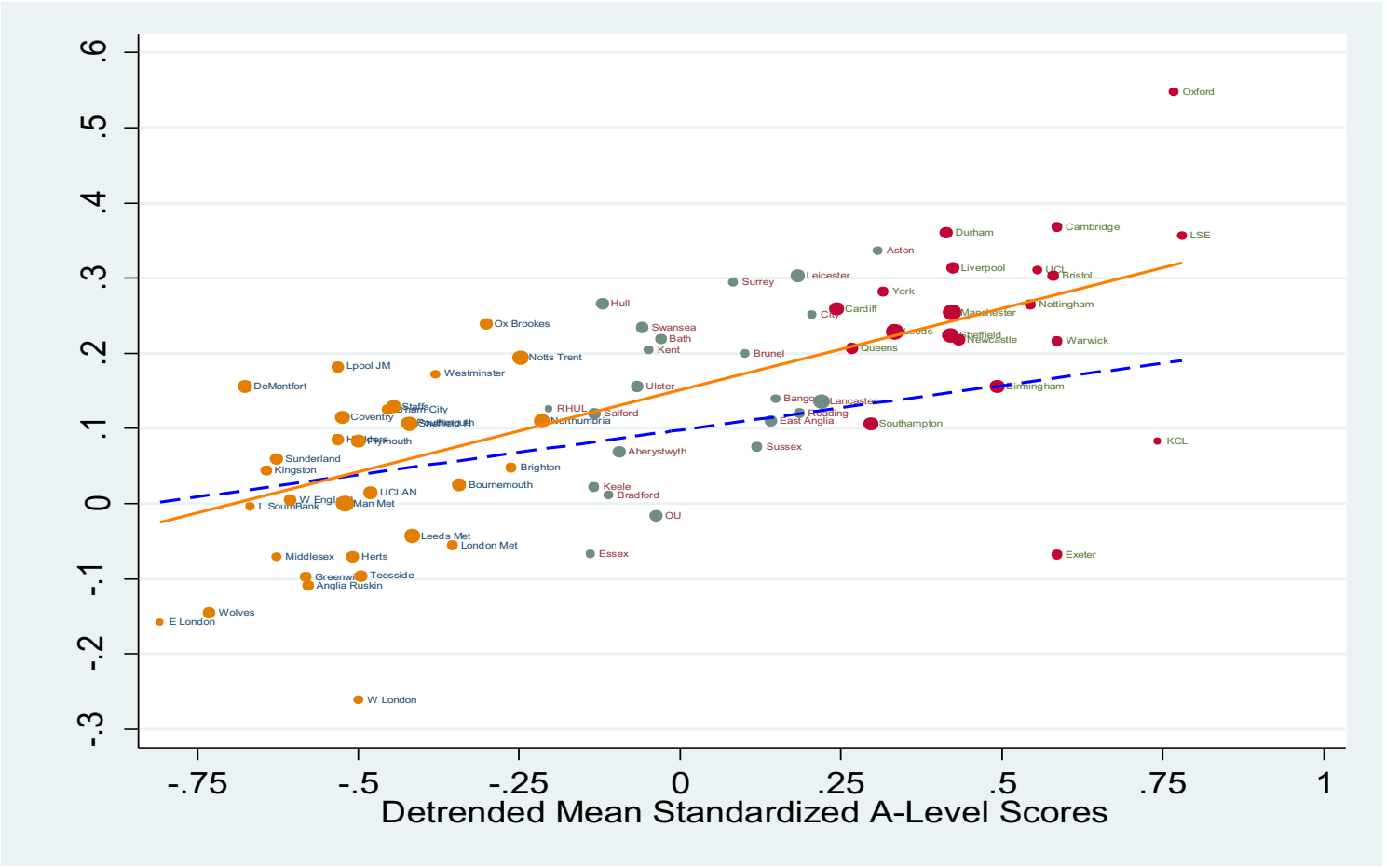
## with selection controls (slope 0.107 (0.013) ), $R^2$=0.394)

# HEI FEs (STEM), slope = 0.205/0.105

## $R^2$=0.612/0.295 without/with selectivity controls

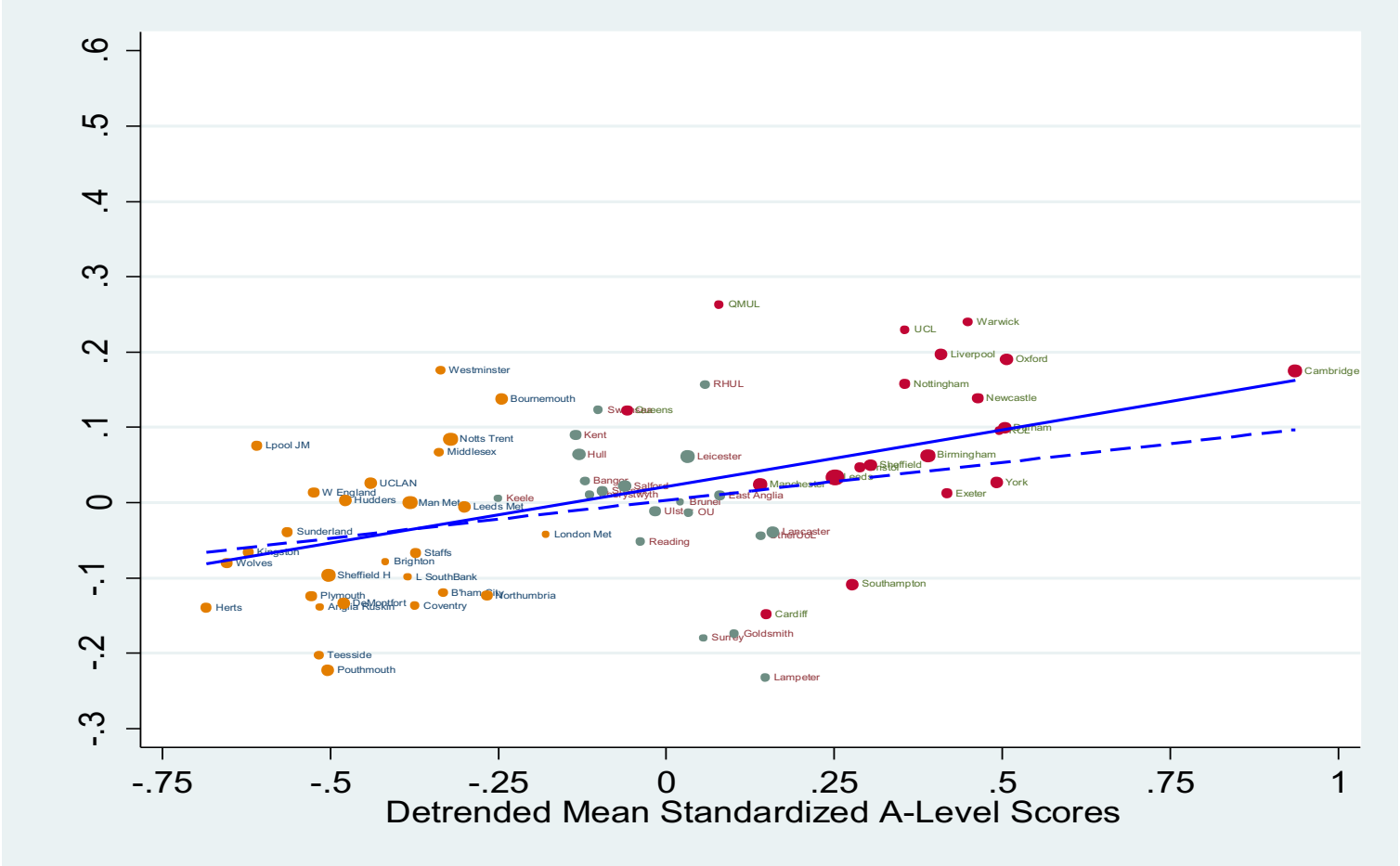


# HEI FEs (Soc Sc), slope = 0.214/0.119
# $R^2$=0.464/0.210 without/with selectivity controls

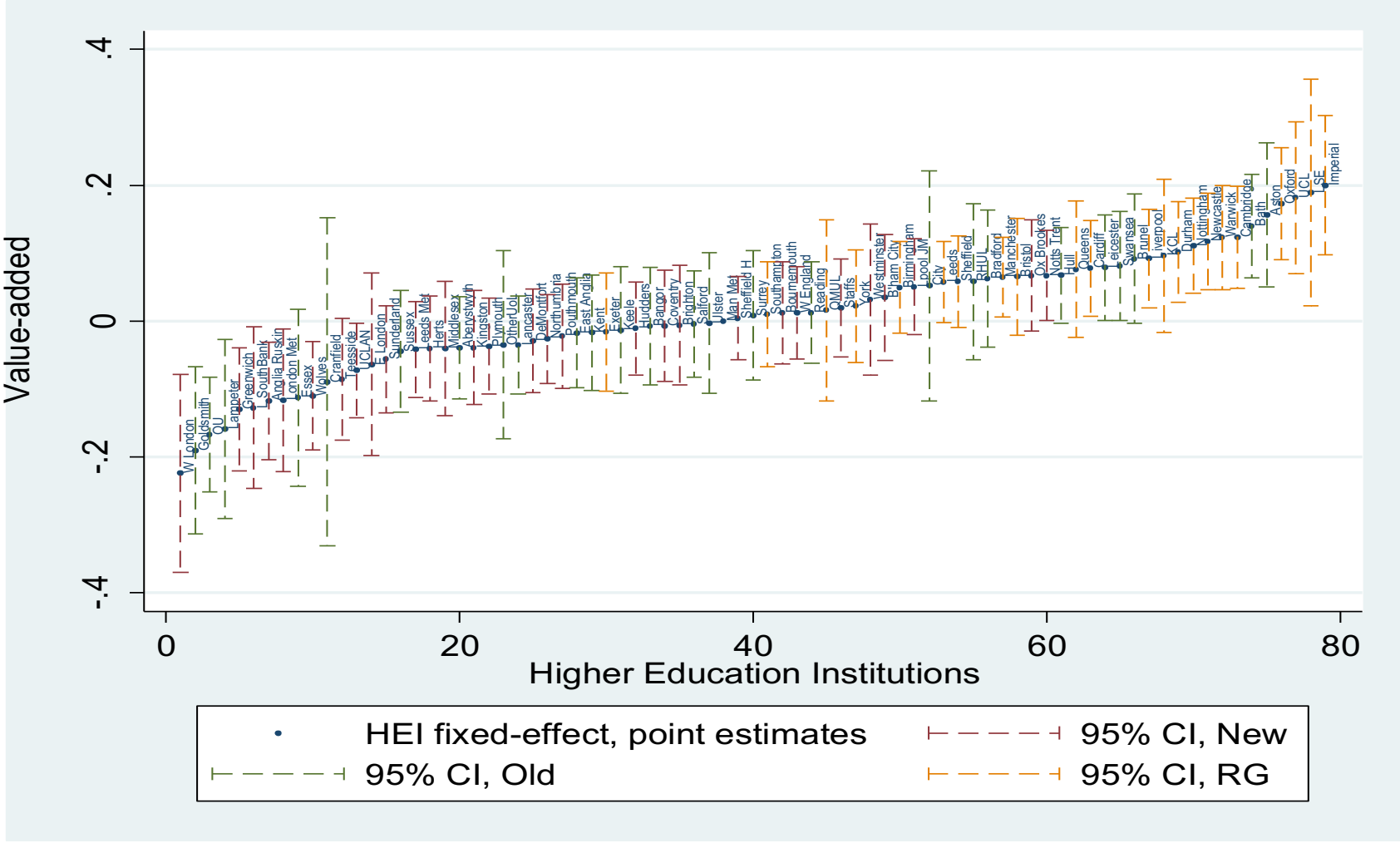


# HEI FEs (Arts & Hum), slope = 0.151/0.100

## $R^2$=0.304/0.162 without/with selectivity controls



# Ranking of HEI value-added (selectivity adj.)



# Inverse-probability weighted regression adjustment (IPWRA)

- IPWRA - MNL version of PS matching (Wooldridge 2007, Imbens & Wooldridge 2009)
  - Step 1: estimate propensity score of treatment using multinomial logit model
  - Step 2: regress outcome on treatment status and covariates, weighting race observations more highly
  - **ATE** "doubly robust"
  - 3 HEI types X 3 broad subjects (relative to New STEM)
  - Very similar results for subj, then HEI type|subj
    - Implies that subject might be treated as exogenous when studying the premium of attending more prestigious universities in the UK

# IPWRA of HEI type & subjects

***Table 7B: IPWRA of HEI type and broad subjects on wages***

| | Treatment Effects | | | OLS | | |
|---|---|---|---|---|---|---|
| | Men (1) | Women (2) | Pooled (3) | Men (4) | Women (5) | Pooled (6) |
| ATE (ref: New - STEM) | | | | | | |
| New - Social Science | 0.075*** | -0.071** | -0.007 | 0.007 | 0.010 | 0.003 |
| | (0.029) | (0.031) | (0.025) | (0.020) | (0.023) | (0.015) |
| New - Arts & Humanities | -0.010 | -0.050 | -0.040 | -0.155*** | -0.065*** | -0.106*** |
| | (0.036) | (0.035) | (0.030) | (0.025) | (0.024) | (0.017) |
| Old - STEM | 0.161*** | -0.047 | 0.054* | 0.038* | -0.053* | -0.001 |
| | (0.031) | (0.037) | (0.028) | (0.022) | (0.028) | (0.017) |
| Old - Social Science | 0.133*** | -0.036 | 0.045 | 0.085*** | 0.032 | 0.055*** |
| | (0.036) | (0.033) | (0.028) | (0.028) | (0.028) | (0.020) |
| Old - Arts & Humanities | 0.003 | -0.156*** | -0.077*** | -0.090*** | -0.121*** | -0.110*** |
| | (0.037) | (0.033) | (0.028) | (0.031) | (0.030) | (0.021) |
| RG - STEM | 0.153*** | 0.039 | 0.102*** | 0.115*** | 0.071*** | 0.092*** |
| | (0.033) | (0.041) | (0.030) | (0.023) | (0.027) | (0.017) |
| RG - Social Science | 0.196*** | 0.049 | 0.108*** | 0.109*** | 0.070** | 0.086*** |
| | (0.041) | (0.038) | (0.031) | (0.027) | (0.029) | (0.020) |
| RG - Arts & Humanities | 0.027 | -0.119*** | -0.065** | -0.104*** | -0.072*** | -0.084*** |
| | (0.039) | (0.038) | (0.032) | (0.030) | (0.028) | (0.020) |
| Female | | | | | | -0.131*** |
| | | | | | | (0.009) |
| Selectivity | | | | 0.039*** | 0.047*** | 0.042*** |
| | | | | (0.014) | (0.014) | (0.010) |
| Observations | 3950 | 4083 | 8138 | 4072 | 4089 | 8161 |
| $R^2$ | | | | 0.404 | 0.355 | 0.391 |

# Discussion

- Selective HEIs don't add as much value as they think
  - Unit change in mean selectivity adds 11% (not 19%)
    - 11%, 12% and 10% for STEM, Soc Sc and Arts & Hum respectively
    - Drive to get low SES kids into elite HEIs - Misguided?
- A lot of variation across HEIs
  - A Cambridge grad could have done as well going to Aston!
- Lower returns to Arts & Hum, relative to STEM/Soc Sc
  - RG/Old premium for STEM and Soc Sc only (with the exception of Old-STEM for women)

# Conclusion

- UK HEIs have been evaluated for research quality for 30 years
  - Research Excellence Framework (REF)
  - Dominated by Oxbridge, Imperial, UCL, LSE etc.
- But not for teaching quality
- Until **NOW** …….
  - **Teaching Excellence Framework (TEF)** outcome announced on 22nd June (e.g. LSE was rated BRONZE while Lancaster/Dundee were rated GOLD)
  - which is what LEO is for

# Further work using LEO (with IFS)

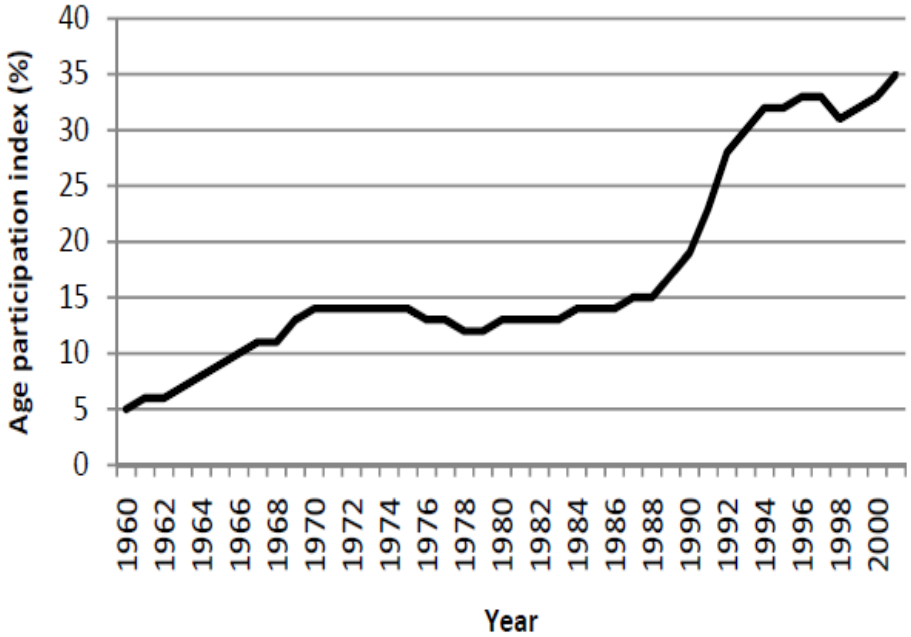
- Income tax records for all 6m+ English taxpayers born 1980+
  - *Will* include self-employed
- Merged into NPD (English) schooling (and FE) records
  - Not Scotland/Wales/NI
  - No non-cog skills in NPD (but they are in LSYPE subset)
  - still not causal, non-cog skills matter for wages and are correlated with cog skills
- Merged to individual HESA HE records
  - But not (yet) to UCAS university choice data
  - And not (yet) SLC records (to get parental income)
- Much bigger than LFS, no measurement error, full controls for cognitive ability, self-employed income
  - But no data on hours so cannot compute wage rate
  - Young cohorts only, so hard to infer lifecycle income

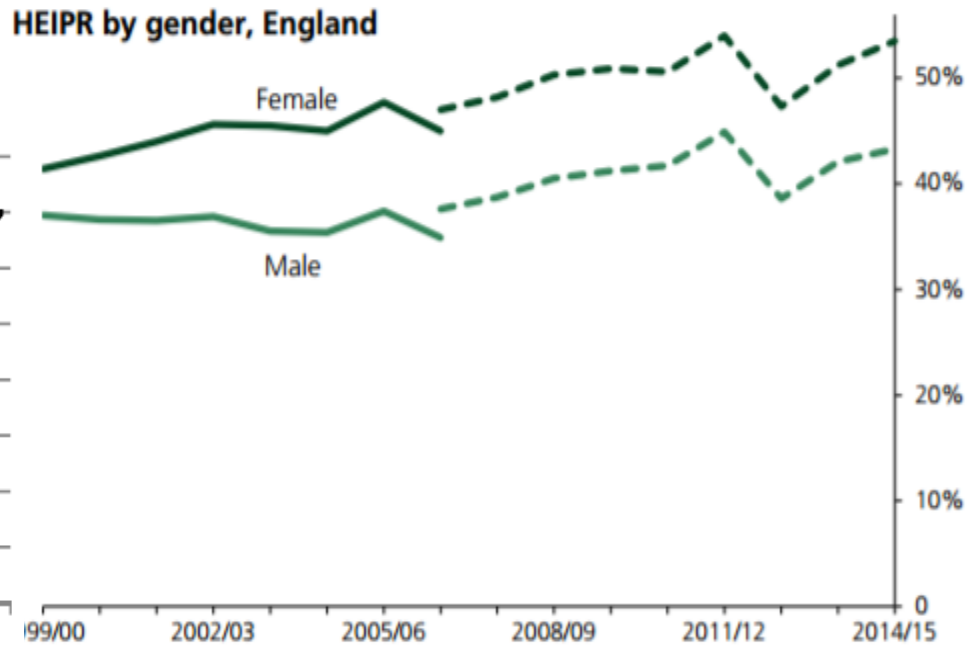
# Appendix

- **Extra Background Information**

# Background to HE participation

- Large and rapid changes to UK HE
  - HE participation rate rose from 7% (1965) of cohort to 14% (1970-90) to 30% (1995).
  - Now close to 50% (thanks to 05/6 definition change)!



# Heterogeneity in UK HE

- 30 universities until 1963
  - 6 ancient (elite) institutions + Victorian institutions
  - Strong tradition of residential university
- Mid-late 60's see about 20 more universities
  - Warwick, Essex, York, ……. Lancaster
  - Some FE colleges become "Polys"
- 1993 end of "binary divide"
  - 35 Polys become Universities
  - "Russell Group" / Old / New
- Lots of heterogeneity in admission criteria
  - Across HEIs
  - But not much across (most) majors | HEI

# Background to HE funding (in England)

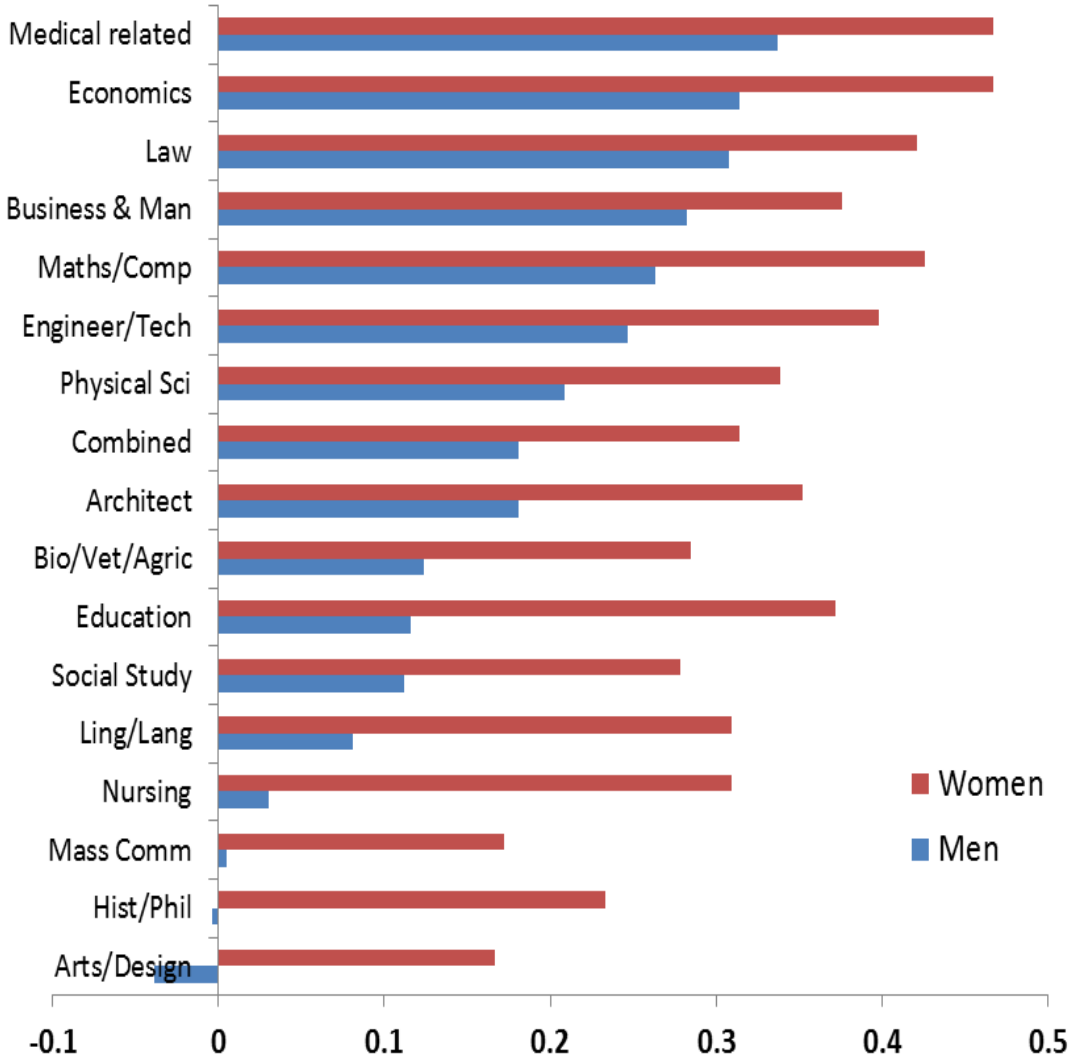
- Mass HE made system financially untenable
  - Pre 1988: Large (means tested) grants and zero fees
    Taxpayer subsidy to unis £6k pppa
  - 1988+: Loans replaced grants (except low income)
    Upfront fees £1k pa for all
  - 2004: £3k pa fees, income contingent loans for fees and subsistence.
    Taxpayer subsidy to unis £3k pppa
  - 2010: £9kpa fees. HEI bursaries (for low income)
    Taxpayer subsidy to unis ends
- Different (and complex) rule for Scotland
  - Excluded from analysis

# Baseline literature

- Greater attention to returns to HE
  - Average, marginal, and observable heterogeneity
  - Changes over time and across cohorts
- But selection bias difficult to overcome in UK
  - Matching/IV: Blundell et al **JRSSa** 2005
    - Rich/small NCDS 27% relative to 2+ A-levels
  - OLS: Walker/Zhu **ScanJE** 2008, **EconEducRev** 2011
    - Poor/small LFS approx 30% relative to 2+ A-levels
  - Matching: Britton et al **IFS WP** 2016
    - Poor/large HMRC large relative returns similar to IWYZ

# Major and HEI relative returns

- Walker/Zhu 2013
  - Lifecycle (allowing for continued SBTC) - net
- LFS
  - Large (OLS) major differentials
  - Much higher, absolute returns
    - despite higher fees and loans
- BHPS
  - Small HEI type effects
    - insignificant



# College choice and college selectivity

- Davies & Guppy (1997, Social Forces): SES predicts entry into selective colleges and lucrative subjects within selective colleges, but not subjects directly; men more likely to enter fields of study with higher economic return.
- Hoxby (2009, JEP): US colleges not more selective over the past 4 decades, except at the very top end; changes mostly due to falling costs of distance and information.
- Smith (2013, EcEdRev): using twins data with application and enrolment information, find that student's probability of bachelor's degree completion within four years increases by 5 percentage points by choosing an institution with a median SAT score 100 points higher than the alternative.
- Goodman et al. (2015, EcEdRev): own college choice is partially affected by college choice made by one's older sibling.

# IPWRA estimates

- Matching to subjects and institutions
  - Based on A-level subjects and scores
  - And on location of home and school
  - And on history of destinations of previous school cohorts
- Looks like sufficient support for *relative* value-added estimation
  - Non-graduate data needed to see if there is enough support to estimate ***absolute returns***
- PG will be more difficult
  - Cell sizes