

Who goes to university and why?

A comparison across cohorts (1985 – 2015)

Oliver Cassagneau-Francis

22nd October 2020

Sciences Po

Introduction

Model

Data

Identification and estimation

Results

What's next and conclusion

Why care about the decision to attend HE?

Lots of evidence of benefits of a university degree

- Increased earnings
- Better health
- Lower levels of crime

Why care about the decision to attend HE?

Lots of evidence of benefits of a university degree

Often taxpayers are paying for these benefits

- £17 billion upfront costs of higher education in the UK in 2017
(Dearden et al., 2017)
- OECD countries' public spending on HE \approx 1% of GDP (OECD, 2020)

Why care about the decision to attend HE?

Lots of evidence of benefits of a university degree

Often taxpayers are paying for these benefits

Persistent gap in HE attainment by socio-economic status

- and hence in beneficiaries of this public spending
- England: children of parents' earning in top 20% **twice as likely to attend university** as children of parents in the bottom 20%
- gap in attainment by parental education up to 30pp (OECD, 2018)
- is HE a barrier to social mobility?

Why care about the decision to attend HE?

Lots of evidence of benefits of a university degree

Though often taxpayers are paying for these benefits

Persistent gap in HE attainment by socio-economic status

Understanding the factors that influence educational attainment is key, not only for educational outcomes but also for wider issues such as inequality and beneficiaries of public funding

Research questions

**How important are earnings expectations for 16–18 year olds
when deciding whether to go to university?**

Research questions

How important are earnings expectations for 16–18 year olds when deciding whether to go to university?

- How much do earnings expectations vary across students?

Research questions

How important are earnings expectations for 16–18 year olds when deciding whether to go to university?

- How much do earnings expectations vary across students?
- How do they compare to other factors in the decision?
i.e. Heckman et al. (2006)'s "psychic costs"

Research questions

How important are earnings expectations for 16–18 year olds when deciding whether to go to university?

What's driving the SES-gap in education attainment?

Research questions

How important are earnings expectations for 16–18 year olds when deciding whether to go to university?

What's driving the SES-gap in education attainment?

- Is it differences in earnings expectations?

Research questions

How important are earnings expectations for 16–18 year olds when deciding whether to go to university?

What's driving the SES-gap in education attainment?

- Is it differences in earnings expectations?
- Or are other factors more important?

Research questions

How important are earnings expectations for 16–18 year olds when deciding whether to go to university?

What is driving the SES-gap in education attainment?

How have these factors changed between the 1980s and today?

Research questions

How important are earnings expectations for 16–18 year olds when deciding whether to go to university?

What is driving the SES-gap in education attainment?

How have these factors changed between the 1980s and today?

- Have students' expectations about the graduate-wage premium changed?

Research questions

How important are earnings expectations for 16–18 year olds when deciding whether to go to university?

What is driving the SES-gap in education attainment?

How have these factors changed between the 1980s and today?

- Have students' expectations about the graduate-wage premium changed?
- What about the relative importance of other factors?

This paper

- Model the decision to go to university

This paper

- Model the decision to go to university
- **Estimate model on data for cohort born in 1989/90**

This paper

- Model the decision to go to university
- **Estimate model on data for cohort born in 1989/90**
 - Schooling, background, and subjective data from before decision
 - Earnings and occupation data after entry to labour market
 - Exploit subjective data on students' attitudes and lifestyles at 16 to directly estimate "psychic costs"

This paper

- Model the decision to go to university
- Estimate model on data for cohort born in 1989/90
- Model structure → estimate **distributions of factors**

This paper

- Model the decision to go to university
- Estimate model on data for cohort born in 1989/90
- Model structure → estimate **distributions of factors**
 - Compare earnings vs other factors (“psychic costs”)
 - → importance of earnings vs other factors in decision

This paper

- Model the decision to go to university
 - Estimate model on data for cohort born in 1989/90
 - Model structure → estimate **distributions of factors**
-
- Compare across different SES groups
 - → what's driving the SES-gap?

This paper

- Model the decision to go to university
- Estimate model on data for cohort born in 1989/90
- Model structure → estimate distributions of factors
- **Re-estimate model on cohort born in 1970**

This paper

- Model the decision to go to university
- Estimate model on data for cohort born in 1989/90
- Model structure → estimate distributions of factors
- **Re-estimate model on cohort born in 1970**
 - Compare across cohorts
 - → have the factors in the HE decision changed? how?

This paper

- Model the decision to go to university
- Estimate model on data for cohort born in 1989/90
- Model structure → estimate distributions of factors
- Re-estimate model on cohort born in 1970
- **Use subjective data to decompose “psychic costs”**

This paper

- Model the decision to go to university
- Estimate model on data for cohort born in 1989/90
- Model structure → estimate distributions of factors
- Re-estimate model on cohort born in 1970
- **Use subjective data to decompose “psychic costs”**
 - data on attitudes for different aspects of life
 - use this to estimate **components** of psychic costs

Preview of results

- non-earnings factors $4\times$ more important than earnings expectations

Preview of results

- non-earnings factors $4\times$ more important than earnings expectations
- **non-earnings factors are also driving the SES gap**

Preview of results

- non-earnings factors $4\times$ more important than earnings expectations
- **non-earnings factors are also driving the SES gap**
 - mean earnings expectations (slightly) decreasing in SES
 - other factors strongly increasing in SES

Preview of results

- non-earnings factors $4\times$ more important than earnings expectations
- non-earnings factors are also driving the SES gap
- **increase in HE driven entirely by non-earnings factors (1985–2015)**

Preview of results

- non-earnings factors $4\times$ more important than earnings expectations
- non-earnings factors are also driving the SES gap
- **increase in HE driven entirely by non-earnings factors (1985–2015)**
 - earnings: $\downarrow 15$ p.p.; other factors: $\uparrow 36$ p.p.
 - mean increase in non-earnings factors $\approx 50\%$ wage increase

Related literature

Estimating *ex ante* returns to education

- rational expectations: e.g. Cunha and Heckman (2007b,a); Heckman et al. (2006)
- surveys to elicit expectations: e.g. Zafar (2009); Delavande and Zafar (2017); Wiswall and Zafar (2015); Arcidiacono et al. (2019); Boneva and Rauh (2019); Wiswall and Zafar (2016)

UK higher education expansion

- 'educational inequality': Blanden and Machin (2004)
- technological change: Blundell et al. (2018)
- wage premium: Walker and Zhu (2008); Green et al. (2016)

Determinants of educational choices

- Cameron and Heckman (1993)

Estimating *ex post* returns to education

- Bhuller et al. (2014)

Roy models with unobserved heterogeneity

- 'extended': D'Haultfoeuille and Maurel (2013)
- 'generalized': Abbring and Heckman (2007); Heckman et al. (2006); Heckman and Navarro (2007); Heckman and Vytlacil (2007)

Expectations (need to review)

- Mackowiak et al. (2018); Fuster et al. (2019)

Introduction

Model

Data

Identification and estimation

Results

What's next and conclusion

Factors in the decision

What do students consider when making educational choices?

Factors in the decision

What do students consider when making educational choices?

- (future) earnings

Factors in the decision

What do students consider when making educational choices?

- (future) earnings
- effort required to *get* a place at university

Factors in the decision

What do students consider when making educational choices?

- (future) earnings
- effort required to *get* a place at university
- life at university:
 - + social life
 - + meet friends / partner
 - ± studying
 - ± leave home / local area
 - not earning / gaining experience
 - stress

Factors in the decision

What do students consider when making educational choices?

- (future) earnings
- effort required to *get* a place at university
- life at university:
 - + social life
 - + meet friends / partner
 - ± studying
 - ± leave home / local area
 - not earning / gaining experience
 - stress
- life after university:
 - + better job / career
 - ± graduate “identity”
 - debt

Utility of attending university or working

Different factors enter utility function *additively*

$$U_{Si} = \alpha Y_{Si}^{ea} + \theta'_{Si}\gamma + \epsilon_{Si} \quad (1)$$

Utility of attending university or working

Different factors enter utility function *additively*

$$U_{Si} = \alpha Y_{Si}^{ea} + \theta'_{Si}\gamma + \epsilon_{Si} \quad (1)$$

- Y_{Si}^{ea} is earnings expectations

Utility of attending university or working

Different factors enter utility function *additively*

$$U_{Si} = \alpha Y_{Si}^{ea} + \theta'_{Si}\gamma + \epsilon_{Si} \quad (1)$$

- Y_S^{ea} is earnings expectations
- θ_S is a vector of expectations about other aspects of life

Utility of attending university or working

Different factors enter utility function *additively*

$$U_{Si} = \alpha Y_{Si}^{ea} + \theta'_{Si}\gamma + \epsilon_{Si} \quad (1)$$

- Y_S^{ea} is earnings expectations
- θ_S is a vector of expectations about other aspects of life
- ϵ_S is a random utility term

Utility of attending university or working

Different factors enter utility function *additively*

$$U_{Si} = \alpha Y_{Si}^{ea} + \theta'_{Si}\gamma + \epsilon_{Si} \quad (1)$$

- Y_S^{ea} is earnings expectations
- θ_S is a vector of expectations about other aspects of life
- ϵ_S is a random utility term
- terms differ for each individual conditional on choice S

A model of educational choice

- Students then compare (expected) utility in each of the two states

$$S_i \equiv \mathbf{1}\{U_{1i} > U_{0i}\}$$

A model of educational choice

- Students then compare (expected) utility in each of the two states

$$S_i \equiv 1\{U_{1i} > U_{0i}\}$$

- This can be written as a difference between **outcomes** and **“costs” à la Roy:**

$$S_i \equiv 1\{\underbrace{\alpha(Y_1^{ea} - Y_0^{ea})}_{\text{outcomes}} - \underbrace{(\theta_{1i} - \theta_{0i})'(-\gamma)}_{\text{“costs”}} + \epsilon_{1i} - \epsilon_{0i} > 0\} \quad (2)$$

A model of educational choice

- Students then compare (expected) utility in each of the two states

$$S_i \equiv \mathbb{1}\{U_{1i} > U_{0i}\}$$

- This can be written as a difference between **outcomes** and **“costs” à la Roy:**

$$S_i \equiv \mathbb{1}\{\underbrace{\alpha(Y_1^{ea} - Y_0^{ea})}_{\text{outcomes}} - \underbrace{(\theta_{1i} - \theta_{0i})'(-\gamma)}_{\text{“costs”}} + \epsilon_{1i} - \epsilon_{0i} > 0\} \quad (2)$$

- α and γ are identified given Y_s^{ea} and θ_{Si} and choices from

$$\Pr(S = 1 | Y_s^{ea}, \theta_{Si}) = \Pr(\alpha(Y_1^{ea} - Y_0^{ea}) + (\theta_{1i} - \theta_{0i})'\gamma > \epsilon_{0i} - \epsilon_{1i})$$

plus some assumptions about errors.

Introduction

Model

Data

Identification and estimation

Results

What's next and conclusion

Main data sources

Two UK longitudinal cohort studies run by CLS at UCL

Main data sources

Two UK longitudinal cohort studies run by CLS at UCL

- Next Steps (LSYPE1) members born in 1989/1990

Main data sources

Two UK longitudinal cohort studies run by CLS at UCL

- Next Steps (LSYPE1) members born in 1989/1990
- British Cohort Study (BCS) members born in 1970

Main data sources

Two UK longitudinal cohort studies run by CLS at UCL

- Next Steps (LSYPE1) members born in 1989/1990
- British Cohort Study (BCS) members born in 1970

Common features:

Main data sources

Two UK longitudinal cohort studies run by CLS at UCL

- Next Steps (LSYPE1) members born in 1989/1990
- British Cohort Study (BCS) members born in 1970

Common features:

- Follow 17,000 people from respective birth cohorts

Main data sources

Two UK longitudinal cohort studies run by CLS at UCL

- Next Steps (LSYPE1) members born in 1989/1990
- British Cohort Study (BCS) members born in 1970

Common features:

- Follow 17,000 people from respective birth cohorts
- Surveys at age 16 and 25
 - (16) Detailed information on schooling, family background
 - (16) Subjective questions about university, life and future
 - (25) Earnings, occupation and qualifications

Main data sources

Two UK longitudinal cohort studies run by CLS at UCL

- Next Steps (LSYPE1) members born in 1989/1990
- British Cohort Study (BCS) members born in 1970

Common features:

- Follow 17,000 people from respective birth cohorts
- Surveys at age 16 and 25
 - (16) Detailed information on schooling, family background
 - (16) Subjective questions about university, life and future
 - (25) Earnings, occupation and qualifications
- Possibility for longer panel as both surveys ongoing

Cohort sample design

Next Steps (England, born 1989/90)

Cohort sample design

Next Steps (England, born 1989/90)

- Sampling “unit” school: 647 of 892 selected took part

Cohort sample design

Next Steps (England, born 1989/90)

- Sampling “unit” school: 647 of 892 selected took part
- Repr. sample oversampled minorities and “deprived” schools

Cohort sample design

Next Steps (England, born 1989/90)

- Sampling “unit” school: 647 of 892 selected took part
- Repr. sample oversampled minorities and “deprived” schools
- 15,770 interviews in sweep 1, from sample of 21,000 (74%)

Cohort sample design

Next Steps (England, born 1989/90)

- Sampling “unit” school: 647 of 892 selected took part
- Repr. sample oversampled minorities and “deprived” schools
- 15,770 interviews in sweep 1, from sample of 21,000 (74%)
- Sweeps annually between 14 and 20, then again at 25

Cohort sample design

Next Steps (England, born 1989/90)

- Sampling “unit” school: 647 of 892 selected took part
- Repr. sample oversampled minorities and “deprived” schools
- 15,770 interviews in sweep 1, from sample of 21,000 (74%)
- Sweeps annually between 14 and 20, then again at 25

British Cohort Study (UK, born 1970)

Cohort sample design

Next Steps (England, born 1989/90)

- Sampling “unit” school: 647 of 892 selected took part
- Repr. sample oversampled minorities and “deprived” schools
- 15,770 interviews in sweep 1, from sample of 21,000 (74%)
- Sweeps annually between 14 and 20, then again at 25

British Cohort Study (UK, born 1970)

- All people born in UK (exc. NI) in one week in April 1970 (16,568)

Cohort sample design

Next Steps (England, born 1989/90)

- Sampling “unit” school: 647 of 892 selected took part
- Repr. sample oversampled minorities and “deprived” schools
- 15,770 interviews in sweep 1, from sample of 21,000 (74%)
- Sweeps annually between 14 and 20, then again at 25

British Cohort Study (UK, born 1970)

- All people born in UK (exc. NI) in one week in April 1970 (16,568)
- Overseas born in same week added in later sweeps

Cohort sample design

Next Steps (England, born 1989/90)

- Sampling “unit” school: 647 of 892 selected took part
- Repr. sample oversampled minorities and “deprived” schools
- 15,770 interviews in sweep 1, from sample of 21,000 (74%)
- Sweeps annually between 14 and 20, then again at 25

British Cohort Study (UK, born 1970)

- All people born in UK (exc. NI) in one week in April 1970 (16,568)
- Overseas born in same week added in later sweeps
- Sweeps at birth, 5, 10, 16, 26 and then every 4 years

Subsample and descriptive statistics: Next Steps

Subsample: all students who were asked subjective questions about university (those with > 5 GCSEs)

	Full sample	Subsample
N	6,628	4,640
Female	0.55	0.57
Degree*	0.58	0.68
<i>Russell group</i> *†	0.26	0.28
Employed*	0.83	0.87
Wage (GBP)*‡	393	424

Notes: * At age 25. † Among degree holders. ‡ Median wage.

Subsample and descriptive statistics: Next Steps

Subsample: all students who were asked subjective questions about university (those with > 5 GCSEs)

	Full sample	Subsample
N	6,628	4,640
Female	0.55	0.57
Degree*	0.58	0.68
<i>Russell group*†</i>	0.26	0.28
Employed*	0.83	0.87
Wage (GBP)*‡	393	424

Notes: * At age 25. † Among degree holders. ‡ Median wage.

Subsample and descriptive statistics: Next Steps

Subsample: all students who were asked subjective questions about university (those with > 5 GCSEs)

	Full sample	Subsample
N	6,628	4,640
Female	0.55	0.57
Degree*	0.58	0.68
<i>Russell group</i> *†	0.26	0.28
Employed*	0.83	0.87
Wage (GBP)*‡	393	424

Notes: * At age 25. † Among degree holders. ‡ Median wage.

Introduction

Model

Data

Identification and estimation

Results

What's next and conclusion

Earnings expectations

- Students have *rational expectations* about their future earnings.

Earnings expectations

- Students have *rational expectations* about their future earnings.
- Only hold limited information (X) when forming expectations, Y_S^{ea}

Earnings expectations

- Students have *rational expectations* about their future earnings.
- Only hold limited information (X) when forming expectations, Y_S^{ea}
- Model Y_S^{ea} as (log-)linear conditional expectation:

$$Y_S^{ea} = \mathbb{E}[Y_S | X] = X' \beta_S$$

where Y_S are realised earnings given choice S

wage distribution

Earnings expectations

- Students have *rational expectations* about their future earnings.
- Only hold limited information (X) when forming expectations, Y_S^{ea}
- Model Y_S^{ea} as (log-)linear conditional expectation:

$$Y_S^{ea} = \mathbb{E}[Y_S | X] = X' \beta_S$$

where Y_S are realised earnings given choice S

wage distribution

- Key assumptions

Earnings expectations

- Students have *rational expectations* about their future earnings.
- Only hold limited information (X) when forming expectations, Y_S^{ea}
- Model Y_S^{ea} as (log-)linear conditional expectation:

$$Y_S^{ea} = \mathbb{E}[Y_S | X] = X' \beta_S$$

where Y_S are realised earnings given choice S

wage distribution

- Key assumptions
 - we observe all relevant variables, X [I plan to test this]

Earnings expectations

- Students have *rational expectations* about their future earnings.
- Only hold limited information (X) when forming expectations, Y_S^{ea}
- Model Y_S^{ea} as (log-)linear conditional expectation:

$$Y_S^{ea} = \mathbb{E}[Y_S | X] = X' \beta_S$$

where Y_S are realised earnings given choice S

wage distribution

- Key assumptions
 - we observe all relevant variables, X [I plan to test this]
 - log-linear wage equation [standard, <cite>]

wage model

Earnings expectations

- Students have *rational expectations* about their future earnings.
- Only hold limited information (X) when forming expectations, Y_S^{ea}
- Model Y_S^{ea} as (log-)linear conditional expectation:

$$Y_S^{ea} = \mathbb{E}[Y_S | X] = X' \beta_S$$

where Y_S are realised earnings given choice S

wage distribution

- Key assumptions
 - we observe all relevant variables, X [I plan to test this]
 - log-linear wage equation [standard, <cite>]
 - rational expectations [standard, <cite>]

wage model

Psychic costs

- Other factors in the utility function: Heckman's “psychic costs”

Psychic costs

- Other factors in the utility function: Heckman's "psychic costs"
- Usually "measured" with family background or residual term

Psychic costs

- Other factors in the utility function: Heckman's "psychic costs"
- Usually "measured" with family background or residual term
- I use subjective, open-ended questions about university

Psychic costs

- Other factors in the utility function: Heckman's "psychic costs"
- Usually "measured" with family background or residual term
- I use subjective, open-ended questions about university
- Similar responses identified and harmonised by survey designers

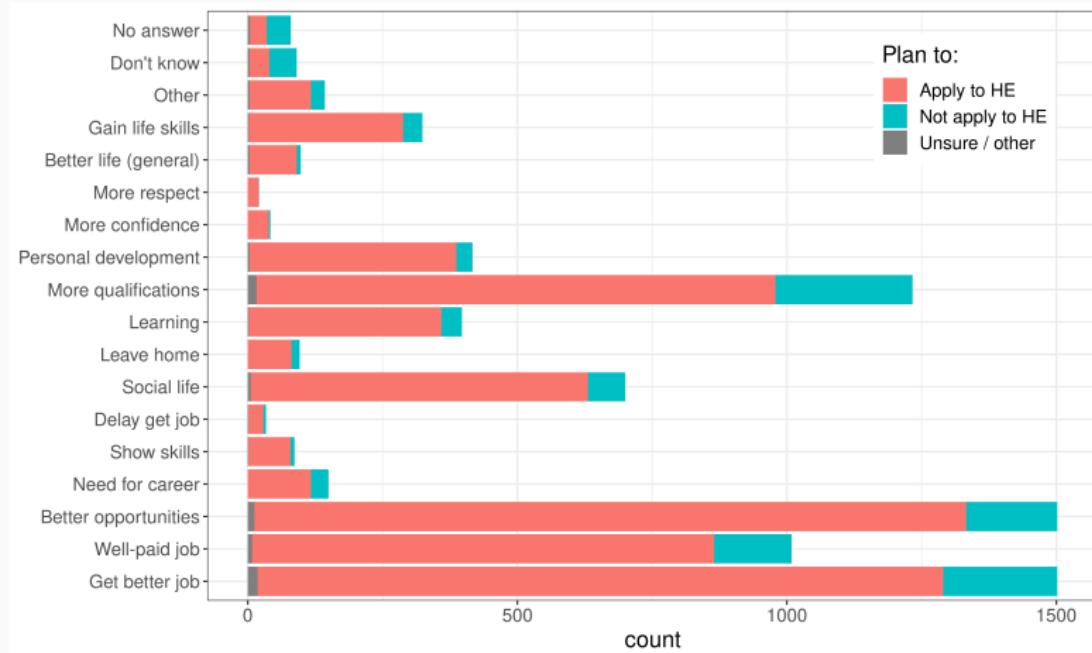
Advantages of attending university

What do you think the advantages, if any, might be for SOMEONE of going to university to study for a degree?

Notes: Asked to all students with >5 GCSEs @A*-C. Open-ended. N = 4,640.

Advantages of attending university

What do you think the advantages, if any, might be for SOMEONE of going to university to study for a degree?



Notes: Asked to all students with >5 GCSEs @A*-C. Open-ended. N = 4,640.

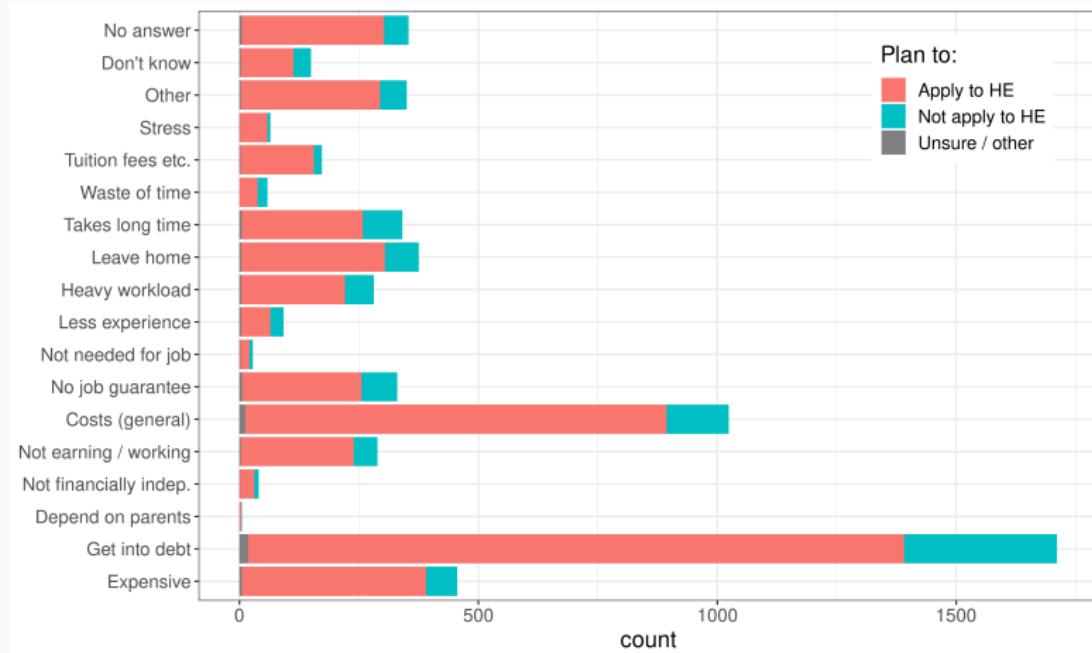
Disadvantages of attending university

What do you think the disadvantages, if any, might be for someone of going to university to study for a degree?

Notes: Asked to all students with >5 GCSEs @A*-C. Open-ended. N = 4,640.

Disadvantages of attending university

What do you think the disadvantages, if any, might be for someone of going to university to study for a degree?



Notes: Asked to all students with >5 GCSEs @A*-C. Open-ended. N = 4,640.

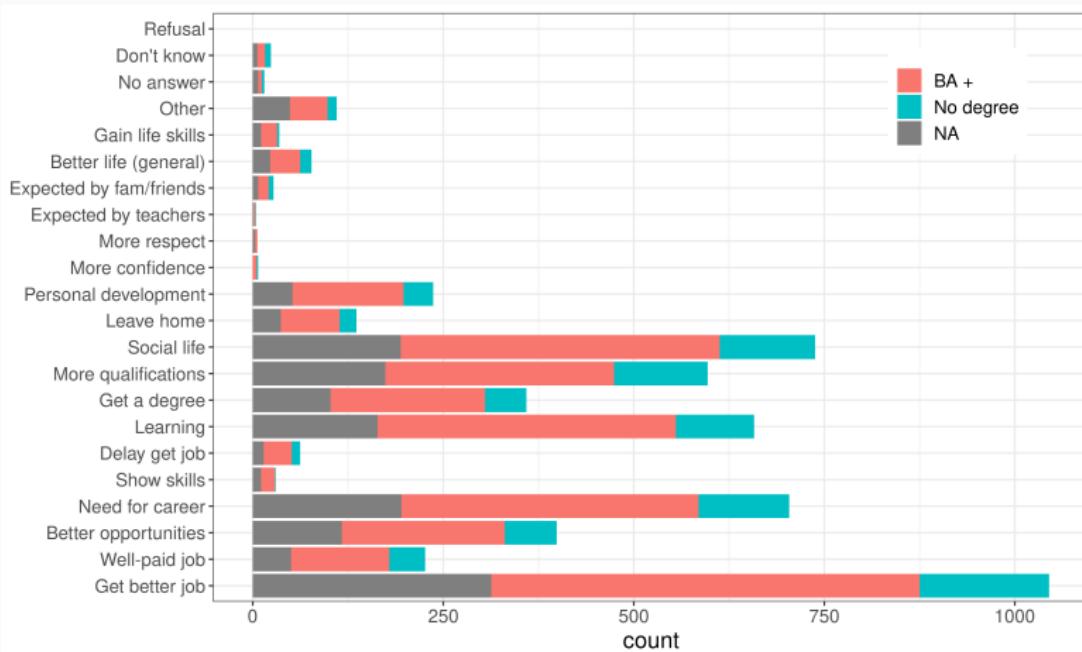
Main reasons for applying to university

Now thinking about yourself. You said you plan to apply for a place at university. What are YOUR main reasons for wanting to go to university?

Notes: Students with >5 GCSEs @A*-C and **who plan to apply**. Open-ended. N = 4,640.

Main reasons for applying to university

Now thinking about yourself. You said you plan to apply for a place at university. What are YOUR main reasons for wanting to go to university?



Notes: Students with >5 GCSEs @A*-C and who plan to apply. Open-ended. N = 4,640.

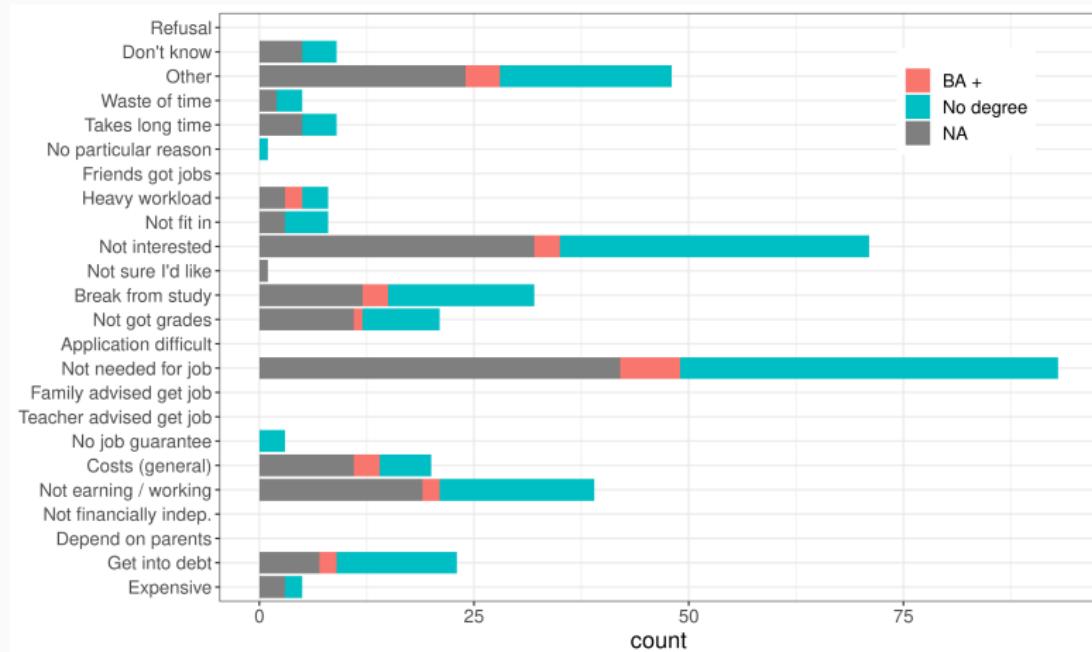
Main reasons for not applying

You have said that you are not planning to apply to university/not likely to ever apply to university. What are the MAIN reasons why you decided not to apply for a place at a university?

Notes: Students with >5 GCSEs @A*-C who **do not plan to apply**. Open-ended. $N = 4,640$.

Main reasons for not applying

You have said that you are not planning to apply to university/not likely to ever apply to university. What are the MAIN reasons why you decided not to apply for a place at a university?



Notes: Students with >5 GCSEs @A*-C who **do not plan to apply**. Open-ended. N = 4,640.

Estimation

Parameter estimates ($\widehat{\alpha}$, $\widehat{\beta_k}$, $\widehat{\Delta\delta} \equiv \widehat{\delta_1 - \delta_0}$)

- Assuming $\epsilon_0 - \epsilon_1 \sim Logit$ allows us to obtain estimates $\widehat{\alpha}$, $\widehat{\Delta\delta}$ by logit regression, given Y_k^{ea} and Z .
- estimate earnings as $\widehat{Y_k^{ea}} = X' \widehat{\beta_k}$, from OLS of Y_k on X

Choosing X : characteristics that affect the HE decision through earnings expectations at age 16

- parents': occupations, ethnicity group, education, income
- no. A-levels taking; gender; whether think high pay is important

Introduction

Model

Data

Identification and estimation

Results

Within and between cohorts: full sample analysis

Decomposing changes in earnings across cohorts

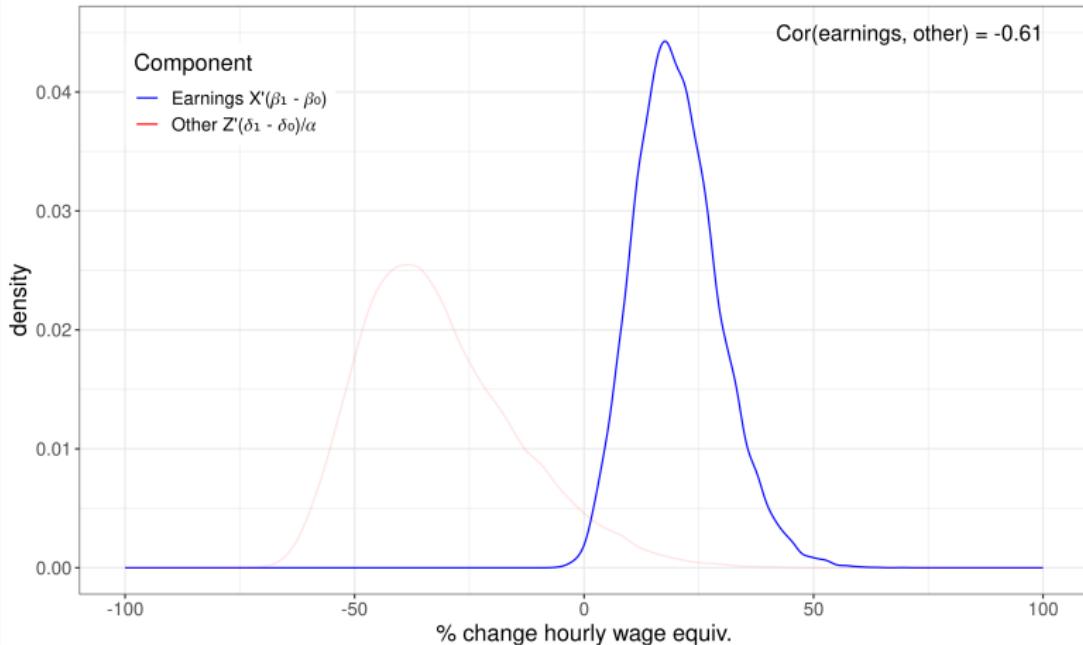
What's next and conclusion

Comparing earnings with other factors

- Compare relative importance of earnings versus other factors
- Strategy:
 - Estimate parameters $\alpha, \beta_k, \Delta\delta$
 - Use estimates and sample covariates to estimate distributions of university-wage premium, and non-earnings factors.
 - Transform these factors so they are equivalent to $\% \Delta$ in earnings
 - Plot the transformed distributions for each cohort and demographics within cohorts
- We can then compare:
 - earnings vs non-earnings factors within cohorts
 - the relative importance of these factors across cohorts

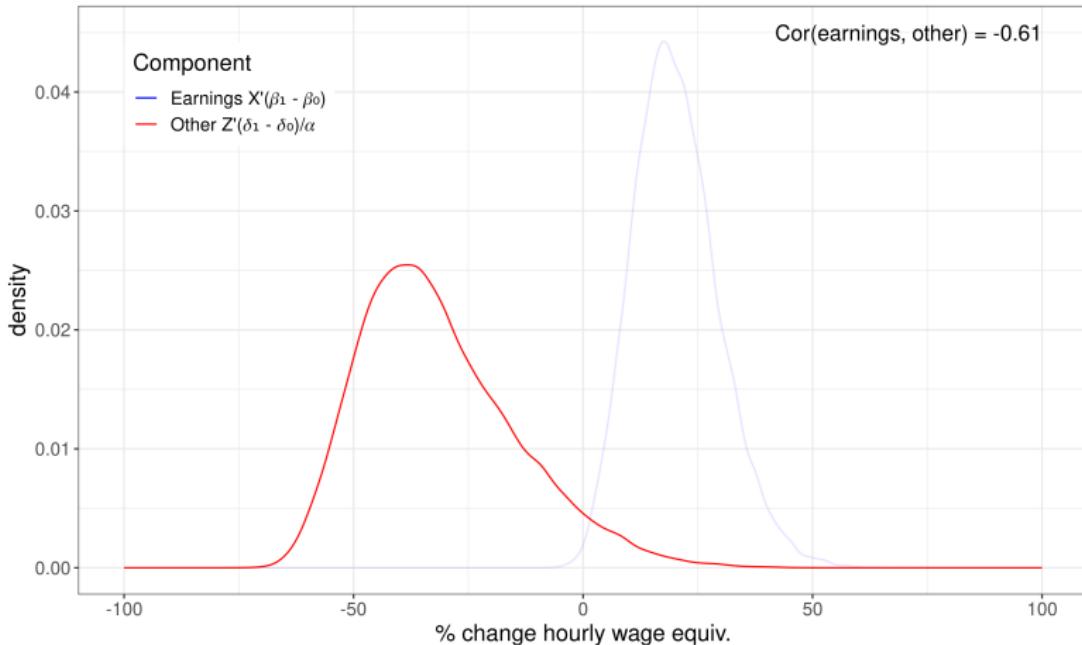
Full sample (1970)

- Earnings expectations exclusively positive (mean 20%)



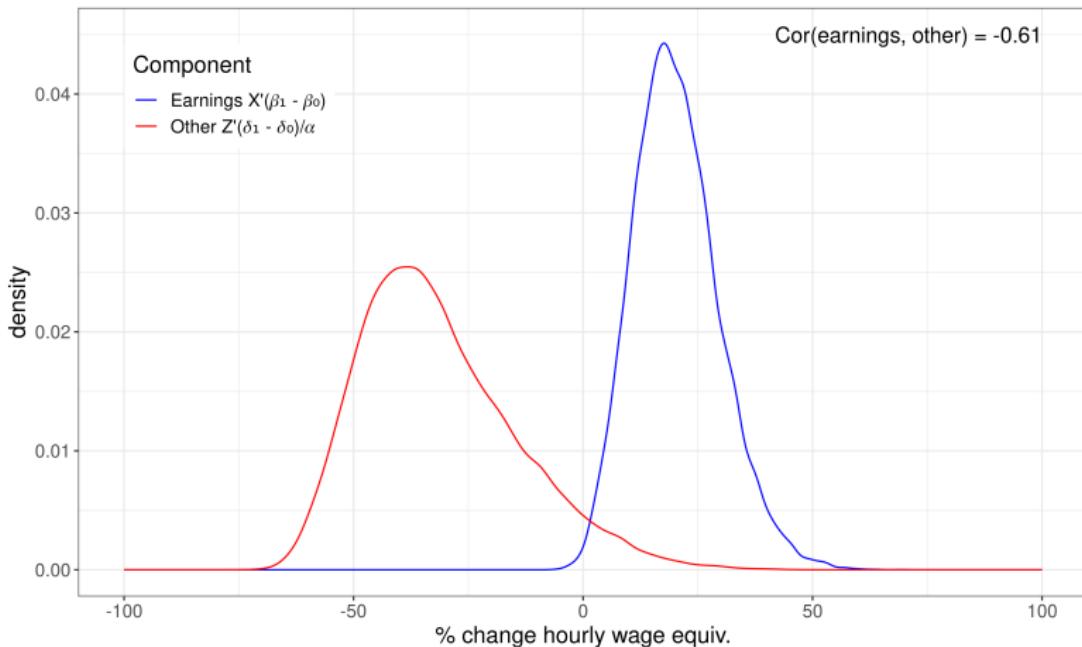
Full sample (1970)

- Earnings expectations exclusively positive (mean 20%)
- Other factors (almost) exclusively negative (mean -31%)



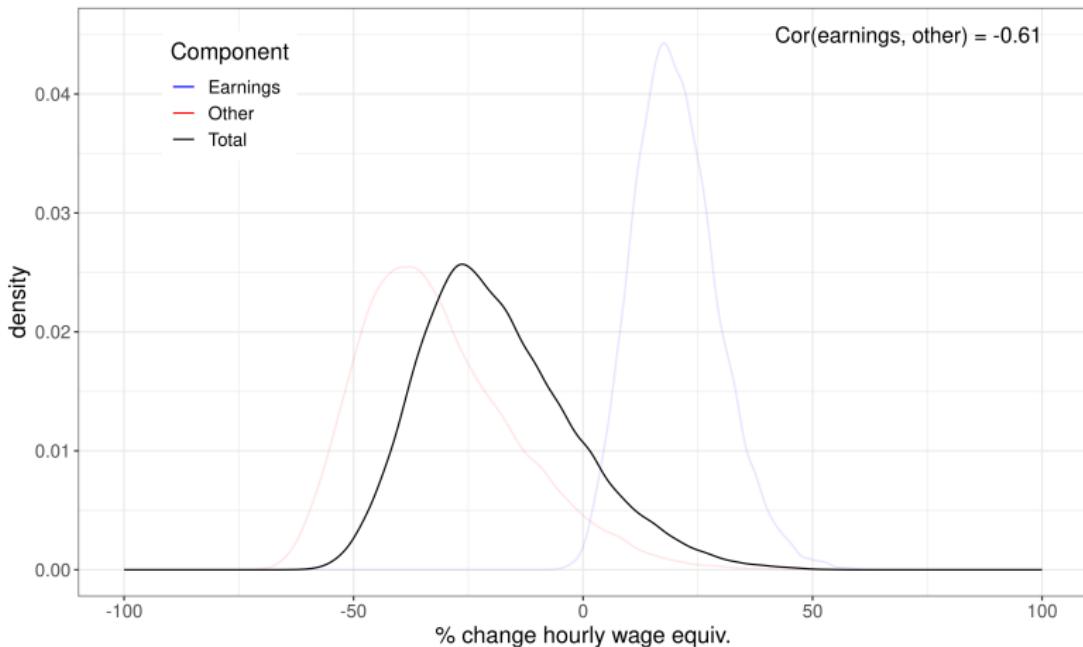
Full sample (1970)

- Earnings expectations exclusively positive (mean 20%)
- Other factors (almost) exclusively negative (mean -31%)
- $sd(\text{earnings})$ approx. half $sd(\text{other factors})$ (9.3 vs 17.0)



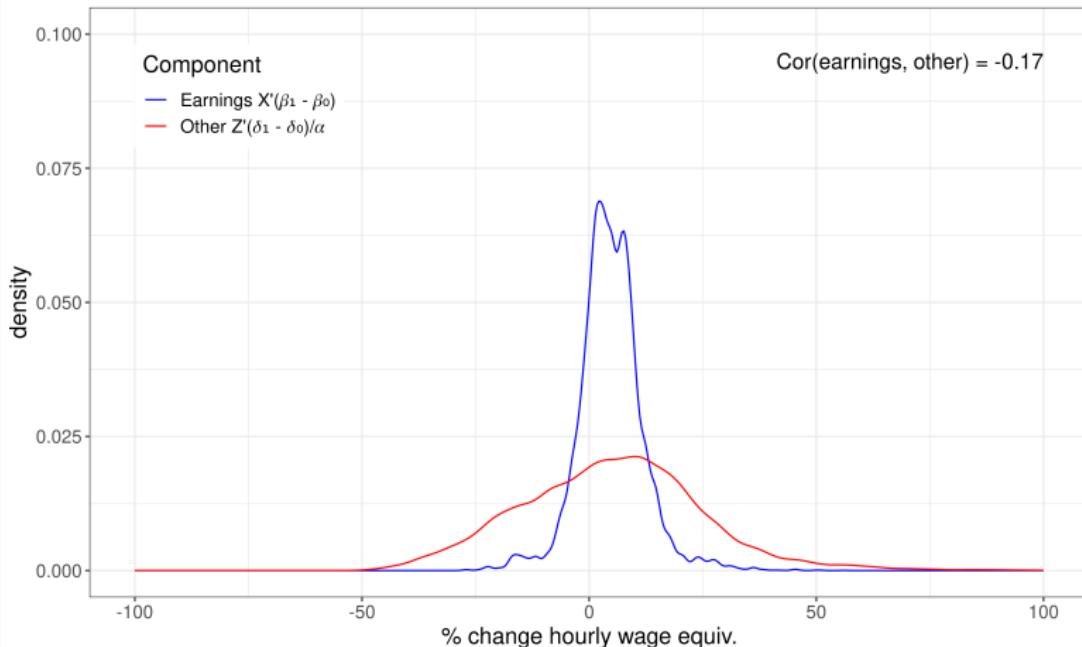
Full sample (1970)

- Earnings expectations exclusively positive (mean 20%)
- Other factors (almost) exclusively negative (mean -31%)
- $sd(\text{earnings})$ approx. half $sd(\text{other factors})$ (9.3 vs 17.0)
- Larger (abs.) mean and variance \Rightarrow other factors more important



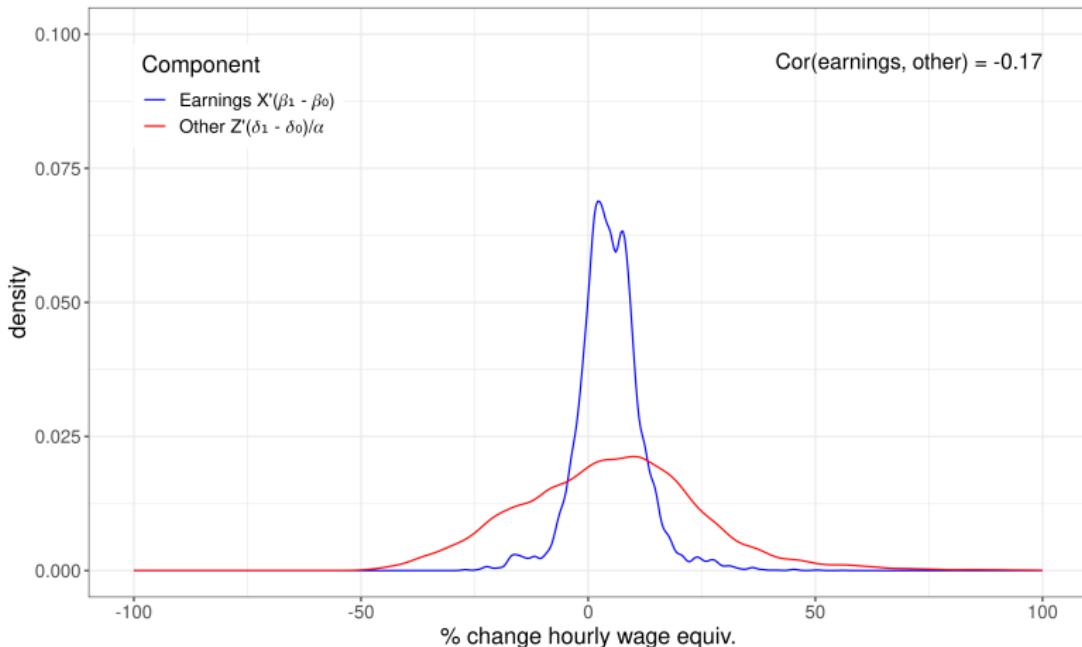
Full sample (1990)

- Both distributions similarly located, positive means (**4.8%** vs **5.6%**)



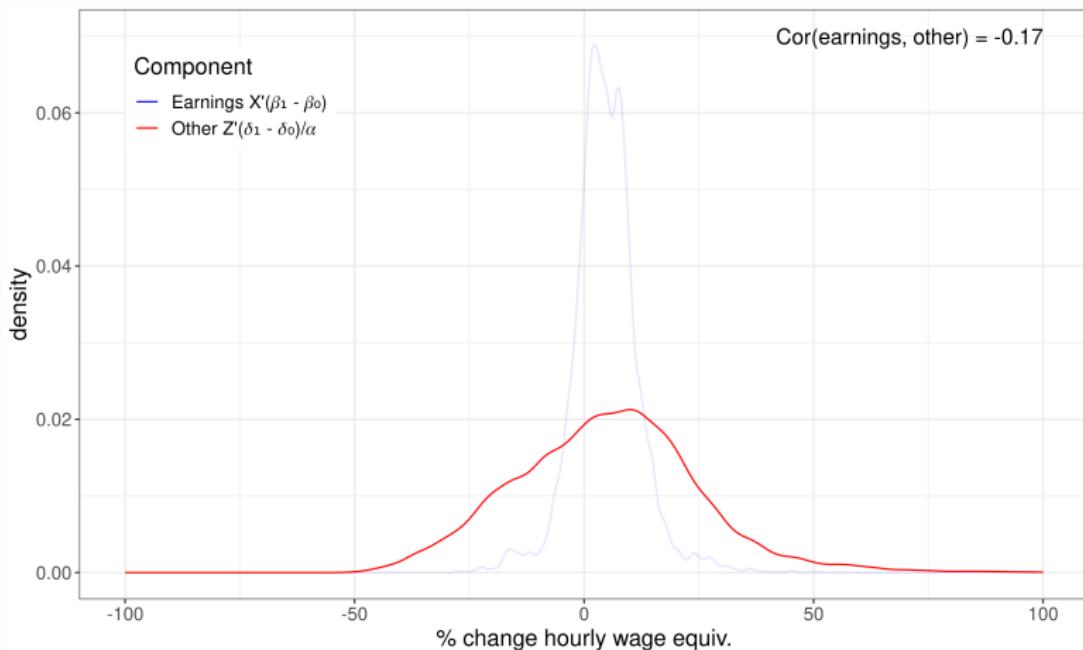
Full sample (1990)

- Both distributions similarly located, positive means (4.8% vs 5.6%)
- Earnings distribution has a much lower std dev. (7.1% vs 41%)



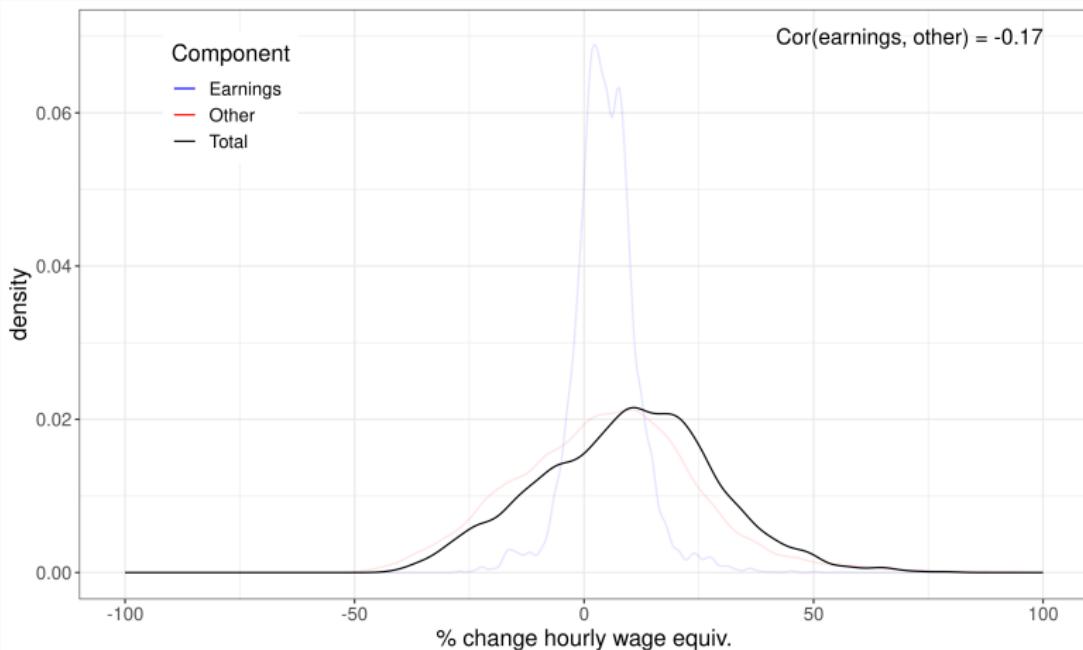
Full sample (1990)

- Both distributions similarly located, positive means (4.8% vs 5.6%)
- Earnings distribution has a much lower std dev. (7.1% vs 41%)
- Other factors distribution possesses significant *positive skew*



Full sample (1990)

- Both distributions similarly located, positive means (4.8% vs 5.6%)
- Earnings distribution has a much lower std dev. (7.1% vs 41%)
- Other factors distribution possesses significant *positive skew*
- Similar impact on rate, other factors more important for variation



Summary and discussion of full sample results

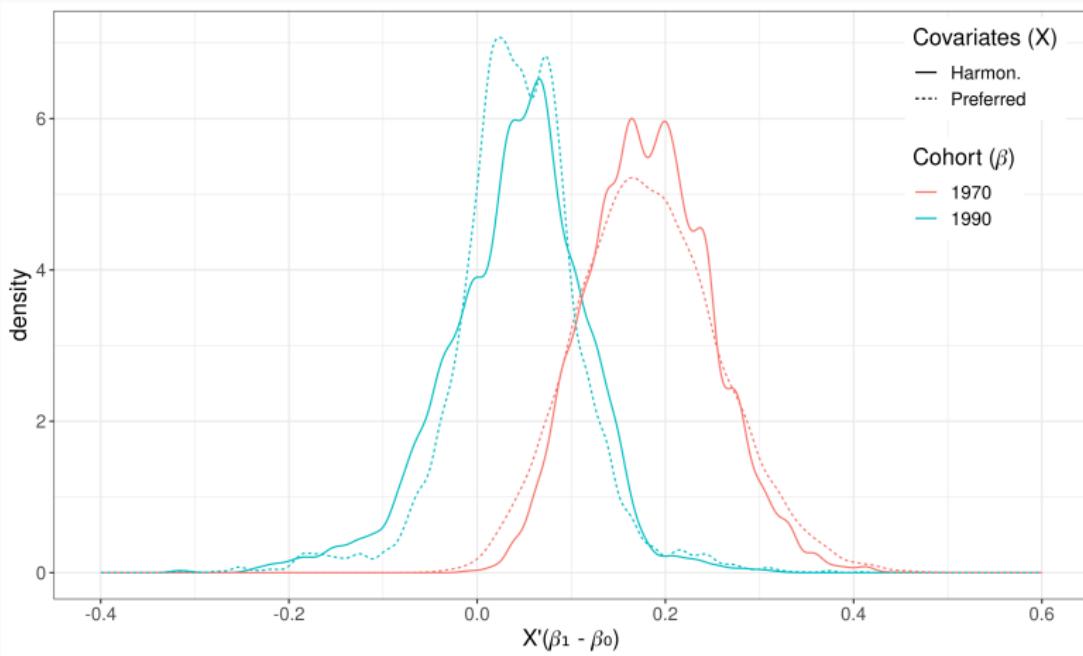
- Shape of earnings and other factors distributions quite different within cohorts:
 - **earnings symmetric, low variance**
 - **other positive skew, larger variance**
- In 1970, **earnings generally positive** and **other factors negative**
- Moving to 1990 we see:
 - a slight decrease in earnings expectations (still mostly +ve)
 - and a large increase in other factors ($-ve \rightarrow$ mostly +ve)
- In both cohorts other factors are responsible for more of the variance in HE attainment (higher variance)
- It is also the changing location of the other factors distribution that drives the large increase in attainment

Decomposing changes in earnings across cohorts

- Now focus only on earnings expectations, and estimate the same specification across both cohorts.
- Requires some combining of groups to harmonise variables across cohorts
- Strategy
 - Estimate identical specification wage model for both cohorts
 - Calculate university wage premium for each cohort
 - Calculate premium for each cohort *using estimated parameters from the other cohort*, i.e. $X_{1970}(\beta_{1,1990} - \beta_{0,1990})$
- Allows us to decompose changes across cohorts into
 - structural changes (changes in $\beta_1 - \beta_0 \equiv \Delta\beta$)
 - compositional changes (changes in distribution of X within cohort)

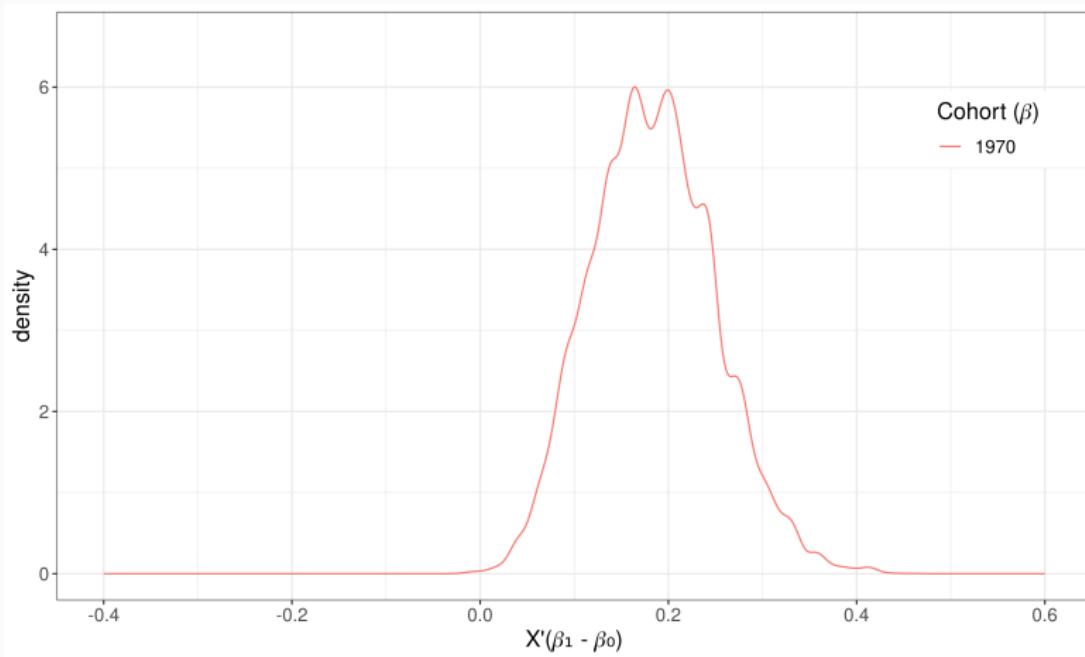
Decomposing earnings expectations: 1970 vs 1990

- “Harmonising” parameters does not affect the distribution estimates very much



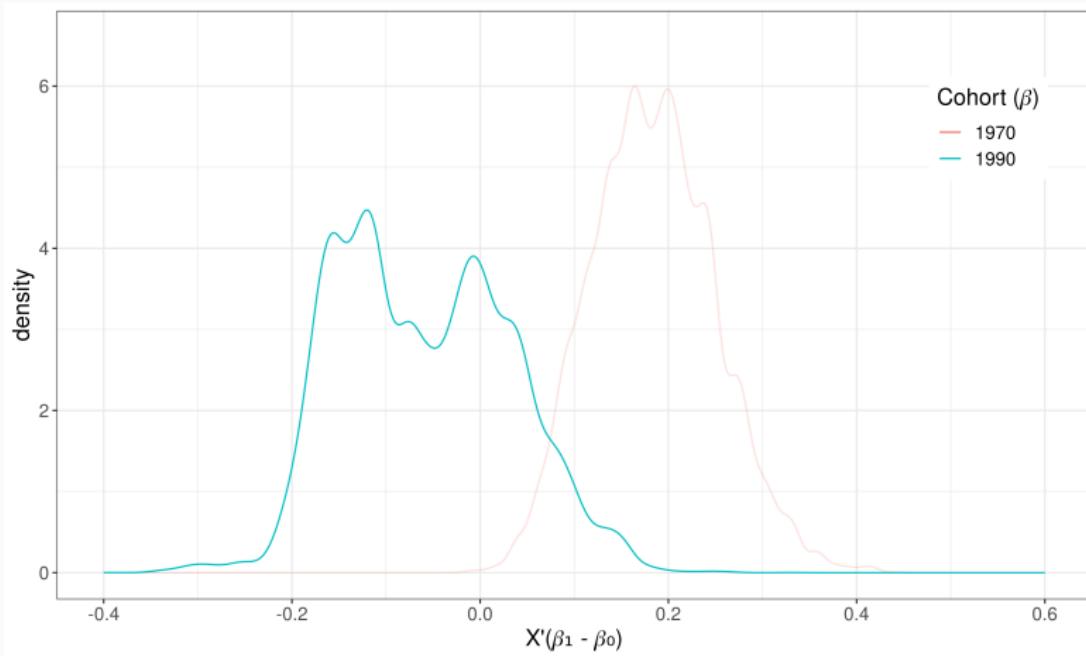
Decomposing earnings expectations: 1970 vs 1990

- Starting from β_{1970} and X_{1970}



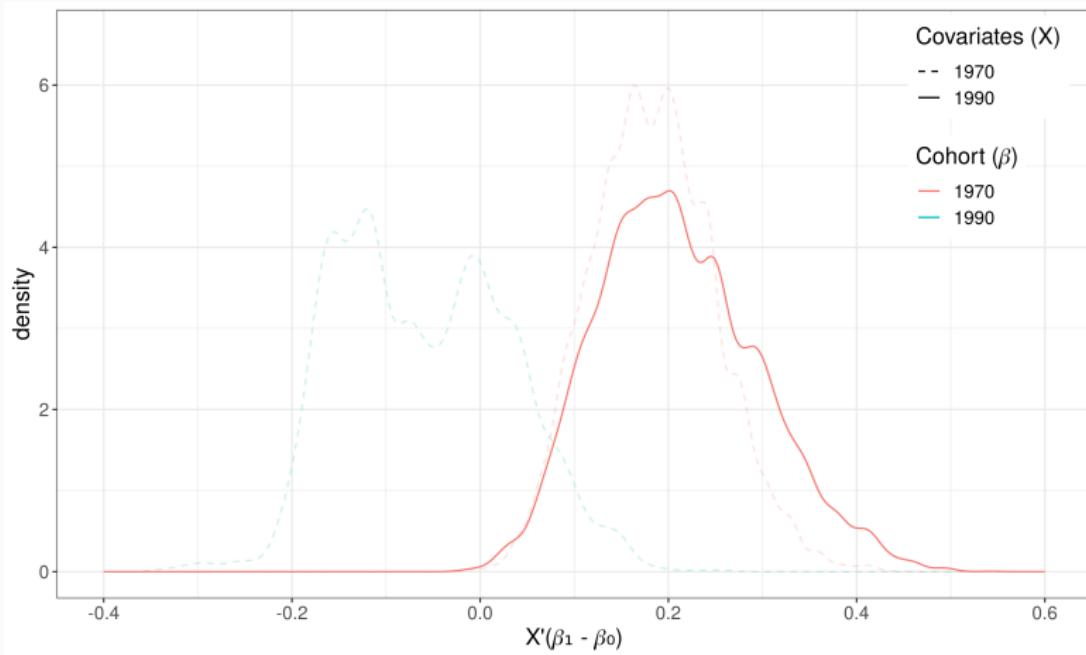
Decomposing earnings expectations: 1970 vs 1990

- Changing to β_{1990} induces a large decrease in earnings expectations...



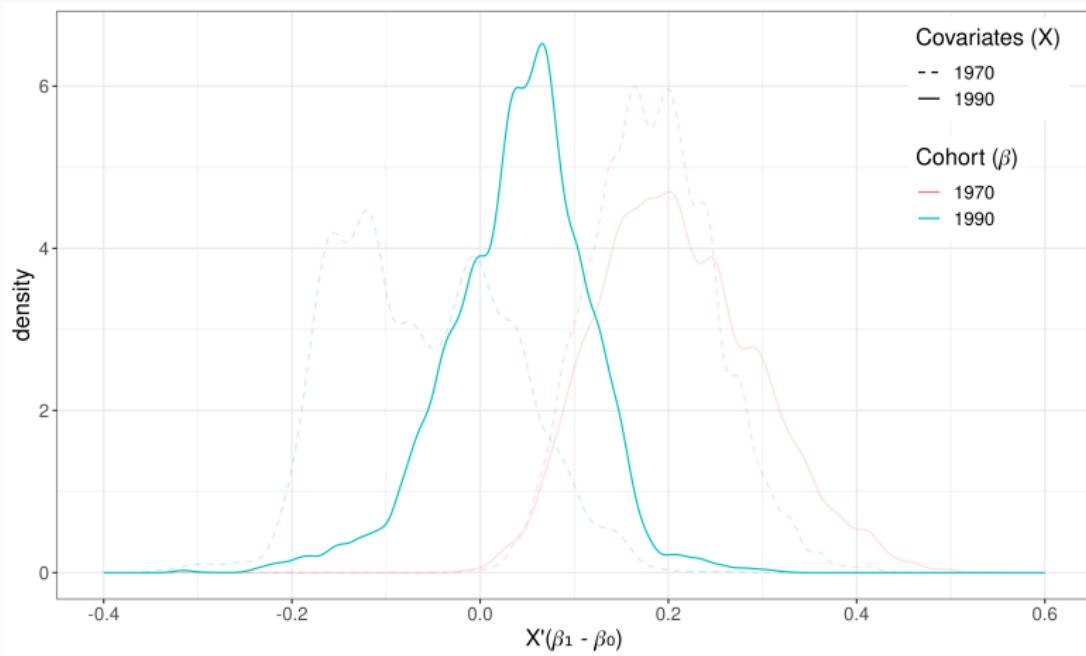
Decomposing earnings expectations: 1970 vs 1990

- ...while changing to X_{1990} (β_{1970}) has little effect.



Decomposing earnings expectations: 1970 vs 1990

- Changing both X and β results in reduced expectations, though by less than just changing β .



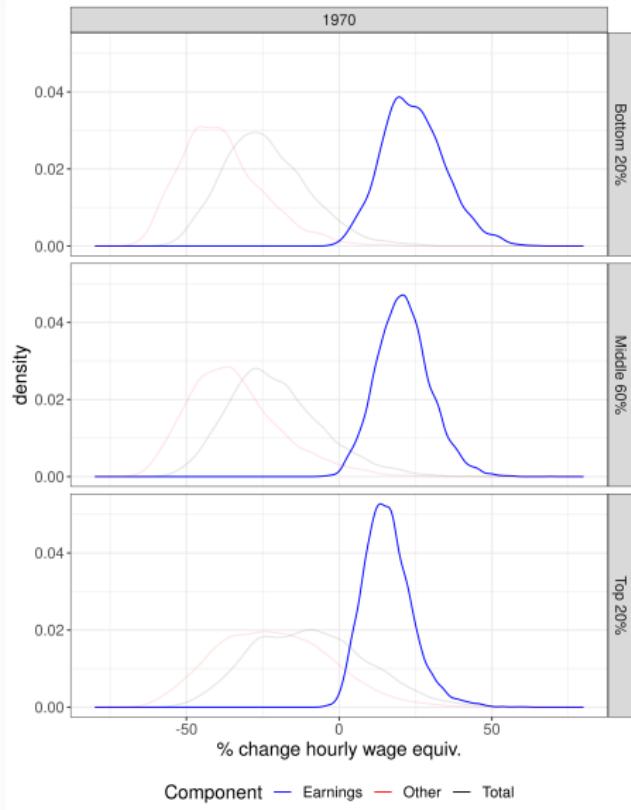
Summary of decomposition

- Purpose of exercise is to compare the effect of changes in parameters versus changes in X on earnings expectations
- This allows two “counterfactual scenarios”:
 - moving the 1970 cohort to 1990 ($X'_{1970}\Delta\beta_{1990}$)
 - moving the 1990 cohort to 1970 ($X'_{1990}\Delta\beta_{1970}$)
- We see that β_{1990} result in lower expectations than β_{1970}
- Changing $\beta_{1970} \rightarrow \beta_{1990}$ results in large decrease in expectations (X_{1970}), while $\beta_{1990} \rightarrow \beta_{1970}$ (X_{1990}) results in increase
- Change in composition ($X_{1970} \rightarrow X_{1990}$) tempers the change due to parameters so the final distribution lies between $X'_{1970}\Delta\beta_{1970}$ and $X'_{1970}\Delta\beta_{1990}$

Comparing factors across SES groups

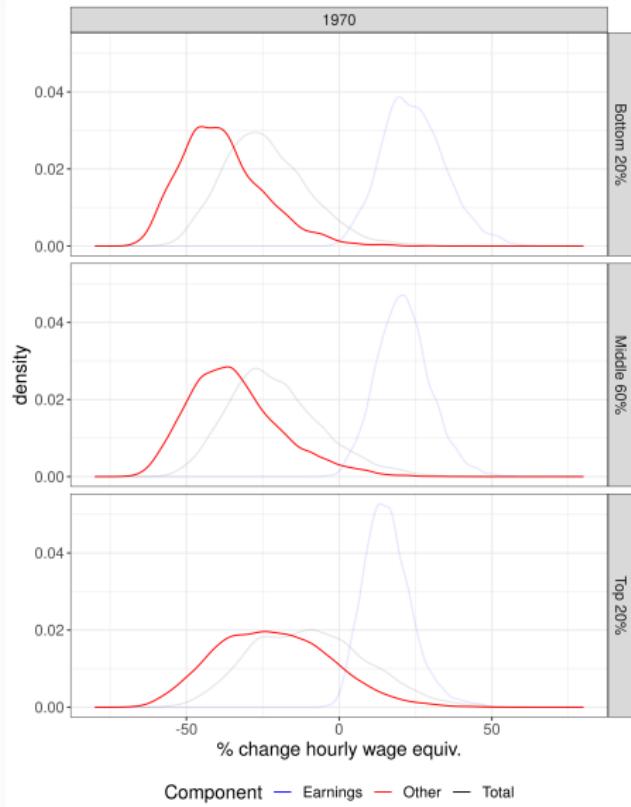
- Aim of exercise is to compare factors in the HE decision across SES groups, and uncover the cause of the SES gap
- Given the longitudinal nature of the surveys, we have detailed information on the family background of students at 16
- Lack a consistent definition of SES, so use parental income
- The following slides present the distributions of earnings expectations and other factors, *conditional on parental income quintiles at 16*
- The parameters are constant across the cohort, but each income group contains only the X, Z of individuals in that group
- **Extension:** re-estimate the model on subsamples by parental income

By parents' income (1970)



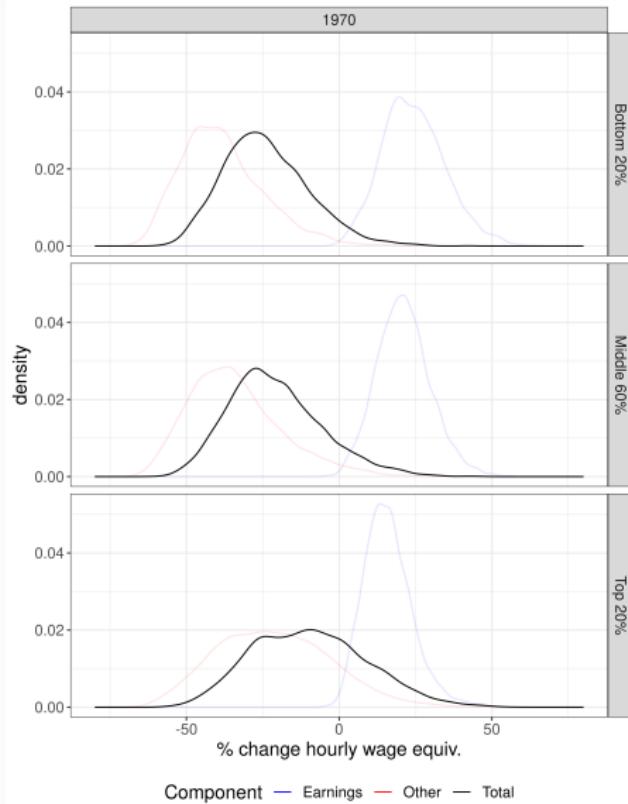
- Mean earnings expectations decreasing in parental income, variance also decreasing

By parents' income (1970)



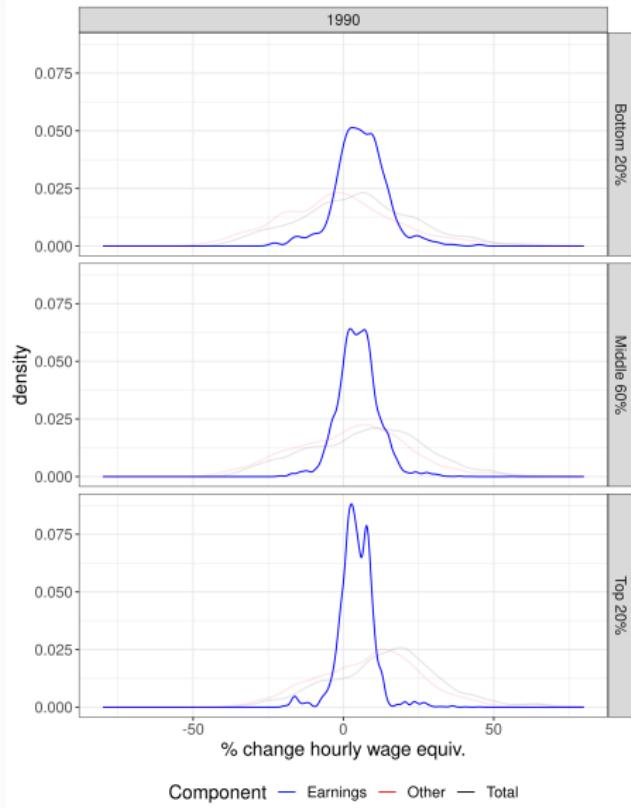
- Mean earnings expectations decreasing in parental income, variance also decreasing
- Other factors very similar distribution for low- and middle-income, higher mean and variance for top 20%

By parents' income (1970)



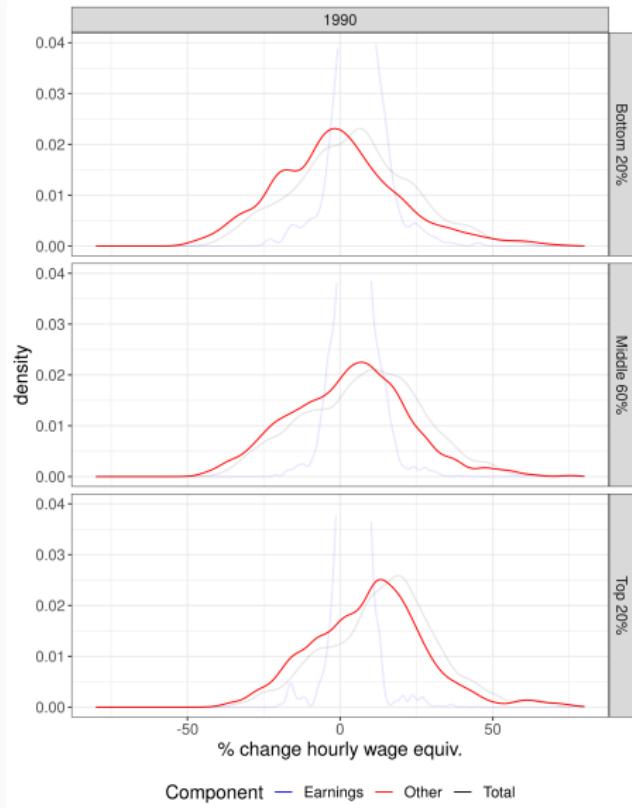
- Mean earnings expectations decreasing in parental income, variance also decreasing
- Other factors very similar distribution for low- and middle-income, higher mean and variance for top 20%
- Combined shows the importance of other factors, and can clearly see HE rate increasing with income

By parents' income (1990)



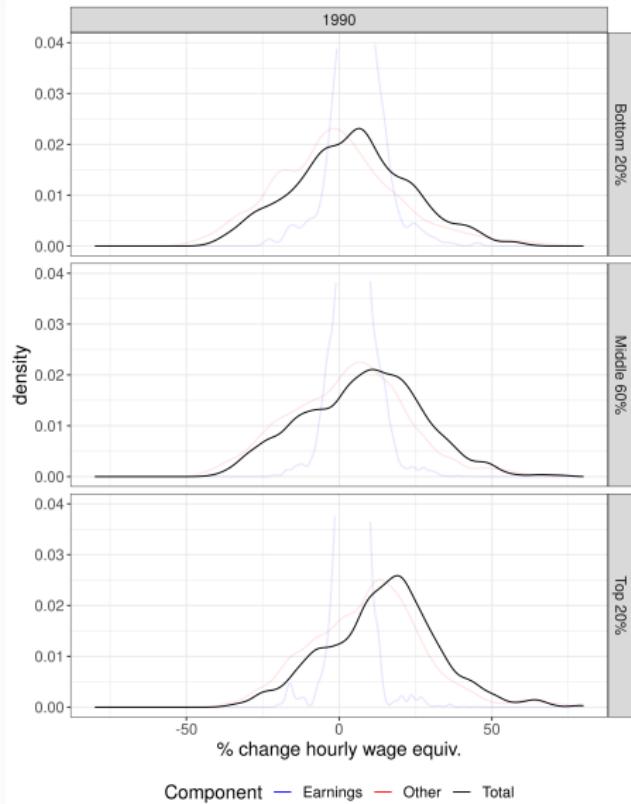
- Mean earnings expectations similar across income groups, variance decreasing in income

By parents' income (1990)



- Mean earnings expectations similar across income groups, variance decreasing in income
- Other factors have very similar-shaped distributions across groups, mean increasing in income

By parents' income (1990)



- Mean earnings expectations similar across income groups, variance decreasing in income
- Other factors have very similar-shaped distributions across groups, mean increasing in income
- Again combining earnings and other factors shows the importance of other factors

Summary of SES analysis

- In both cohorts, other factors are the main contributor to the SES gap in HE attainment
- Mean earnings expectations actually slightly decreasing in income for 1970 cohort
- Though the variance of earnings expectations is decreasing in income for both cohorts
- Meanwhile, clear that other factors contribution to the HE decision is increasing in earnings
- Further work:
 - Robustness check estimating model separately on subsamples (allow decomposition analysis)
 - Understand the main components of other factors (“psychic costs”)

Introduction

Model

Data

Identification and estimation

Results

What's next and conclusion

Still to do

- Decompose other factors into meaningful components of “psychic costs”
 - Create a mapping, $f : Z \rightarrow \theta$ (via PCA or manual classification)
 - Re-estimate model and redo analysis on components of θ
- Allow for unobserved heterogeneity in earnings
 - May be missing information students use to forecast their earnings
 - Following Heckman et al. (2006), allow for this missing data by explicitly including it in the model of earnings expectations

Allowing for unobserved heterogeneity

Cunha et al. (2004) show how to allow for unobserved heterogeneity in earnings. Their strategy proceeds as follows:

- Impose a factor structure on the errors from earnings and SEU equations:

$$U_s = \eta\alpha_s + \varepsilon_s \quad \forall s \in \{1, 0, SEU\}$$

where $U_k = Y_k - \beta_k X$ and $U_{SEU} = SEU^* - Y^{ea} - \theta' \Delta \gamma$

- The factor loadings α_1 and α_0 can be identified from:

$$\text{Cov}(U_1, U_{SEU}) = \alpha_1 \alpha_{SEU} \sigma_\eta^2$$

$$\text{Cov}(U_0, U_{SEU}) = \alpha_0 \alpha_{SEU} \sigma_\eta^2$$

normalising $\alpha_{SEU} = 1$ and $\sigma_\eta^2 = 1$, assuming $\eta \perp\!\!\!\perp (\varepsilon_1, \varepsilon_0, \varepsilon_{SEU})$

- Then test whether component can predict schooling choice: i.e. is it used by the individual in HE decision?

Conclusion

- Summary of findings so far
 - Earnings expectations (as measured here) are much less important than other factors in the decision to continue to higher education (in line with recent results from other research)
 - Over a period of huge growth in HE attainment, earnings expectations in the UK have fallen slightly and it is other factors that have driven this increase
 - A similar story explains the continuing SES gap in attainment: earnings expectations are broadly similar across parental income groups, and other factors are the culprit
- Next steps
 - Decompose other factors into meaningful components

Thank you

Thank you :)

References i

- Abbring, J. H. and Heckman, J. J. (2007). Chapter 72 Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation. In *Handbook of Econometrics*, volume 6, pages 5145–5303. Elsevier.
- Arcidiacono, P., Hotz, V. J., Maurel, A., and Romano, T. (2019). Ex Ante Returns and Occupational Choice. page 47.
- Bhuller, M., Mogstad, M., and Salvanes, K. G. (2014). Life Cycle Earnings, Education Premiums and Internal Rates of Return. page 34.
- Blanden, J. and Machin, S. (2004). Educational Inequality and the Expansion of UK Higher Education. *Scottish Journal of Political Economy*, 51(2):230–249.
- Blundell, R., Green, D., and Jin, W. (2018). The UK Education Expansion and Technological Change. page 71.
- Boneva, T. and Rauh, C. (2019). Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns. *Journal of Political Economy*, page 83.
- Cameron, S. V. and Heckman, J. J. (1993). Determinants of Young Male Schooling and Training Choices. Working Paper 4327, National Bureau of Economic Research.

References ii

- Cunha, F. and Heckman, J. (2007a). The Evolution of Labor Earnings Risk in the U.S. Economy. page 53.
- Cunha, F. and Heckman, J. J. (2007b). Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling. *Labour Economics*, 14(6):870–893.
- Cunha, F., Heckman, J. J., and Navarro, S. (2004). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers*, 57(2):191–261.
- Dearden, L., van der Erve, L., Belfield, C., and Britton, J. (2017). Higher Education funding in England: past, present and options for the future. Technical report. ISBN: 9781911102489.
- Delavande, A. and Zafar, B. (2017). University Choice: The Role of Expected Earnings, Non-pecuniary Outcomes and Financial Constraints. page 56.
- D'Haultfoeuille, X. and Maurel, A. (2013). Inference on an extended Roy model, with an application to schooling decisions in France. *Journal of Econometrics*, 174(2):95–106.
- Fuster, A., Perez-Truglia, R., and Zafar, B. (2019). Expectations with Endogenous Information Acquisition: An Experimental Investigation. page 68.

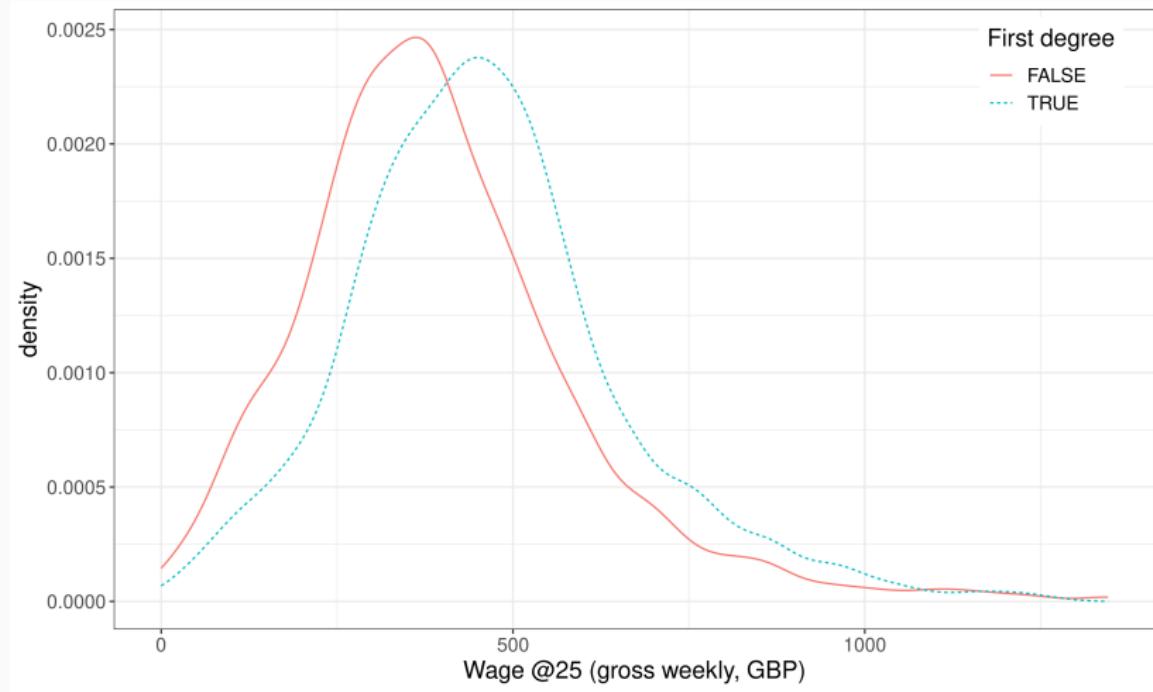
References iii

- Green, D. A., Jin, W. M., and Blundell, R. (2016). The UK wage premium puzzle: how did a large increase in university graduates leave the education premium unchanged? Working Paper Series, IFS. Series: Working Paper Series.
- Heckman, J. J., Lochner, L. J., and Todd, P. E. (2006). Chapter 7 Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. In *Handbook of the Economics of Education*, volume 1, pages 307–458. Elsevier.
- Heckman, J. J. and Navarro, S. (2007). Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics*, 136(2):341–396.
- Heckman, J. J. and Vytlacil, E. J. (2007). Chapter 71 Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6, pages 4875–5143. Elsevier.
- Mackowiak, B., Matejka, F., and Wiederholt, M. (2018). Survey: Rational Inattention, a Disciplined Behavioral Model. page 42.
- OECD (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*. PISA. OECD.

References iv

- OECD (2020). Public spending on education (indicator). *Education at a glance: Educational finance indicators*.
- Walker, I. and Zhu, Y. (2008). The College Wage Premium and the Expansion of Higher Education in the UK. *The Scandinavian Journal of Economics*, 110(4):695–709.
- Wiswall, M. and Zafar, B. (2015). Determinants of College Major Choice: Identification using an Information Experiment. *The Review of Economic Studies*, 82(2):791–824.
- Wiswall, M. and Zafar, B. (2016). Human Capital Investments and Expectations about Career and Family. Technical Report w22543, National Bureau of Economic Research, Cambridge, MA.
- Zafar, B. (2009). How do College Students Form Expectations? *SSRN Electronic Journal*.

Wage distributions (Next Steps)



back

Wage model

$$Y_S = X'\beta_S + U_S \quad (3)$$

back