High Education Sorting and Social Mobility

Jack Britton
(IFS)

Mónica Costa-Dias
(UoB and IFS)

David Goll
(IFS and UCL)

Annual Conference on Structural Dynamic Models
Bonn, August 2021

Motivation

- Education is a key determinant of economic wellbeing
- And a main facilitator of social (inter-generational) mobility
- While returns to HE are large, they are also highly heterogeneous
  - ⋆ by the characteristics of HE programmes
  - ⋆ those of students
  - ⋆ and how the two align
- Those from poorer backgrounds benefit less from the HE system (Black et al. 2015, Campbell et al. 2019, Chetty et al. 2020)
- Here we investigate why and what policy can do about it

Questions

1. How do students sort into HE programmes in England?

   → In particular, to what extent is sorting determined by the abilities of students, the characteristics of programmes and the returns to the investment?

2. In turn, what explains the variation in the labour market outcomes of University graduates?

3. To what extent is sorting into HE keeping low SES children from benefiting?

4. How does HE policy affect all this?

What we do

Build a structural model of life-cycle labour supply, earnings and education with an embedded equilibrium model of the HE market

→ Separates preferences of students for HE programmes, from those of universities for students

Estimates based only on demand side fail to account for (unobserved) choice sets

Exploit instrumental variation and structure of market for identification of preferences and returns to HE

Use model to understand implications of implemented reforms

And to do counterfactual analysis of HE policy

Key findings

1. Equilibrium model of HE market fits data well and predicts well reduced for estimates of the impact of tuition and loan reforms
2. Strong assortativeness between students' skills and programmes' characteristics
3. Some complementarity in earnings between student's skills and the characteristics of programmes
4. Once selection is accounted for, impact of quality is small but significant
5. Demand-side policies have small impacts on the sorting of students and social mobility
6. Supply-side policies (quotas) can be more effective, but at a cost

Background and contributions

Returns to university quality and field of study Dale & Krueger 2002, 2014, Back & Smith 2006, Broecke 2012, Hastings et al. 2013, Kierkeboen et al 2017, Anelli 2018, Dillon & Smith 2020)

High education choices (Keane and Wolpin 1997, 2001, Arcidiacono 2004, Wiswall and Zafar 2015, Delavande and Zafar 2019)

Modelling HE market (Arcidiacon 2005, Epple et al 2006, Fu 2014, Kapor 2020)

Identification in matching markets (Agarval 2015, Agarval and Diamond 2017)

Grant, tuition and loan policies (Denning et al 2019, Epple et al. 2006, Hubner 2012, Azmat and Simion 2020)

Institutional features: the HE market

- Students apply to University in their final HS year, before sitting final exams
  - ⋆ Choose up to 5 programmes (subject×university combinations)
  - ⋆ Student ability demonstrated on test scores at age 16 (GCSEs), University interviews and tests
- The University system is public
  - ⋆ Universities have a set number of places
  - ⋆ Charge tuitions that are regulated by law
- Sorting
  - ⋆ Universities make offers (conditional on final results)
  - ⋆ Once results are known, students take place of their choice to which they match the offer
  - ⋆ Remaining open places allocated by 'clearing'

HE funding policy

2006-2011

- Tuition fees capped at £3,000 per year, apply to every student
- No up-front costs: student loans cover tuition + living costs (up to a cap)
- Loans repaid at 9% of income above a threshold after graduation (£15,000 in 2011)
- Interest rate averaged 1.5%
- Outstanding debt forgiven after 25 years

In 2012

- Big fee increases (trebled to £9,000)
- Interest rates increased to RPI + 3%
- Repayment threshold raised (£15,000 to £21,000)
- Loan term extended (25 to 30 years)

repayment schedule

Data: Longitudinal Educational Outcomes (LEO)

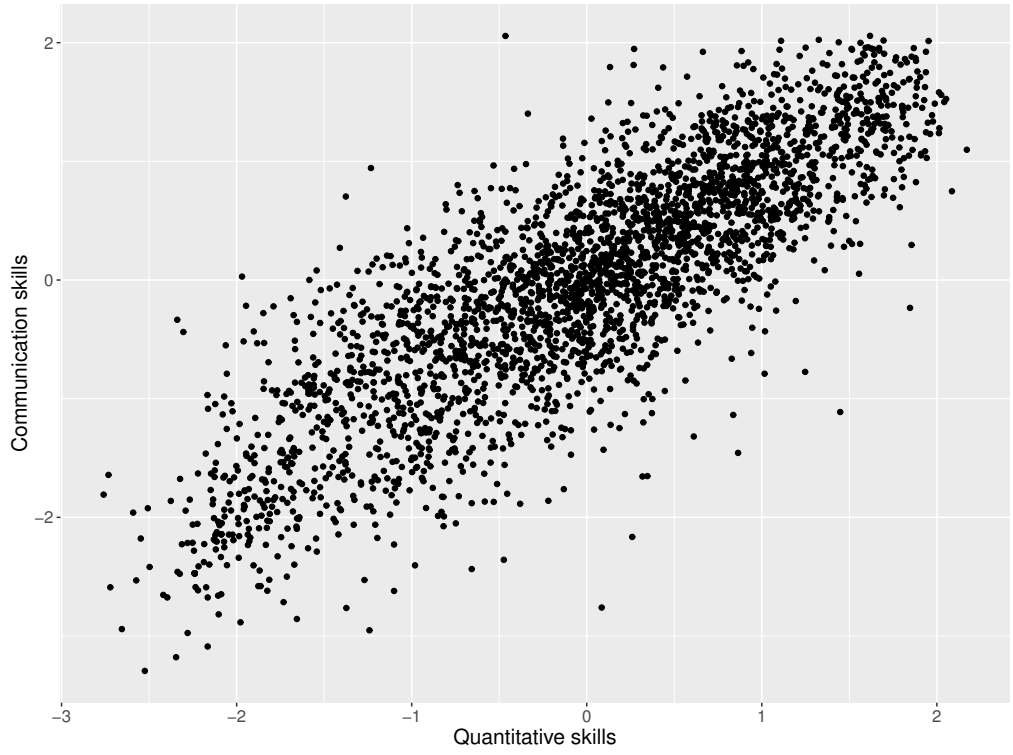
- English admin data, links information from three sources
  - ⋆ National Pupil Database (NPD): school id, test scores at ages 11 (SATs), 16 (GCSEs) and 18 (A-levels), gender, cohort, ethnicity, place of residence, local index of deprivation (IDACI)
  - ⋆ Higher Education Statistics Agency (HESA): University and field of study for those who enrol in HE
  - ⋆ HMRC tax records: working status and annual earnings
- Study cohorts born in 1988-91, entering university in 2006-09
- Over 2.5 million children in total, of whom about 35% enrol in 3 year University programmes
- Follow their lives up to age 28
- Use earlier cohorts to obtain age-profiles of earnings and employment beyond that age (Britton, Dearden and Waltmann, 2020)

Skills of students

- Two dimension of skills: communication and quantitative
- Use factor analysis to summarise test scores at 11 and 16
- Exclusion restrictions: Maths and English scores at 16

Skills of students

- Two dimension of skills: communication and quantitative
- Use factor analysis to summarise test scores at 11 and 16
- Exclusion restrictions: Maths and English scores at 16



Communication and quantitative skills (corr=0.9)

Heterogeneity in HE programmes

- By subject

  STEM: Science, Engineering, Maths and Medicine

  LEM: Law, Economics and Management

  AHSS: Arts, Humanities and Social Sciences

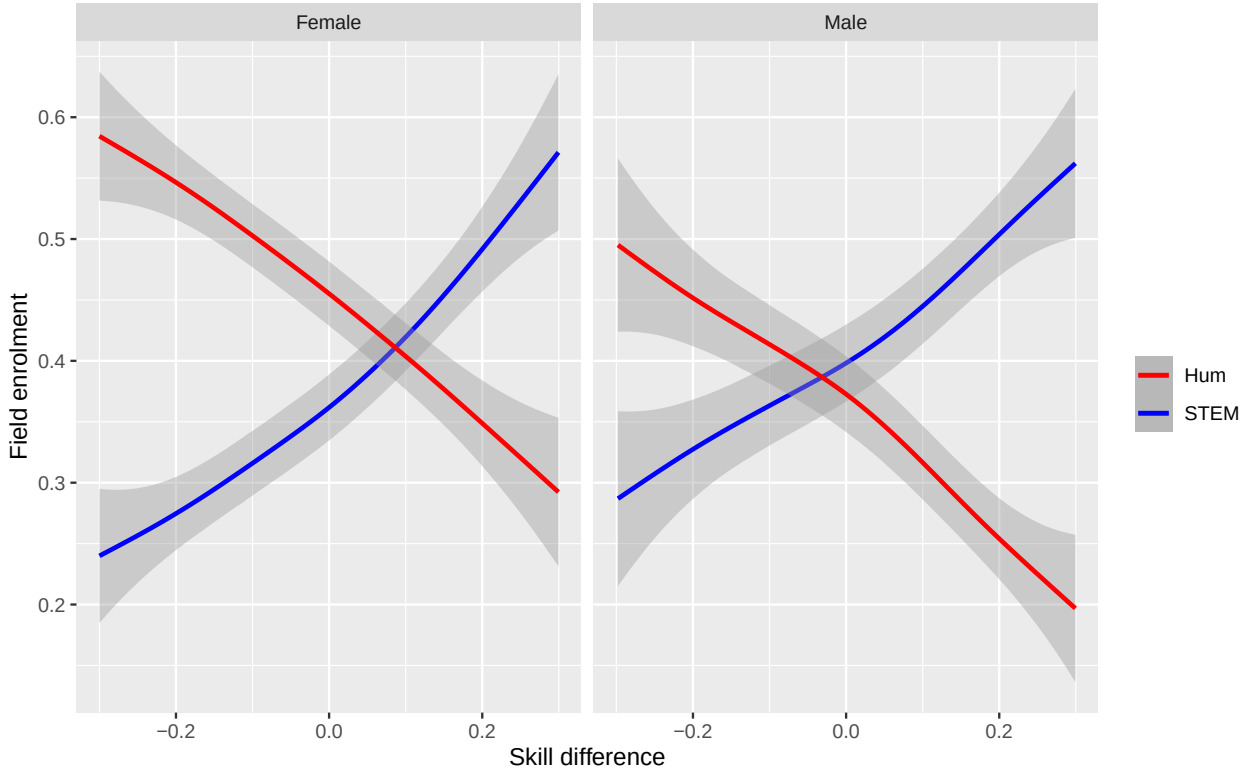
- By University: all 150 institutes in England

- Summary measure of **programme quality**

  - ⋆ Scrapped data from *Complete University Guide* – league tables for all UK programmes
  - ⋆ First principal component from 5 quality indicators: spending in academic services and facilities, research quality, student/staff ratio, index of student satisfaction

High Education Sorting

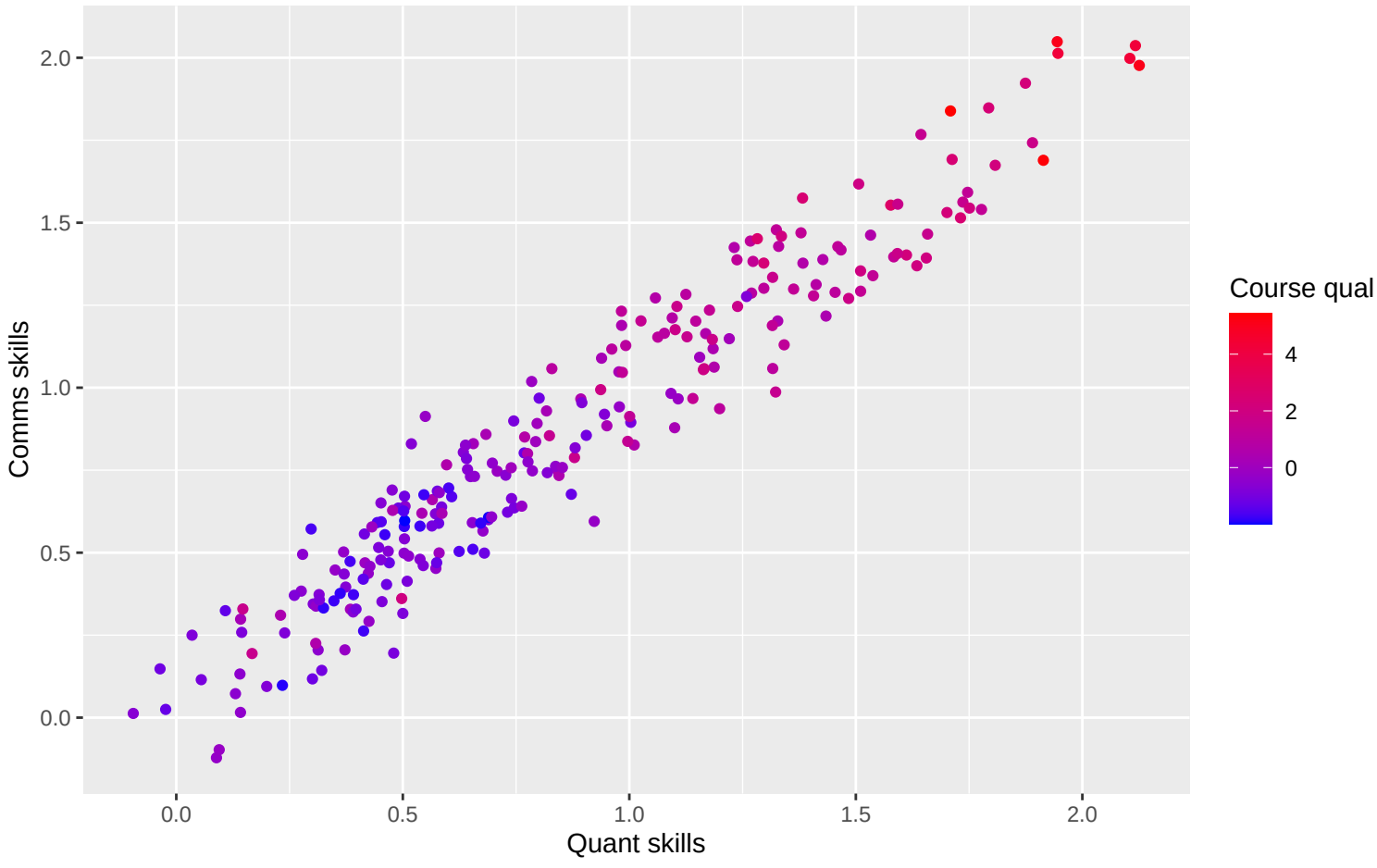
Relative skills matter for subject choice



Relative skills and subject sorting

High Education Sorting

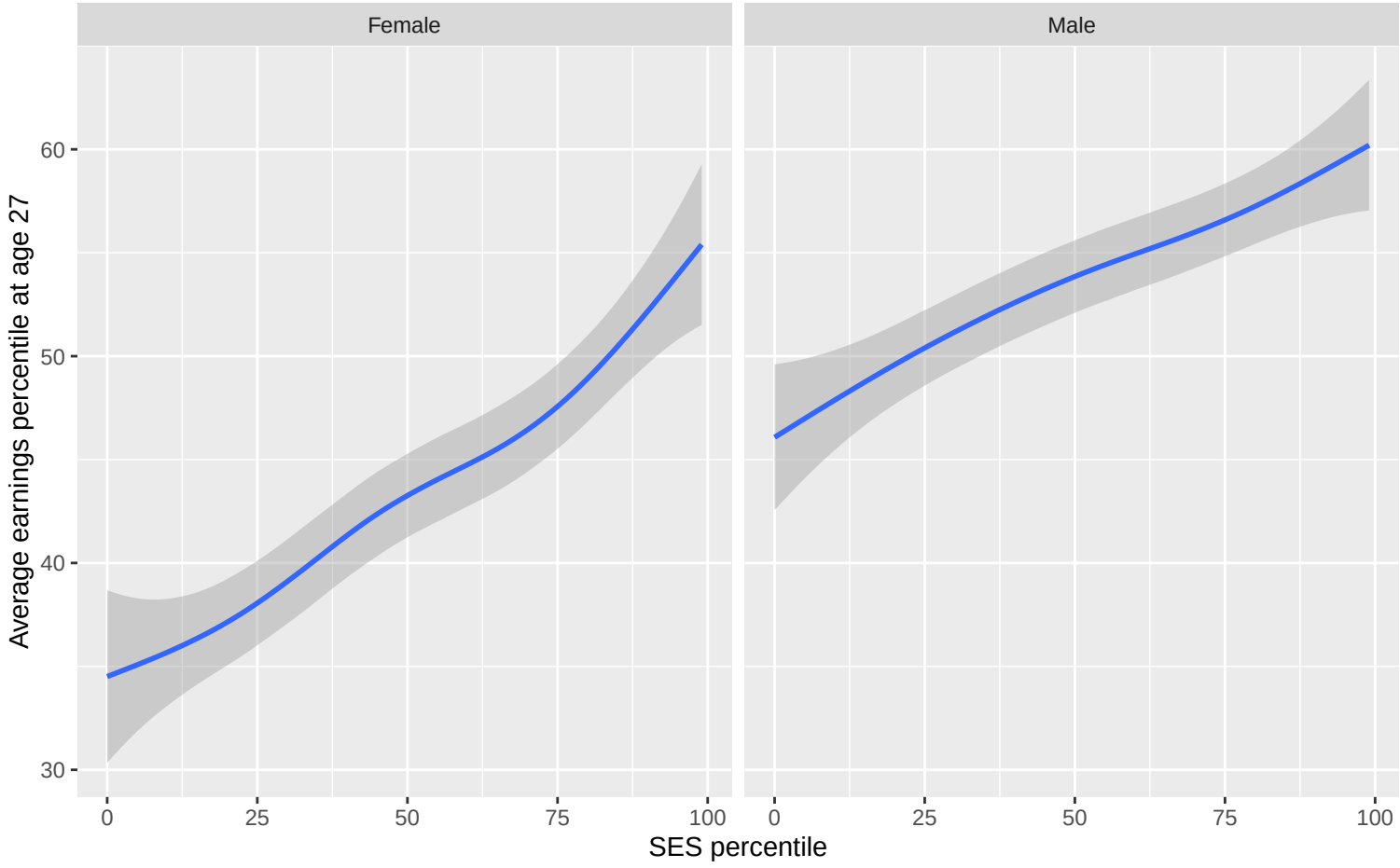
While the quality index captures programme selectivity



Average skill composition by course

Socio-economic gradient

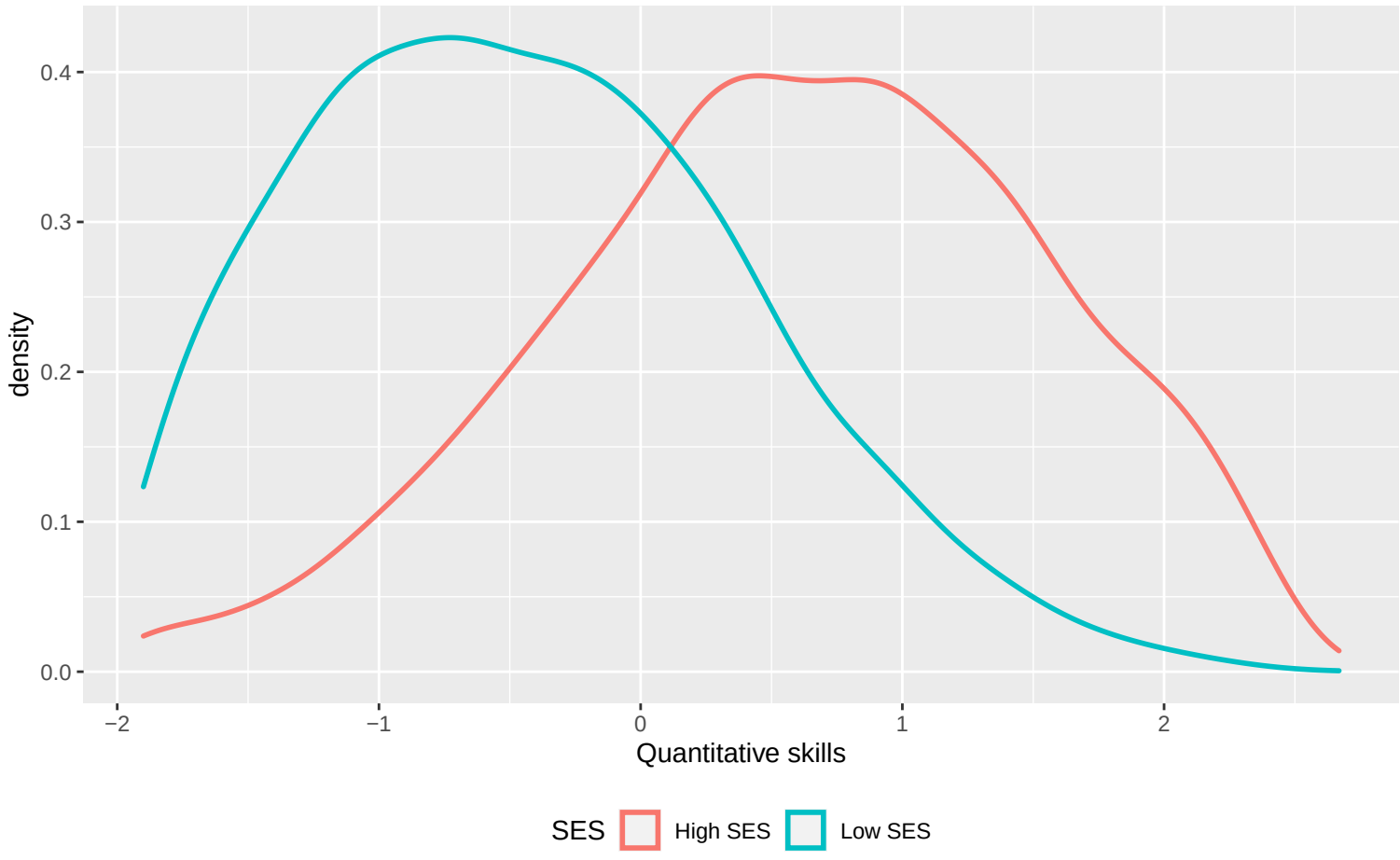
Background disadvantage is strongly reflected on later earnings



Parental SES and child's income

Socio-economic gradient

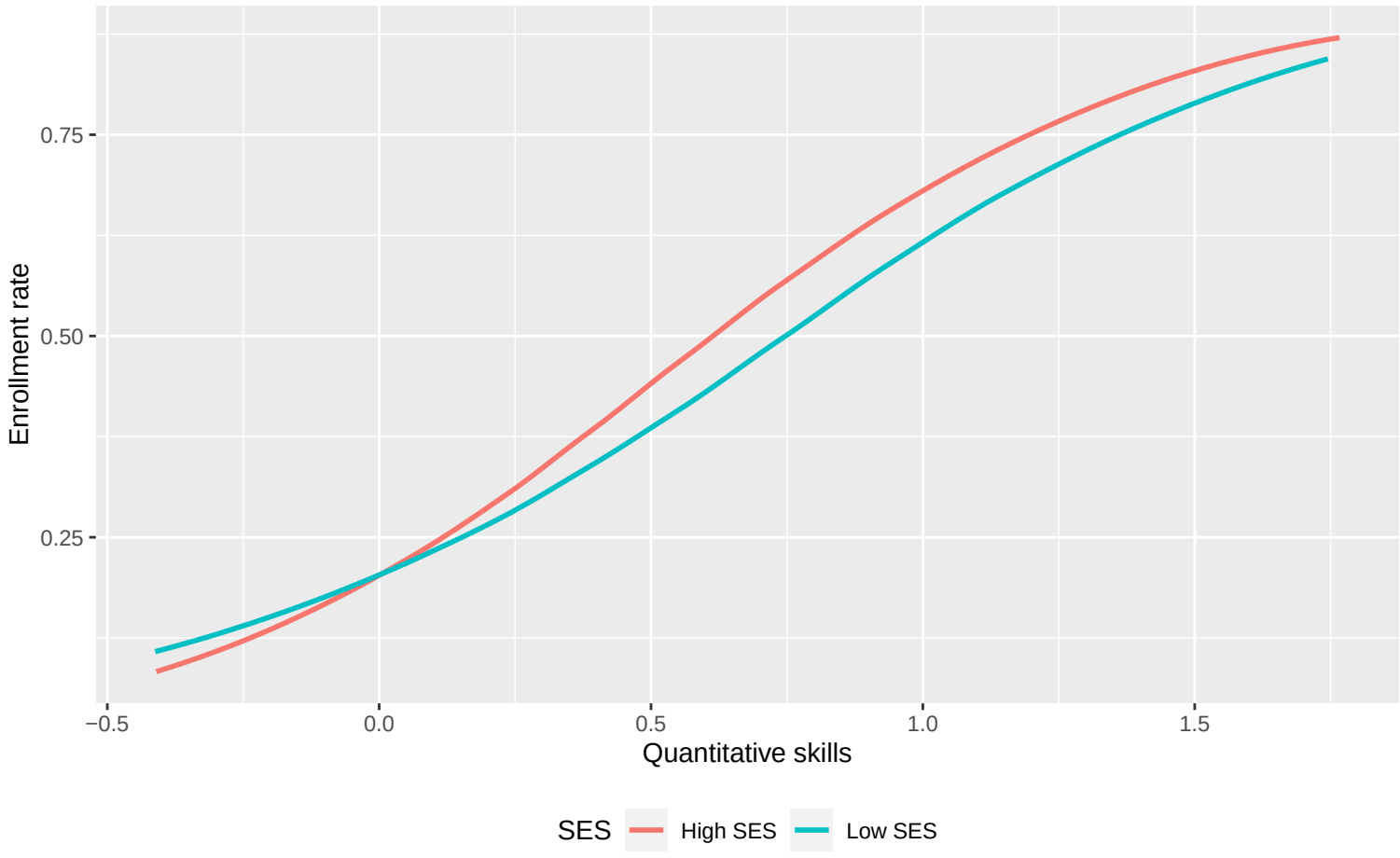
... but also on skills



Distribution of quantitative skills by parental SES

Socio-economic gradient

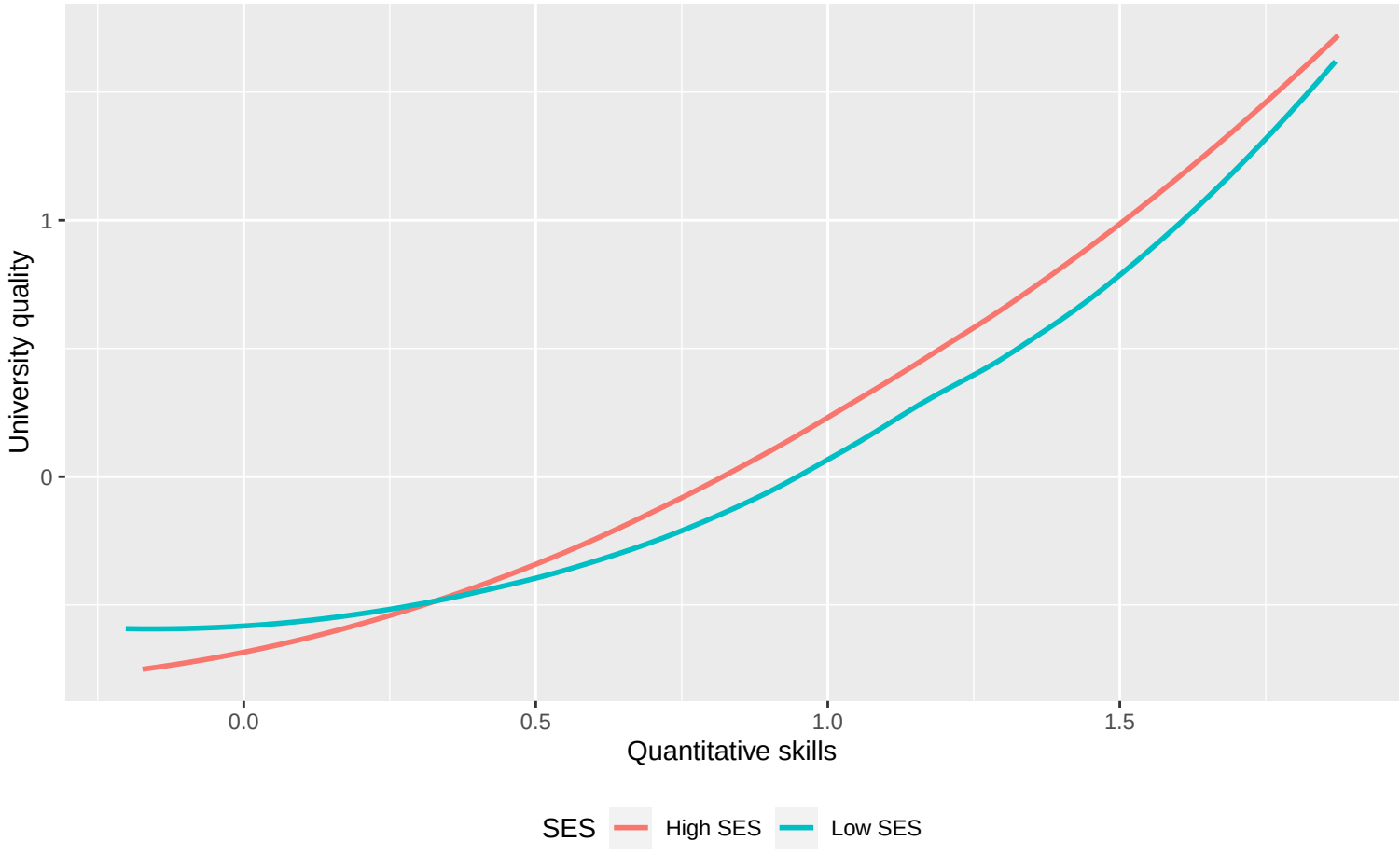
... participation in HE



Enrollment rates by quantitative skills and parental SES

Socio-economic gradient

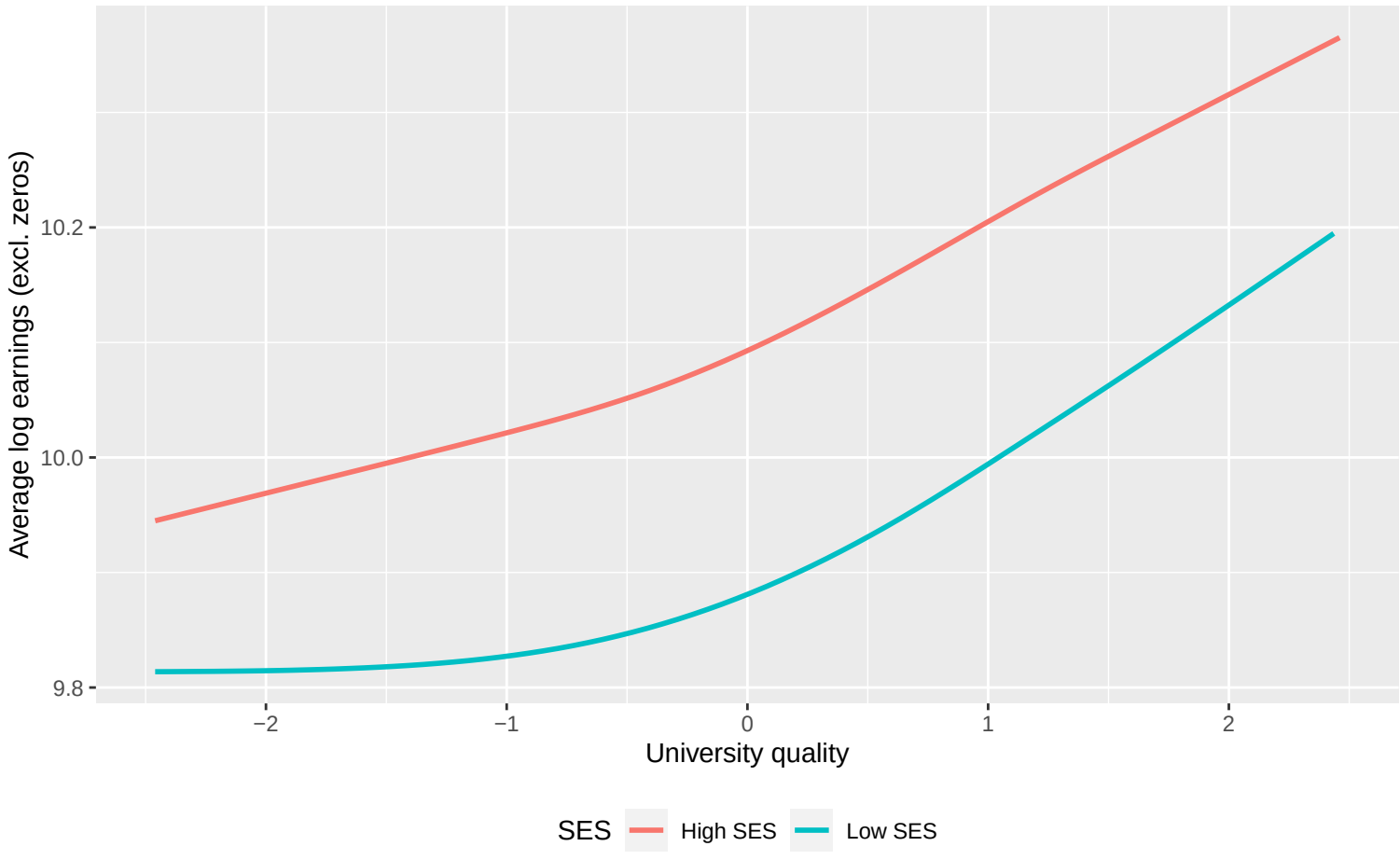
... the selectivity of degree



Average programme quality among participants, by quantitative skills by parental SES

Socio-economic gradient

... and earnings conditional on selectivity



Average earnings by programme quality and parental SES

Model

main features

Life in two stages. Stage 1: education

- ⋆ Heterogeneous students decide whether to participate in HE
- ⋆ Participants and programmes meet in many-to-one frictionless matching market
- ⋆ Tuition fees set exogenously and uniform across programmes
- ⋆ Preferences and capacity constraints determine equilibrium sorting
- ⋆ Non-transferable utility

Stage 2: work

- ⋆ Simplest life-cycle model of labour supply, consumption and savings
- ⋆ Earnings depend on education attainment and the characteristics of both students and programmes

Model

HE programmes rank applicants

- Programmes $j \in J$ offer education valued in work
  - ⋆ by subject F_j: LEM, STEM and AHSS
  - ⋆ with quality Q_j
- Care about reputation, reflected in the skill composition of their students:
  - ⋆ observed mathematical and communication skills (S^m, S^c)
  - ⋆ unobserved (to the econometrician) productive ability θ
- Have strict preference for filling available spots
- Subject-specific preferences for student i

$$W_{ji} = \gamma^m_{F_j} S^m_i + \gamma^c_{F_j} S^c_i + \gamma^\theta_{F_j} \theta_i + \epsilon_{F_j,i}$$

ϵ: unobserved preferences – skills that do not matter for earnings or errors on the part of programmes in assessing student's skills

Model

Students rank programmes

- Students $i \in I$ characterised by
  - ⋆ Skills (S_i^m, S_i^c, θ_i)
  - ⋆ Traits X_i: socio-economic background $(\text{SES}_i, \text{private}_i)$, gender g_i
- Draw value from attending HE and its returns
  - ⋆ Varies with nature of the programme (F, Q)
  - ⋆ ... distance from home – distance_{ij}
  - ⋆ ... choices of past cohorts in secondary school – share_{ij}
- Value for student i of joining programme j

$$U_{ij} = u(F_j, Q_j, \text{distance}_{ij}, \text{share}_{ij}, X_i) + \eta_{ij} + \text{E}V_{ij}$$

where η are unobserved preferences for programme j, split in two components: preferences for subject and university

Model

Earnings

Pre-tax earnings t years out of education with qualification j in field f

$$\begin{aligned}
\ln\left(y_{it}^{e}\right) &= \alpha_{0f} + \alpha_{1f} X_i + \alpha_{2f} S_i^m + \alpha_{3f} S_i^c + \alpha_{4f} Q_j + \alpha_{5f} Q_j^2 \\
&\quad + \alpha_{6f} Q_j S_i^m + \alpha_{7f} Q_j S_i^c + \alpha_{8f} \ln(t+1) + \alpha_{\theta f} \theta_i + \epsilon_{fit} \\
\epsilon_{fit} &= \rho_f \epsilon_{fit-1} + \xi_{fit}
\end{aligned}$$

Model

EV is solution to simple life-cycle problem

- University students use their loans to consume and pay fees
- Working life starts at 22, or 19 for non-graduates
- During the working life, individuals consume, supply labour and save:

$$
\begin{aligned}
V_{it}^{j} \quad = \quad & \max_{c_{it}, d_{it}, a_{it+1}} \left[\ln(c_{it}) + g\left(t, g_i, f, \lambda_{it}\right) d_{it} + \beta \mathrm{E} V_{it+1}^{j} \right] \\
\text{st} \quad & a_{it+1} \quad = \quad R a_{it} + d_{it} y_{it} - P\left(d_{it} y_{it}, l_{it}\right) - T\left(d_{it} y_{it} - P\right) - c_{it} \\
& l_{it+1} \quad = \quad R_l l_{it} - P\left(d_{it} y_{it}, l_{it}\right) \\
& l_{it} \quad \geq \quad 0
\end{aligned}
$$

Match equilibrium

- Centralised market with non-transferable utility, a la Gale-Shapley
- No prices in this market: programme selectivity rule (or capacity constraints) act as prices to clear market (Azevedo and Leshno 2016)
- A stable match exists (Gale-Shapley, Abdulkadiroglu et al. 2015, Azevedo and Leshno 2016)
- We assume stability, which means

$$U_{ij} > U_{i,\mu(i)} \implies \min_{i' \in \mu^{-1}(j)} \{W_{ji'}\} > W_{ji}$$

where

  - $\mu(i) : I \to J$ is the programme attended by student i
  - $\mu^{-1}(j)$ is the set of students matched to course j

Estimation

We use a simulated minimum distance estimator to estimate the parameters of the model:

$$\hat{\Theta} = \min_{\Theta} (M - M(\Theta))' \mathbf{W} (M - M(\Theta))$$

Where Θ includes

1. Student utility parameters
2. Earnings returns parameters
3. Course utility parameters

Two parameters are fixed outside the model: discount rate $\beta = 0.95$ and the t coefficient in the earnings equation

Identification

Excluded variables from earnings and university preferences

- ⋆ Distance to university, choices of past cohorts in school
- ⋆ Varies demand for programmes without changing choice sets
- ⋆ Varies education attainment without affecting potential earnings

Sorting patterns

- ⋆ Reveal whether two course types are equally desirable (Diamond 2017)
- ⋆ Include moments on covariance of student & course chars

Many-to-one matching

- ⋆ If within variation $<<$ between variation $\rightarrow$ universities have strong preference for characteristic
- ⋆ Include moments on within vs between variation in student chars and correlations between one's characteristics and those of their peers

Parameter estimates: earnings equation

	No Univ	STEM	LEM	AHSS
intercept	9.898	10.064	10.104	9.988
female	-0.310	-0.148	-0.186	-0.132
SES	0.324	0.205	0.213	0.231
private	0.051	-0.022	-0.003	0.002
math skills	0.101	0.089	0.125	0.109
communication skills	0.051	0.046	0.030	0.007
programme quality		0.021	0.026	0.017
programme quality squared		-0.003	0.002	-0.002
quality $\times$ math		0.031	0.008	0.020
quality $\times$ communication		-0.002	0.009	0.009
unobserved skill	0.044	0.044	0.044	0.044

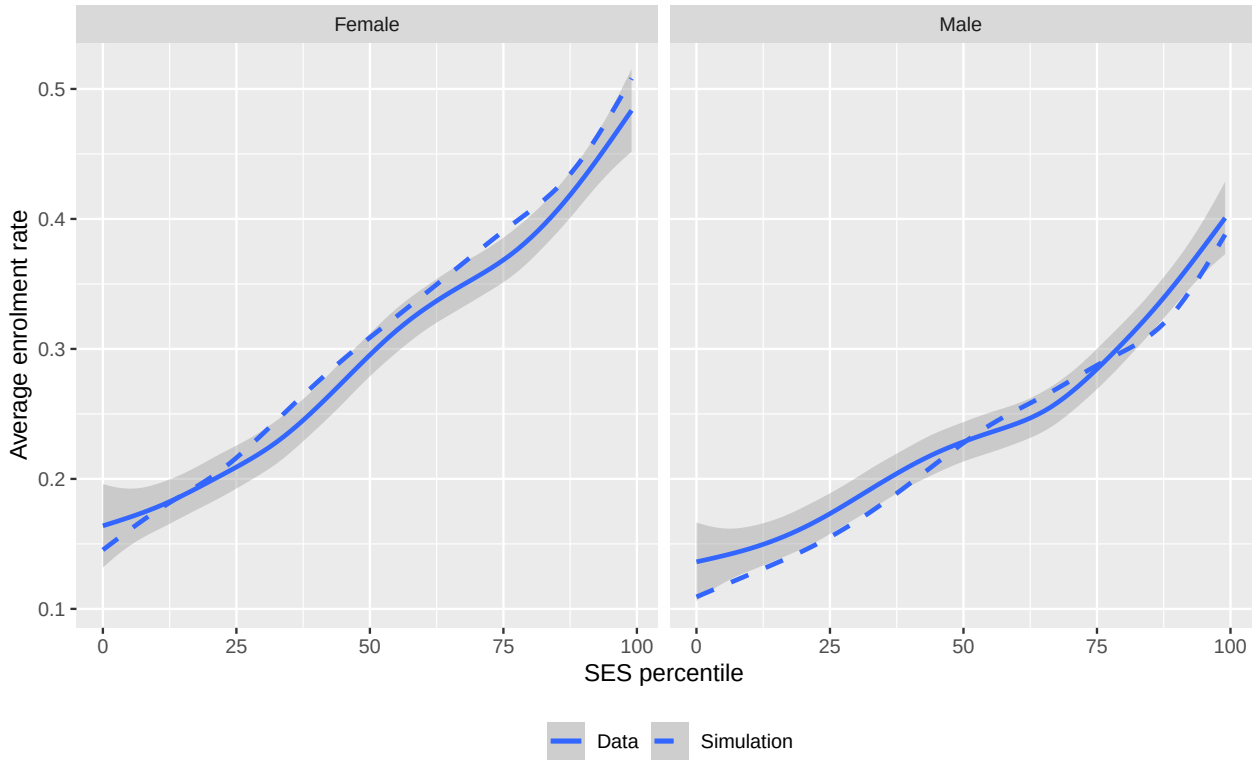
Parameter estimates: student preferences

	STEM	LEM	AHSS
intercept	0.107	-0.756	1.825
female	-0.042	-0.763	-0.088
SES		0.792	
private school		0.369	
distance		-0.503	
distance×SES		0.752	
distance×private		0.304	
school share	0.180	0.218	
Variance: field preference		1.007	
Variance: university preference		0.571	

Parameter estimates: programme preferences

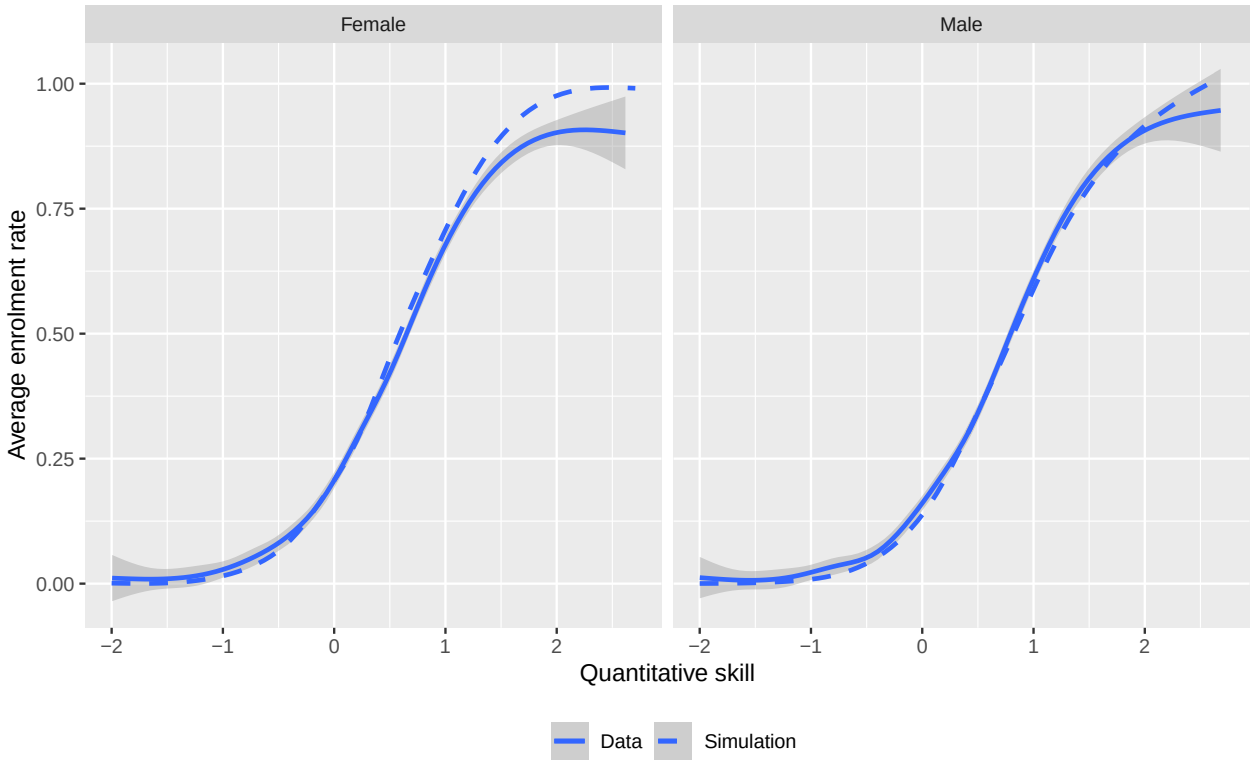
	STEM	LEM	AHSS
math skill	1.598	0.599	0.107
communication skill	0.057	0.814	1.462
unobserved skill		0.050	

Model fit



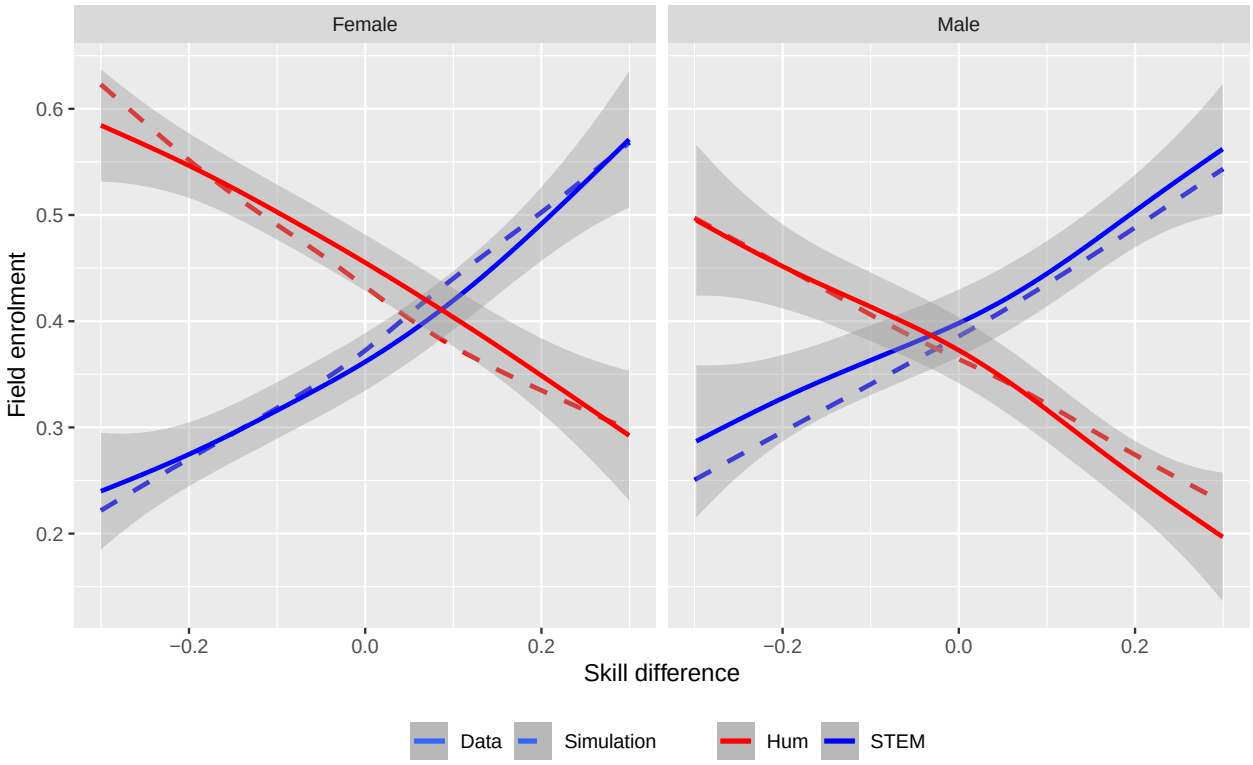
Enrolment by SES and by gender

Model fit



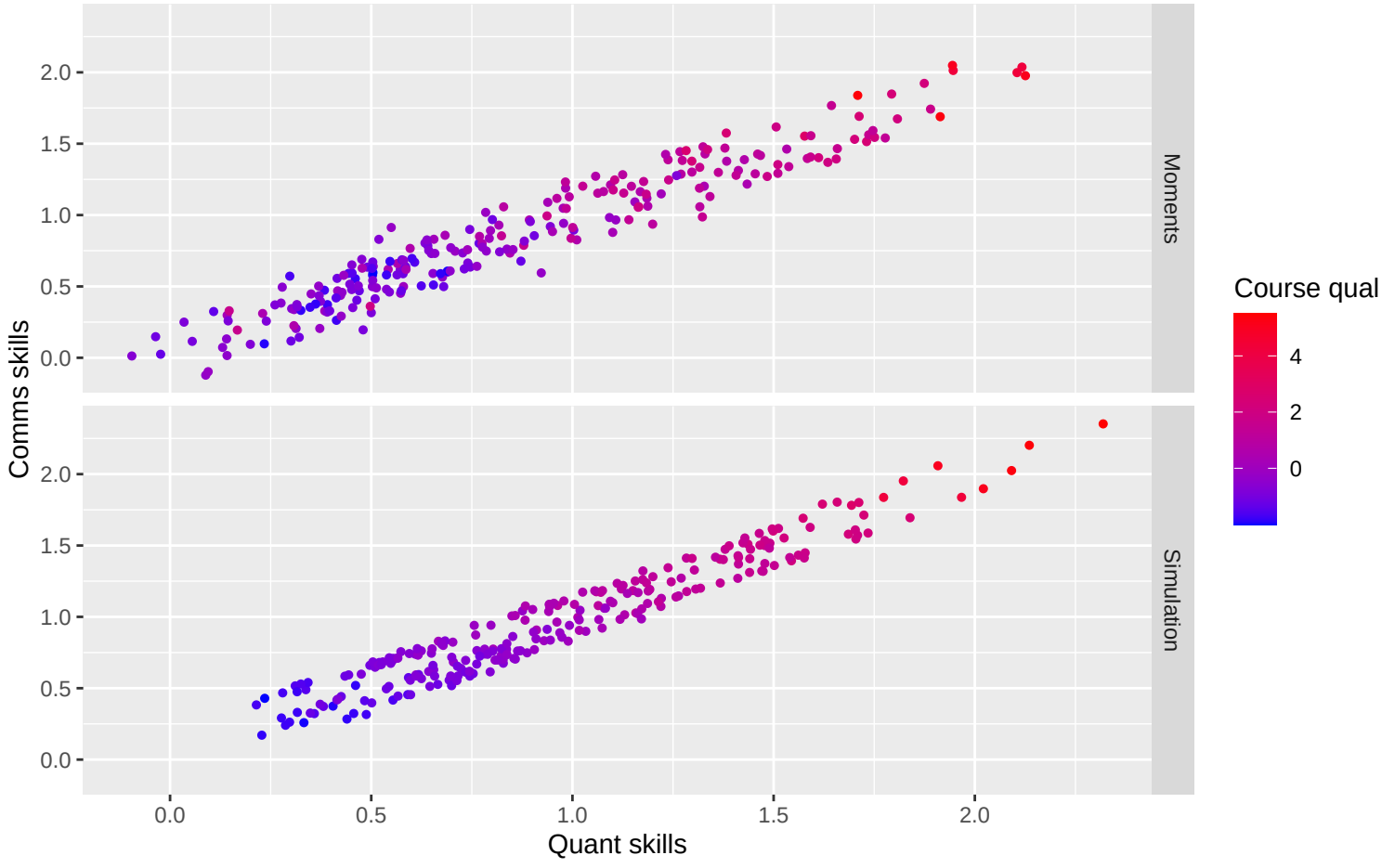
Enrolment by quantitative skills, by gender

Model fit



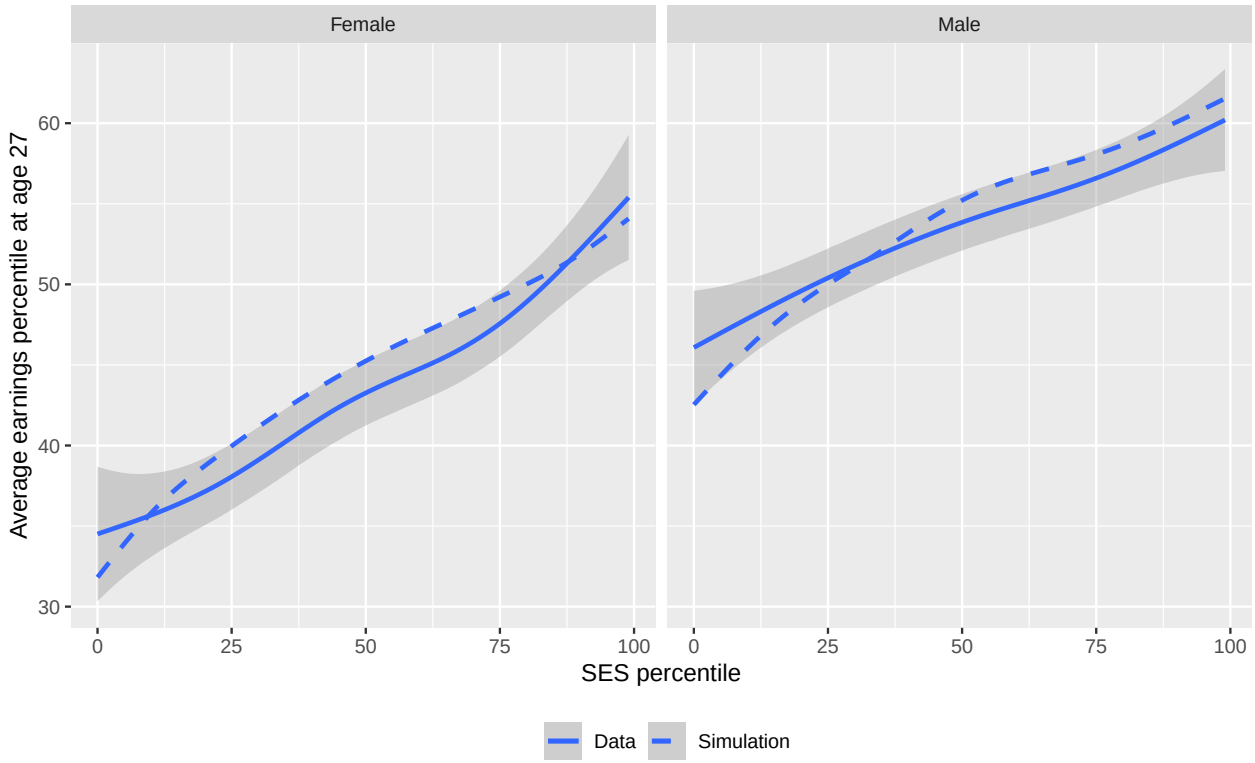
Relative skills and subject choice, by gender

Model fit



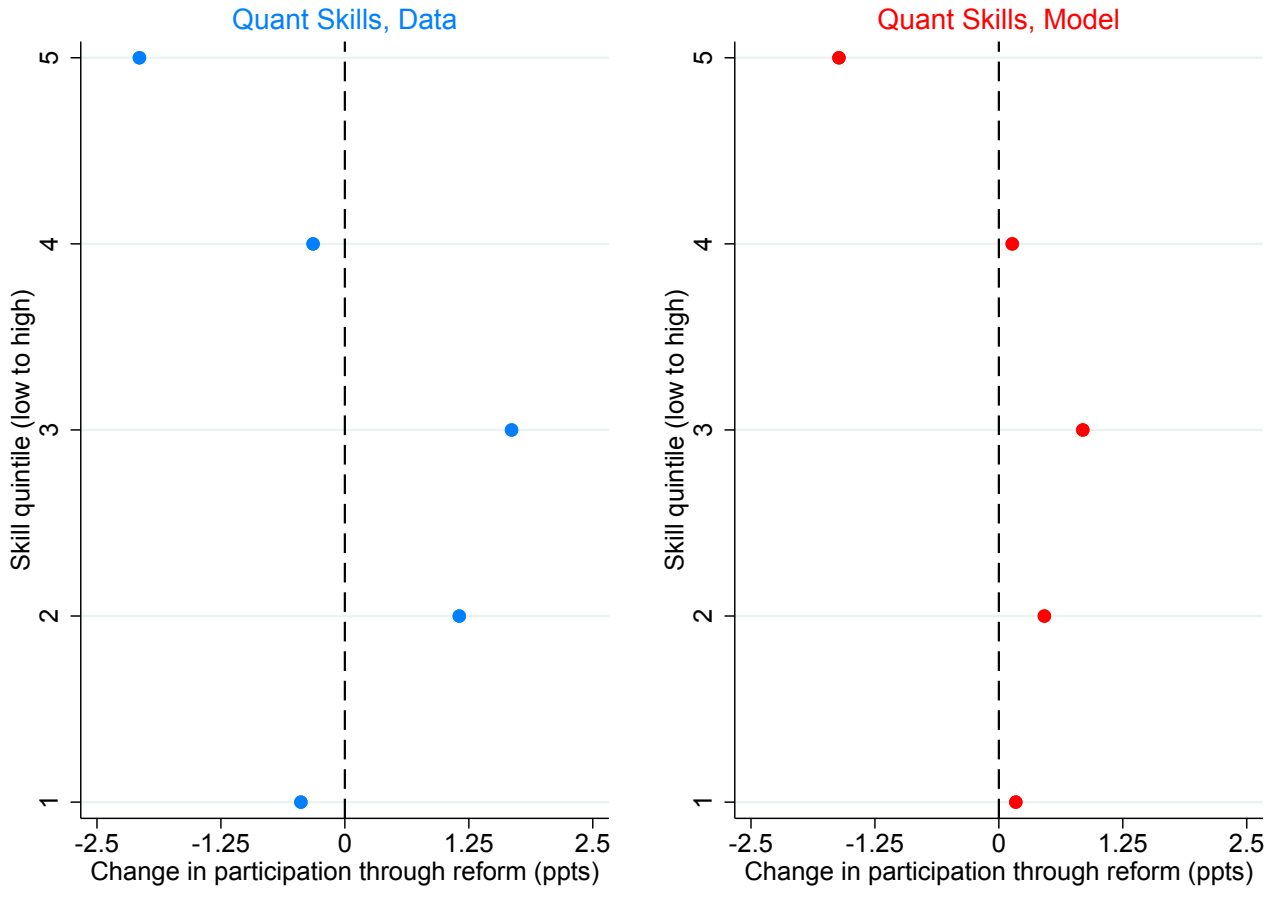
Skill composition by quality of programme

Model fit



Social mobility, by gender

Model fit: out of sample



Enrolment effects of 2012 reforms

Counterfactual policy reforms

- **Policy 1:** Additional maintenance grants of £4,000 per year, on top of existing loans, for low SES (bottom 50%) students
- **Policy 2:** As above, but only if studying STEM
- **Policy 3:** No student loans for low SES students - tuition fees set to £0 and maintenance loans converted to grants (but held at current levels)
- **Policy 4:** A quota system: programmes to reserve 50% of their places for low SES students.

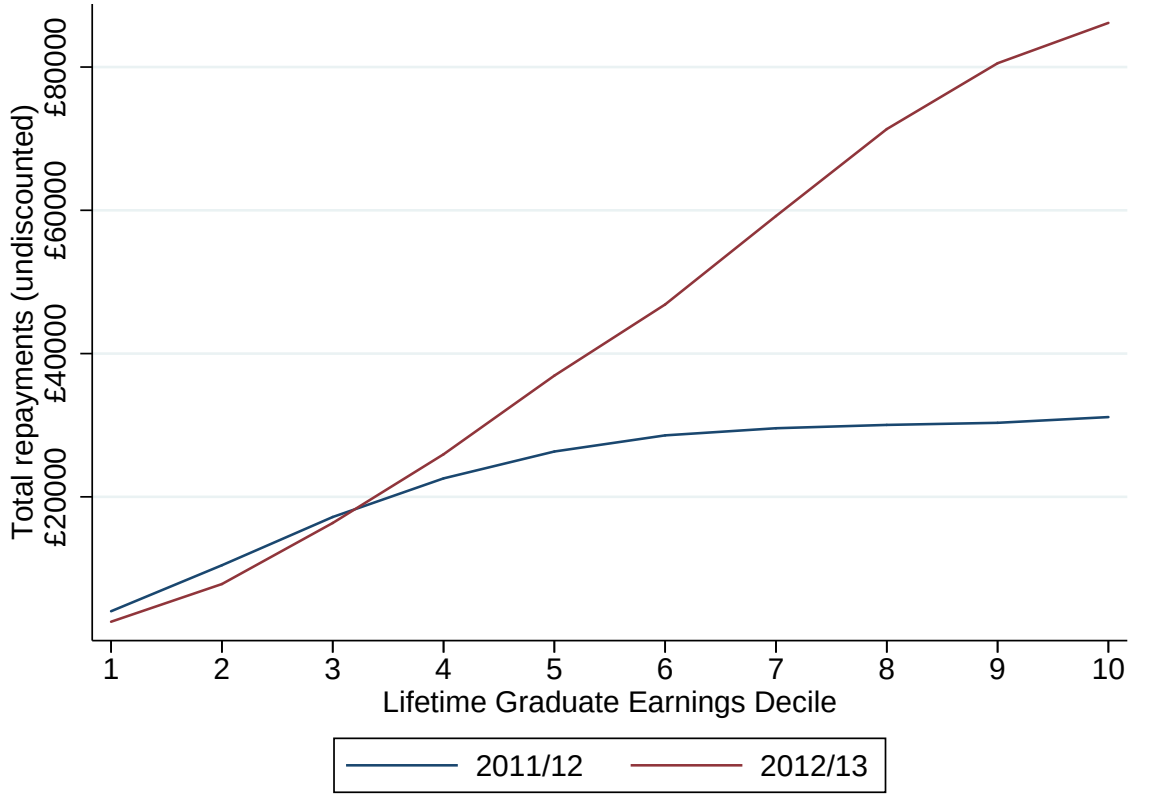
Counterfactual policy reforms

	Baseline	Policy 1 (Grants)	Policy 2 (STEM Grants)	Policy 3 (Loan write-off)	Policy 4 (Quota)
			Low SES		
HE participation	0.198	0.202	0.198	0.199	0.279
STEM share	0.445	0.443	0.551	0.454	0.441
Undermatch	1.46	1.59	1.66	1.49	0.66
Δ Av. Earnings (%)	-	0.24	1.01	0.08	4.21
			High SES		
HE participation	0.478	0.471	0.475	0.476	0.391
STEM share	0.417	0.415	0.347	0.413	0.421
Undermatch	1.24	1.22	1.18	1.24	1.92
Δ Av. Earnings (%)	-	-0.32	-0.90	-0.25	-5.21

Conclusion

- We develp a model of the HE market and life cycle labour supply and savings
- Our model is able to capture sorting patterns in the data as well as the impacts of the 2021 reform
- We find that SES remains a strong determinant of HE sorting and earnings even as we account for the skills of students and the characteristics of programmes
- We use the model to better understand impact of 2012 reforms
- And to simulate a big reform to fees/maintenance
- We find only muted impacts to supply side reforms
- Aggressive demand-side reforms can make a difference

How did the 2012 reform affect enrolment?



Lifetime student loan repayments

Cost increased the most among the highest paid

back