The role of earnings expectations versus non-pecuniary factors in university attendance

Oliver Cassagneau-Francis

18th November 2020

Sciences Po

Motivation

Why care about the decision to attend higher education?

Motivation

Why care about the decision to attend higher education?

- **Lots of evidence of benefits of a university degree**
  - Increased earnings
  - Better health

Motivation

Why care about the decision to attend higher education?

- 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)

Motivation

Why care about the decision to attend higher education?

- 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 who enjoys these state-sponsored benefits
  - England: children of parents' earning in top 20% **twice as likely to attend university** as children of parents in the bottom 20%
  - "SES-gap" in attainment up to 30pp across OECD countries (OECD, 2018)
  - is HE actually hindering social mobility?

Motivation

Why care about the decision to attend higher education?

- 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

Understanding the factors that influence educational attainment is key, not only for educational outcomes but also for wider issues such as inequality and who benefits from public spending

Research questions + contributions

- **How important are earnings expectations vs other factors in decision to attend university?**
- What's driving the SES-gap in educational attainment?
- How have these factors changed over a period of expansion in HE attainment (1980s to today)?

Research questions + contributions

- **How important are earnings expectations vs other factors in decision to attend university?**
- Previous work:
  - US: Heckman et al. (2006) estimate earnings and "psychic costs"
  - France: D'Haultfoeuille and Maurel (2013) similar exercise
  - both rely on family background heterogeneity to estimate non-pecuniary factors
- This paper:
  - I exploit data on expectations about non-pecuniary aspects
    $\rightarrow$ more relevant heterogeneity
    $\rightarrow$ can **decompose** non-pecuniary factors [in progress]
  $\Rightarrow$ and find a large role for non-pecuniary factors
- What's driving the SES-gap in educational attainment?
- How have these factors changed over a period of expansion in HE attainment (1980s to today)?

Research questions + contributions

- **How important are earnings expectations vs other factors in decision to attend university?**
- Previous work:
  - US: Heckman et al. (2006) estimate earnings and "psychic costs"
  - France: D'Haultfoeuille and Maurel (2013) similar exercise
  - both rely on family background heterogeneity to estimate non-pecuniary factors
- This paper:
  - I exploit data on expectations about non-pecuniary aspects
    $\rightarrow$ more relevant heterogeneity
    $\rightarrow$ can **decompose** non-pecuniary factors [in progress]

  $\Rightarrow$ and find a large role for non-pecuniary factors
- What's driving the SES-gap in educational attainment?
- How have these factors changed over a period of expansion in HE attainment (1980s to today)?

Research questions + contributions

- How important are earnings expectations vs other factors in decision to attend university?
- **What's driving the SES-gap in educational attainment?**
- How have these factors changed over a period of expansion in HE attainment (1980s to today)?

Research questions + contributions

- How important are earnings expectations vs other factors in decision to attend university?
- **What's driving the SES-gap in educational attainment?**
- Previous work:
  - UK: Boneva and Rauh (2019) survey students aged 13–18 in 2017
  - 10% response rate, students still at university
    $\Rightarrow$ find equal role for pec. and non-pec. factors in SES gap
- This paper:
  - I use data on students *realised* wages + non-pec. expectations
  - representative sample
    $\Rightarrow$ find non-pec. factors alone explain SES gap
- How have these factors changed over a period of expansion in HE attainment (1980s to today)?

Research questions + contributions

- How important are earnings expectations vs other factors in decision to attend university?
- **What's driving the SES-gap in educational attainment?**
- Previous work:
  - UK: Boneva and Rauh (2019) survey students aged 13–18 in 2017
  - 10% response rate, students still at university
    $\Rightarrow$ find equal role for pec. and non-pec. factors in SES gap
- This paper:
  - I use data on students *realised* wages + non-pec. expectations
  - representative sample
    $\Rightarrow$ find non-pec. factors alone explain SES gap
- How have these factors changed over a period of expansion in HE attainment (1980s to today)?

Research questions + contributions

- How important are earnings expectations vs other factors in decision to attend university?
- What's driving the SES-gap in educational attainment?
- **How have these factors changed over a period of expansion in HE attainment (1980s to today)?**

▸ HE exp.

Research questions + contributions

- How important are earnings expectations vs other factors in decision to attend university?
- What's driving the SES-gap in educational attainment?
- **How have these factors changed over a period of expansion in HE attainment (1980s to today)?**
- Previous work:
  - returns: Walker and Zhu (2008); Green et al. (2016)
    $\Rightarrow$ graduate wage premium flat over this period
- This paper:
  - I compare factors across cohorts spanning this period
    $\Rightarrow$ slight decrease in expected wage premium
    $\Rightarrow$ large increase in expected non-pec. benefits

▸ wage premium

Research questions + contributions

- How important are earnings expectations vs other factors in decision to attend university?
- What's driving the SES-gap in educational attainment?
- **How have these factors changed over a period of expansion in HE attainment (1980s to today)?**
- Previous work:
  - returns: Walker and Zhu (2008); Green et al. (2016)
    $\Rightarrow$ graduate wage premium flat over this period
- This paper:
  - I compare factors across cohorts spanning this period
    $\Rightarrow$ slight decrease in expected wage premium
    $\Rightarrow$ large increase in expected non-pec. benefits

Road map

- Model the decision to go to university
- Panel data for cohort born in 1989/90
- Model + data $\rightarrow$ estimate distributions of factors
- Re-estimate model on cohort born in 1970

Road map

- **Model the decision to go to university**
  - Extended-Roy model of educational choice at age 16 / 17
  - Explicitly include earnings expectations and other factors
- Panel data for cohort born in 1989/90
- Model + data $\rightarrow$ estimate distributions of factors
- Re-estimate model on cohort born in 1970

Road map

- Model the decision to go to university
- **Panel data for cohort born in 1989/90**
  - Schooling, background, and non-pec. expectations from before decision
  - Choices: did they *complete* university? (96% completion rate)
  - Earnings and occupation after entry to labour market (age 25)
- Model + data $\rightarrow$ estimate distributions of factors
- Re-estimate model on cohort born in 1970

Road map

- Model the decision to go to university
- Panel data for cohort born in 1989/90
- Model + data $\rightarrow$ estimate **distributions of factors**
  - Compare earnings vs other factors ("psychic costs")
  - $\rightarrow$ quantify their relative importance in decision
  - Compare across different groups (e.g. SES)
  - $\rightarrow$ which factors are driving the SES-gap?
- Re-estimate model on cohort born in 1970

Road map

- Model the decision to go to university
- Panel data for cohort born in 1989/90
- Model + data $\rightarrow$ estimate **distributions of factors**
  - Compare earnings vs other factors ("psychic costs")
  - $\rightarrow$ quantify their relative importance in decision
  - Compare across different groups (e.g. SES)
  - $\rightarrow$ which factors are driving the SES-gap?
- Re-estimate model on cohort born in 1970

Road map

- Model the decision to go to university
- Panel data for cohort born in 1989/90
- Model + data $\rightarrow$ estimate **distributions of factors**
  - Compare earnings vs other factors ("psychic costs")
  - $\rightarrow$ quantify their relative importance in decision
  - Compare across different groups (e.g. SES)
  - $\rightarrow$ which factors are driving the SES-gap?
- Re-estimate model on cohort born in 1970

Road map

- Model the decision to go to university
- Panel data for cohort born in 1989/90
- Model + data $\rightarrow$ estimate distributions of factors
- **Re-estimate model on cohort born in 1970**
  - Compare across cohorts
  - $\rightarrow$ understand drivers of expansion of HE in England

Introduction

Extended Roy-model including psychic costs

Large-cohort panel data

Identification and estimation

Results

Conclusion

Factors in the decision to attend university

What do students consider when making educational choices?

Factors in the decision to attend university

What do students consider when making educational choices?

- (future) earnings

Factors in the decision to attend university

What do students consider when making educational choices?

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

Factors in the decision to attend university

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
  - $\pm$ studying
  - $\pm$ leave home / local area
  - $-$ not earning / gaining experience
  - $-$ stress

Factors in the decision to attend university

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
  - $\pm$ studying
  - $\pm$ leave home / local area
  - $-$ not earning / gaining experience
  - $-$ stress
- life after university:
  - $+$ better job / career
  - $\pm$ graduate “identity”
  - $-$ debt

Utility of attending university or working

Different factors enter utility function *additively*

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

Utility of attending university or working

Different factors enter utility function *additively*

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

- Y_S^{ea} is earnings expectations

Utility of attending university or working

Different factors enter utility function *additively*

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

- 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*

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

- 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*

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

- 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

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

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:

$$S_i \equiv \mathbb{1}\{\alpha(\underbrace{Y_{1i}^{ea} - Y_{0i}^{ea}}_{\text{outcomes}}) - \underbrace{(\theta_{1i} - \theta_{0i})'}_{\text{"costs"}}(-\gamma) + \epsilon_{1i} - \epsilon_{0i} > 0\} \qquad (2)$$

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:

$$S_i \equiv \mathbb{1}\{\alpha(\underbrace{Y_{1i}^{ea} - Y_{0i}^{ea}}_{\text{outcomes}}) - \underbrace{(\theta_{1i} - \theta_{0i})'}_{\text{"costs"}}(-\gamma) + \epsilon_{1i} - \epsilon_{0i} > 0\} \qquad (2)$$

- probability of attending university conditional on Y_s^{ea} and θ_{Si} is

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

Cohort studies

Next Steps (main data source)

- UK cohort study run by CLS at UCL

Cohort studies

Next Steps (main data source)

- UK cohort study run by CLS at UCL
- Follows 15,770 people born in 1989/1990 (“1990 cohort”)

Cohort studies

Next Steps (main data source)

- UK cohort study run by CLS at UCL
- Follows 15,770 people born in 1989/1990 (“1990 cohort”)
- Surveys at age 16 and 25

Cohort studies

Next Steps (main data source)

- UK cohort study run by CLS at UCL
- Follows 15,770 people born in 1989/1990 (“1990 cohort”)
- Surveys at age 16 and 25
  (16) Detailed information on schooling, family background

Cohort studies

Next Steps (main data source)

- UK cohort study run by CLS at UCL
- Follows 15,770 people born in 1989/1990 (“1990 cohort”)
- Surveys at age 16 and 25
  (16) Detailed information on schooling, family background
  (16) Subjective questions about university, life and future

Cohort studies

Next Steps (main data source)

- UK cohort study run by CLS at UCL
- Follows 15,770 people born in 1989/1990 (“1990 cohort”)
- 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

Cohort studies

Next Steps (main data source)

- UK cohort study run by CLS at UCL
- Follows 15,770 people born in 1989/1990 (“1990 cohort”)
- 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 ongoing

Cohort studies

Next Steps (main data source)

- UK cohort study run by CLS at UCL
- Follows 15,770 people born in 1989/1990 (“1990 cohort”)
- 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 ongoing

British Cohort Study (comparison over time)

- Similar study also run by CLS
- Cohort born in one week in April 1970
- Waves every 4 years since birth

Next Steps (1990)

Sample design

- 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

Key feature: subjective, open-ended questions about advantages and disadvantages of attending university

- subjective: captures which aspects **each** student considers
- open-ended: students can mention anything, not leading q’s
- questions specifically about attending university
- Similar responses identified and harmonised by survey designers
- BUT only asked to some students

Next Steps (1990)

Sample design

- 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

Key feature: subjective, open-ended questions about advantages and disadvantages of attending university

- subjective: captures which aspects **each** student considers
- open-ended: students can mention anything, not leading q’s
- questions specifically about attending university
- Similar responses identified and harmonised by survey designers
- BUT only asked to some students

Next Steps (1990)

Sample design

- 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

Key feature: subjective, open-ended questions about advantages and disadvantages of attending university

- subjective: captures which aspects **each** student considers
- open-ended: students can mention anything, not leading q’s
- questions specifically about attending university
- Similar responses identified and harmonised by survey designers
- BUT only asked to some students

Next Steps (1990)

Sample design

- 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

Key feature: subjective, open-ended questions about advantages and disadvantages of attending university

- subjective: captures which aspects **each** student considers
- open-ended: students can mention anything, not leading q’s
- questions specifically about attending university
- Similar responses identified and harmonised by survey designers
- BUT only asked to some students

Next Steps (1990)

Sample design

- 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

Key feature: subjective, open-ended questions about advantages and disadvantages of attending university

- subjective: captures which aspects **each** student considers
- open-ended: students can mention anything, not leading q’s
- questions specifically about attending university
- Similar responses identified and harmonised by survey designers
- BUT only asked to some students

Next Steps (1990)

Sample design

- 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

Key feature: subjective, open-ended questions about advantages and disadvantages of attending university

- subjective: captures which aspects **each** student considers
- open-ended: students can mention anything, not leading q’s
- questions specifically about attending university
- Similar responses identified and harmonised by survey designers
- BUT only asked to some students

Next Steps (1990)

Sample design

- 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

Key feature: subjective, open-ended questions about advantages and disadvantages of attending university

- subjective: captures which aspects **each** student considers
- open-ended: students can mention anything, not leading q’s
- questions specifically about attending university
- Similar responses identified and harmonised by survey designers
- BUT only asked to some students

Next Steps (1990)

Sample design

- 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

Key feature: subjective, open-ended questions about advantages and disadvantages of attending university

- subjective: captures which aspects **each** student considers
- open-ended: students can mention anything, not leading q’s
- questions specifically about attending university
- Similar responses identified and harmonised by survey designers
- BUT only asked to some students

Next Steps (1990)

Sample design

- 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

Key feature: subjective, open-ended questions about advantages and disadvantages of attending university

- subjective: captures which aspects **each** student considers
- open-ended: students can mention anything, not leading q’s
- questions specifically about attending university
- Similar responses identified and harmonised by survey designers
- BUT only asked to some students

Subsample and descriptive statistics

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
*Russell group**†	0.26	0.28
Employed*	0.83	0.87
Wage (GBP)*‡	393	424

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

What are these open-ended questions? And what were the responses?

Subsample and descriptive statistics

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
*Russell group**†	0.26	0.28
Employed*	0.83	0.87
Wage (GBP)*‡	393	424

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

What are these open-ended questions? And what were the responses?

Subsample and descriptive statistics

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
*Russell group**†	0.26	0.28
Employed*	0.83	0.87
Wage (GBP)*‡	393	424

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

What are these open-ended questions? And what were the responses?

Advantages of attending university

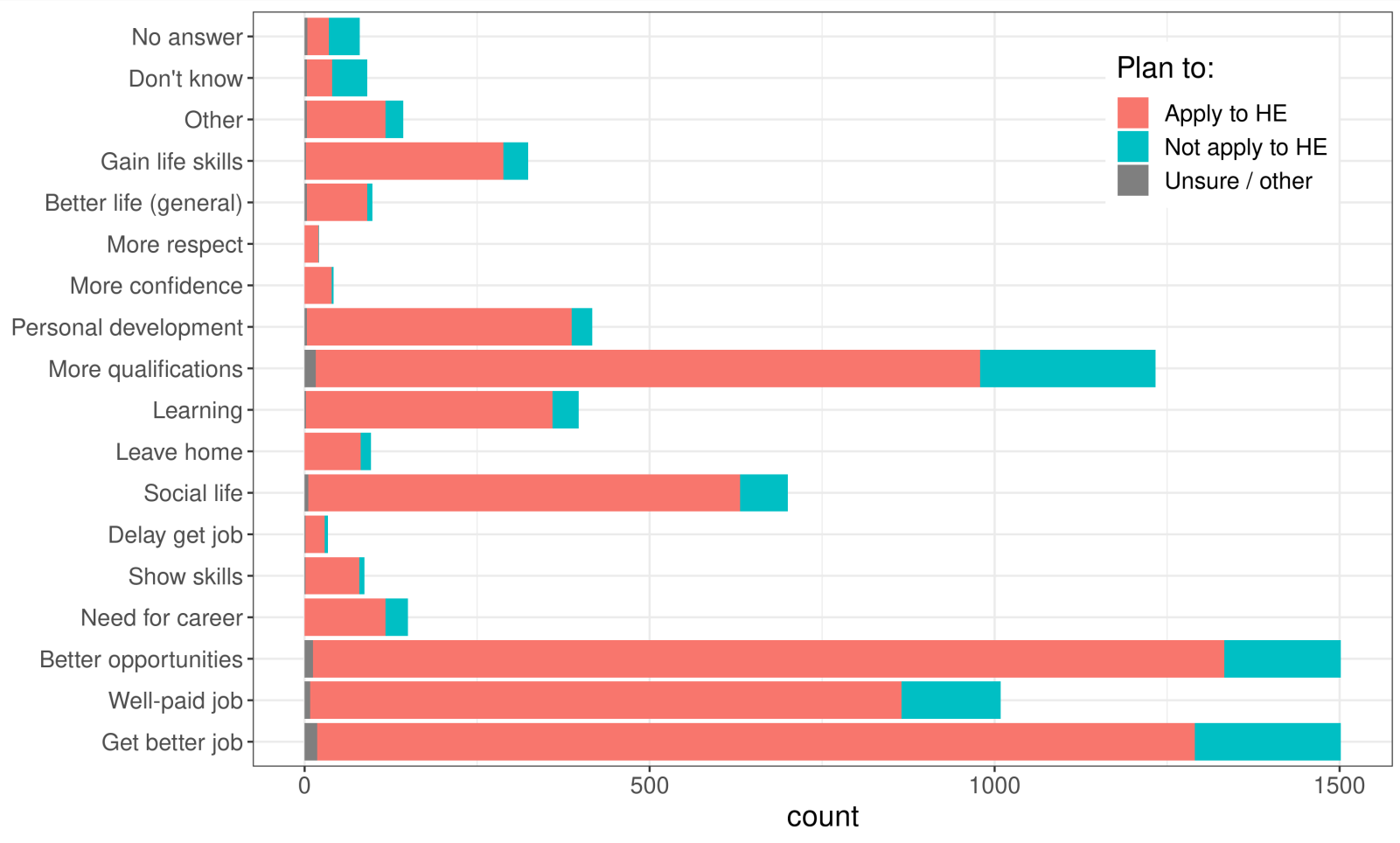
What do you think the advantages, if any, might be for SOMEONE of going to university to study for a degree?

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?

main reasons



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

Disadvantages of attending university

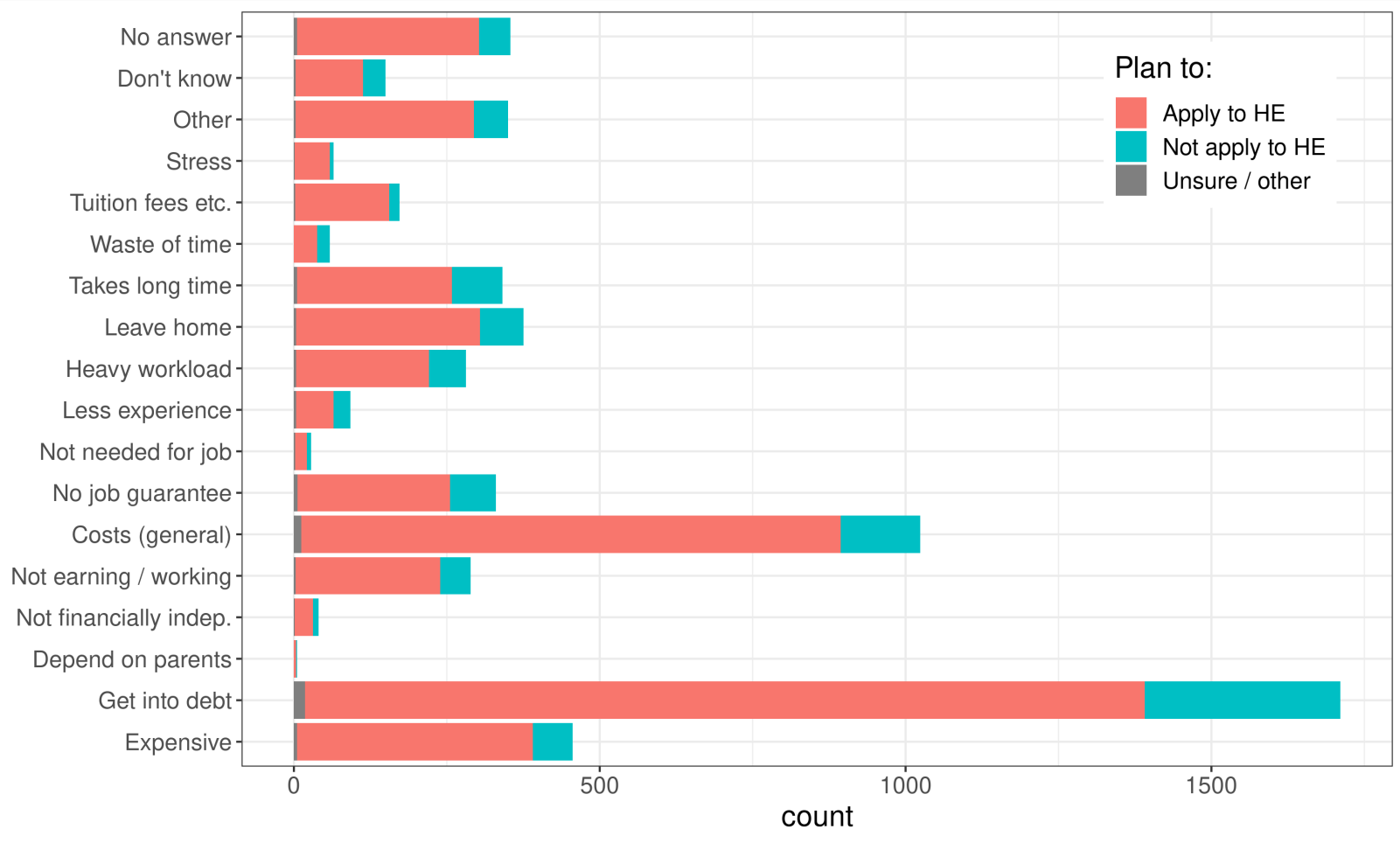
What do you think the disadvantages, if any, might be for someone of going to university to study for a degree?

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?

▸ main reasons



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

Earnings expectations, Y_S^{ea}

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

Earnings expectations, Y_S^{ea}

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

Earnings expectations, Y_S^{ea}

- 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 at age 25 given choice S

wage distribution

Earnings expectations, Y_S^{ea}

- 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 at age 25 given choice S

wage distribution

- Key assumptions

Earnings expectations, Y_S^{ea}

- 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 at age 25 given choice S

wage distribution

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

Earnings expectations, Y_S^{ea}

- 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 at age 25 given choice S

wage distribution

- Key assumptions
  - we observe all relevant variables, X [I plan to test/relax this]
  - log-linear wage equation [standard, since Mincer (1974)]

Earnings expectations, Y_S^{ea}

- 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 at age 25 given choice S

wage distribution

- Key assumptions
  - we observe all relevant variables, X [I plan to test/relax this]
  - log-linear wage equation [standard, since Mincer (1974)]
  - rational expectations [standard, e.g. Heckman et al. (2006)]

Psychic costs, θ_S

- **Other factors in the utility function: Heckman's "psychic costs"**

Psychic costs, θ_S

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

Psychic costs, θ_S

- Other factors in the utility function: Heckman's "psychic costs"
- Usually "measured" with family background or residual term
- **I use responses to the open-ended questions**

Psychic costs, θ_S

- Other factors in the utility function: Heckman's "psychic costs"
- Usually "measured" with family background or residual term
- **I use responses to the open-ended questions**
  - each harmonised "response" is a component of $\theta_1 - \theta_0$

Psychic costs, θ_S

- Other factors in the utility function: Heckman's "psychic costs"
- Usually "measured" with family background or residual term
- **I use responses to the open-ended questions**
  - each harmonised "response" is a component of $\theta_1 - \theta_0$
  - value 1 if mentioned by student; 0 otherwise

Psychic costs, θ_S

- Other factors in the utility function: Heckman's "psychic costs"
- Usually "measured" with family background or residual term
- **I use responses to the open-ended questions**
  - each harmonised "response" is a component of $\theta_1 - \theta_0$
  - value 1 if mentioned by student; 0 otherwise
  - also include family background and other characteristics

α and γ

- Recall the conditional probability of attending university

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

α and γ

- Recall the conditional probability of attending university

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

- α and γ are identified by an assumption on the errors, given $Y_{1i}^{ea} - Y_{0i}^{ea}$ and $\theta_{1i} - \theta_{0i}$

α and γ

- Recall the conditional probability of attending university

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

- α and γ are identified by an assumption on the errors, given $Y_{1i}^{ea} - Y_{0i}^{ea}$ and $\theta_{1i} - \theta_{0i}$
- assume errors follow an extreme-value distribution

α and γ

- Recall the conditional probability of attending university

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

- α and γ are identified by an assumption on the errors, given $Y_{1i}^{ea} - Y_{0i}^{ea}$ and $\theta_{1i} - \theta_{0i}$
- assume errors follow an extreme-value distribution
- then $(\epsilon_0 - \epsilon_1) \sim$ Logistic and

$$\Pr(S = 1 | Y_s^{ea}, \theta_{Si}) = \left(1 + e^{-(\alpha(Y_{1i}^{ea} - Y_{0i}^{ea}) + (\theta_{1i} - \theta_{0i})'\gamma)}\right)^{-1}$$

α and γ

- Recall the conditional probability of attending university

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

- α and γ are identified by an assumption on the errors, given $Y_{1i}^{ea} - Y_{0i}^{ea}$ and $\theta_{1i} - \theta_{0i}$
- assume errors follow an extreme-value distribution
- then $(\epsilon_0 - \epsilon_1) \sim$ Logistic and

$$\Pr(S = 1 | Y_s^{ea}, \theta_{Si}) = \left(1 + e^{-(\alpha(Y_{1i}^{ea} - Y_{0i}^{ea}) + (\theta_{1i} - \theta_{0i})'\gamma)}\right)^{-1}$$

- standard in discrete-choice models

Estimation

Parameter estimates $(\widehat{\alpha}, \widehat{\beta_k}, \widehat{\gamma})$

- $\epsilon_0 - \epsilon_1 \sim$ *Logit* allows us to obtain estimates $\widehat{\alpha}, \widehat{\gamma}$ by logit regression, given Y_S^{ea} and θ_S.

Estimation

Parameter estimates $(\widehat{\alpha}, \widehat{\beta_k}, \widehat{\gamma})$

- $\epsilon_0 - \epsilon_1 \sim$ *Logit* allows us to obtain estimates $\widehat{\alpha}, \widehat{\gamma}$ by logit regression, given Y_S^{ea} and θ_S.
- estimate earnings as $\widehat{Y_S^{ea}} = X' \widehat{\beta_S}$, from OLS of Y_S on X

Estimation

Parameter estimates $(\widehat{\alpha}, \widehat{\beta_k}, \widehat{\gamma})$

- $\epsilon_0 - \epsilon_1 \sim$ *Logit* allows us to obtain estimates $\widehat{\alpha}, \widehat{\gamma}$ by logit regression, given Y_S^{ea} and θ_S.
- estimate earnings as $\widehat{Y_S^{ea}} = X' \widehat{\beta_S}$, from OLS of Y_S on X

Choosing components of X and θ

- X contains characteristics that affect the HE decision through earnings expectations at age 16

Estimation

Parameter estimates $(\widehat{\alpha}, \widehat{\beta_k}, \widehat{\gamma})$

- $\epsilon_0 - \epsilon_1 \sim$ *Logit* allows us to obtain estimates $\widehat{\alpha}, \widehat{\gamma}$ by logit regression, given Y_S^{ea} and θ_S.
- estimate earnings as $\widehat{Y_S^{ea}} = X' \widehat{\beta_S}$, from OLS of Y_S on X

Choosing components of X and θ

- X contains characteristics that affect the HE decision through earnings expectations at age 16
  - parents': occupations, ethnicity group, education, income

Estimation

Parameter estimates $(\widehat{\alpha}, \widehat{\beta_k}, \widehat{\gamma})$

- $\epsilon_0 - \epsilon_1 \sim$ *Logit* allows us to obtain estimates $\widehat{\alpha}, \widehat{\gamma}$ by logit regression, given Y_S^{ea} and θ_S.
- estimate earnings as $\widehat{Y_S^{ea}} = X' \widehat{\beta_S}$, from OLS of Y_S on X

Choosing components of X and θ

- X contains 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

Estimation

Parameter estimates $(\widehat{\alpha}, \widehat{\beta_k}, \widehat{\gamma})$

- $\epsilon_0 - \epsilon_1 \sim$ *Logit* allows us to obtain estimates $\widehat{\alpha}, \widehat{\gamma}$ by logit regression, given Y_S^{ea} and θ_S.
- estimate earnings as $\widehat{Y_S^{ea}} = X' \widehat{\beta_S}$, from OLS of Y_S on X

Choosing components of X and θ

- X contains 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
- recall: θ contains responses to open-ended questions (plus components of X)

Comparing earnings with other factors

- Compare relative importance of earnings versus other factors

Comparing earnings with other factors

- Compare relative importance of earnings versus other factors
- Strategy:

Comparing earnings with other factors

- Compare relative importance of earnings versus other factors
- Strategy:
  - Estimate parameters $\alpha, \beta_k, \Delta\delta$

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 expected graduate-wage premium, and non-earnings factors.

Comparing earnings with other factors

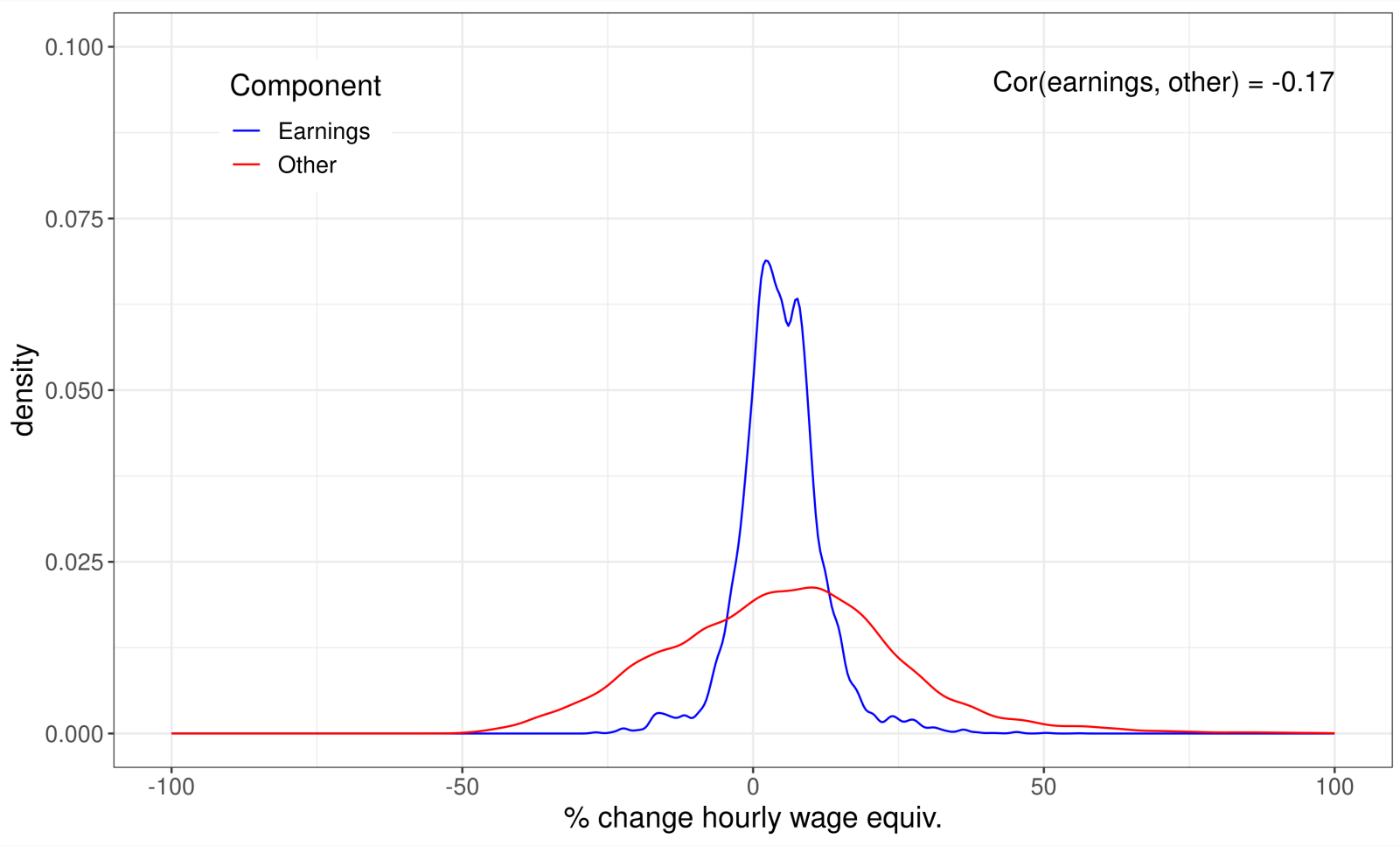
- Compare relative importance of earnings versus other factors
- Strategy:
  - Estimate parameters α, β_k, $\Delta\delta$
  - Use estimates and sample covariates to estimate distributions of expected graduate-wage premium, and non-earnings factors.
  - Transform these factors so they are equivalent to $\%\Delta$ in earnings

Comparing earnings with other factors

- Compare relative importance of earnings versus other factors
- Strategy:
  - Estimate parameters α, β_k, $\Delta\delta$
  - Use estimates and sample covariates to estimate distributions of expected graduate-wage premium, and non-earnings factors.
  - Transform these factors so they are equivalent to $\%\Delta$ in earnings
  - Plot the transformed distributions

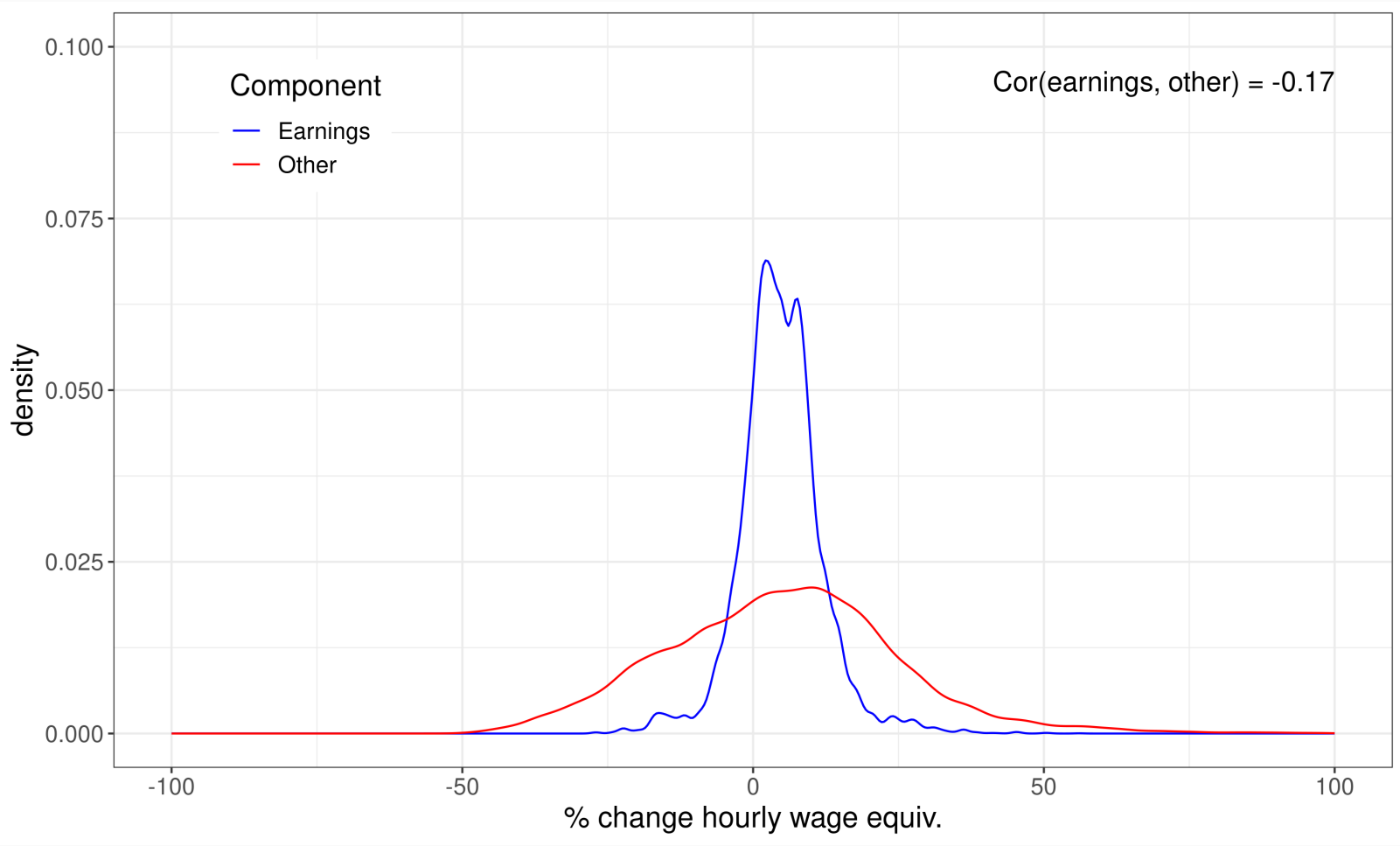
Full sample (1990): earnings vs other factors

- Both distributions similarly located, with positive means (4.8% vs 4.9%)



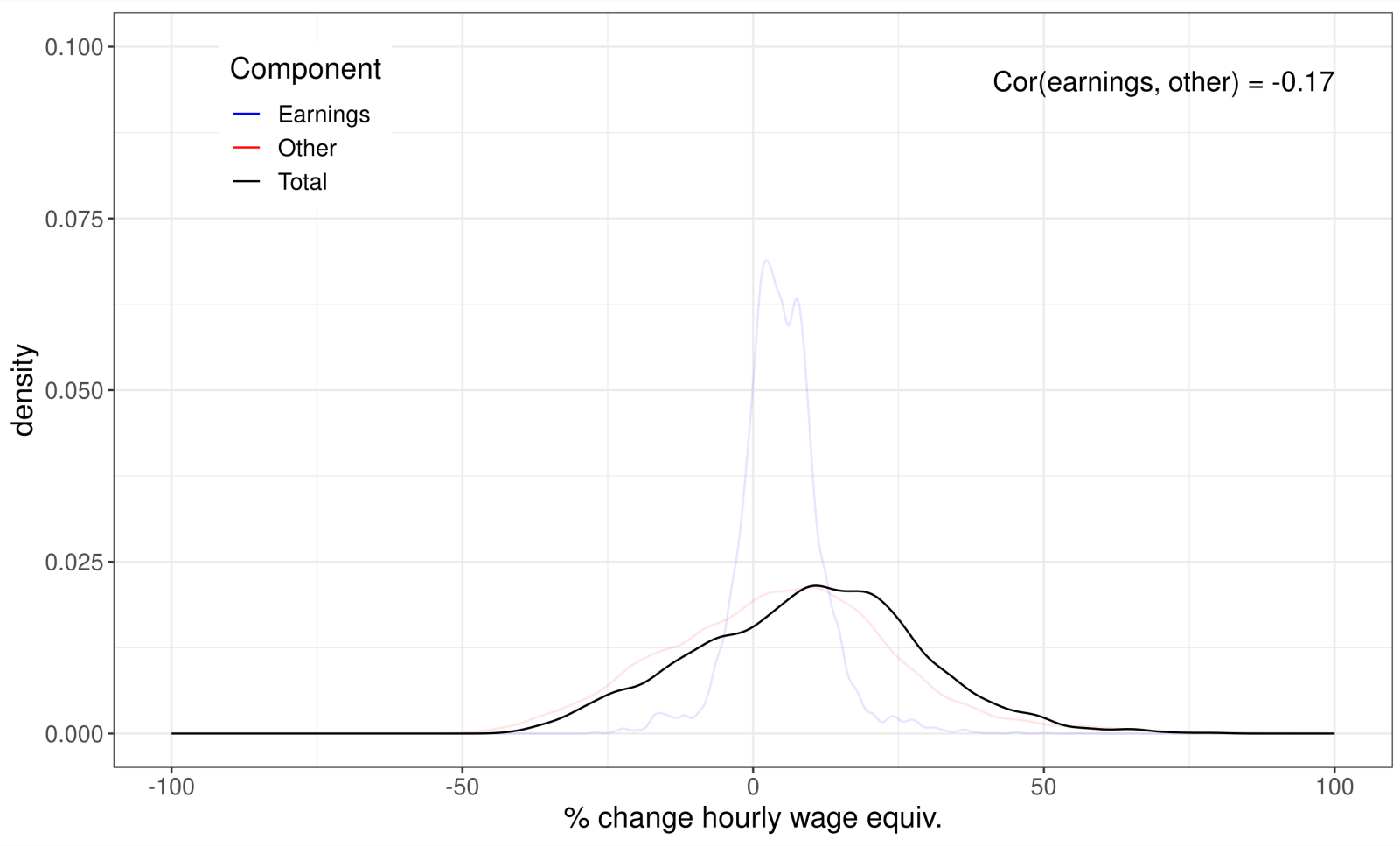
Full sample (1990): earnings vs other factors

- Both distributions similarly located, with positive means (4.8% vs 4.9%)
- Earnings distribution has a much lower std dev. (7.1% vs 20%)



Full sample (1990): earnings vs other factors

- Both distributions similarly located, with positive means (4.8% vs 4.9%)
- Earnings distribution has a much lower std dev. (7.1% vs 20%)
  $\Rightarrow$ Main determinant of decision is other factors: SD(total) = 19.2%



Counterfactual exercise

Variation in factors: suppose everyone has...

- ...non-pecuniary factors equal to first quartile

  $\rightarrow$ only 24.0% have high-enough earnings expectations to choose to attend university
- ...non-pecuniary factors equal to third quartile

  $\rightarrow$ now 99.5% attend university
- earnings expectations equal to first quartile

  $\rightarrow$ 61.9% have high-enough non-pec. factors to attend university
- earnings expectations equal to third quartile

  $\rightarrow$ still only 74.8% would attend university

Counterfactual exercise

Variation in factors: suppose everyone has...

- ...non-pecuniary factors equal to first quartile

  $\rightarrow$ only 24.0% have high-enough earnings expectations to choose to attend university
- ...non-pecuniary factors equal to third quartile

  $\rightarrow$ now 99.5% attend university
- earnings expectations equal to first quartile

  $\rightarrow$ 61.9% have high-enough non-pec. factors to attend university
- earnings expectations equal to third quartile

  $\rightarrow$ still only 74.8% would attend university

Counterfactual exercise

Variation in factors: suppose everyone has...

- ...non-pecuniary factors equal to first quartile

  $\rightarrow$ only 24.0% have high-enough earnings expectations to choose to attend university
- ...non-pecuniary factors equal to third quartile

  $\rightarrow$ now 99.5% attend university
- earnings expectations equal to first quartile

  $\rightarrow$ 61.9% have high-enough non-pec. factors to attend university
- earnings expectations equal to third quartile

  $\rightarrow$ still only 74.8% would attend university

Counterfactual exercise

Variation in factors: suppose everyone has...

- ...non-pecuniary factors equal to first quartile

  $\rightarrow$ only 24.0% have high-enough earnings expectations to choose to attend university
- ...non-pecuniary factors equal to third quartile

  $\rightarrow$ now 99.5% attend university
- earnings expectations equal to first quartile

  $\rightarrow$ 61.9% have high-enough non-pec. factors to attend university
- earnings expectations equal to third quartile

  $\rightarrow$ still only 74.8% would attend university

Counterfactual exercise

Variation in factors: suppose everyone has...

- ...non-pecuniary factors equal to first quartile

  $\rightarrow$ only 24.0% have high-enough earnings expectations to choose to attend university
- ...non-pecuniary factors equal to third quartile

  $\rightarrow$ now 99.5% attend university
- earnings expectations equal to first quartile

  $\rightarrow$ 61.9% have high-enough non-pec. factors to attend university
- earnings expectations equal to third quartile

  $\rightarrow$ still only 74.8% would attend university

Counterfactual exercise

Variation in factors: suppose everyone has...

- ...non-pecuniary factors equal to first quartile

  $\rightarrow$ only 24.0% have high-enough earnings expectations to choose to attend university
- ...non-pecuniary factors equal to third quartile

  $\rightarrow$ now 99.5% attend university
- earnings expectations equal to first quartile

  $\rightarrow$ 61.9% have high-enough non-pec. factors to attend university
- earnings expectations equal to third quartile

  $\rightarrow$ still only 74.8% would attend university

Counterfactual exercise

Variation in factors: suppose everyone has...

- ...non-pecuniary factors equal to first quartile

  $\rightarrow$ only 24.0% have high-enough earnings expectations to choose to attend university
- ...non-pecuniary factors equal to third quartile

  $\rightarrow$ now 99.5% attend university
- earnings expectations equal to first quartile

  $\rightarrow$ 61.9% have high-enough non-pec. factors to attend university
- earnings expectations equal to third quartile

  $\rightarrow$ still only 74.8% would attend university

Counterfactual exercise

Variation in factors: suppose everyone has...

- ...non-pecuniary factors equal to first quartile

  $\rightarrow$ only 24.0% have high-enough earnings expectations to choose to attend university
- ...non-pecuniary factors equal to third quartile

  $\rightarrow$ now 99.5% attend university
- earnings expectations equal to first quartile

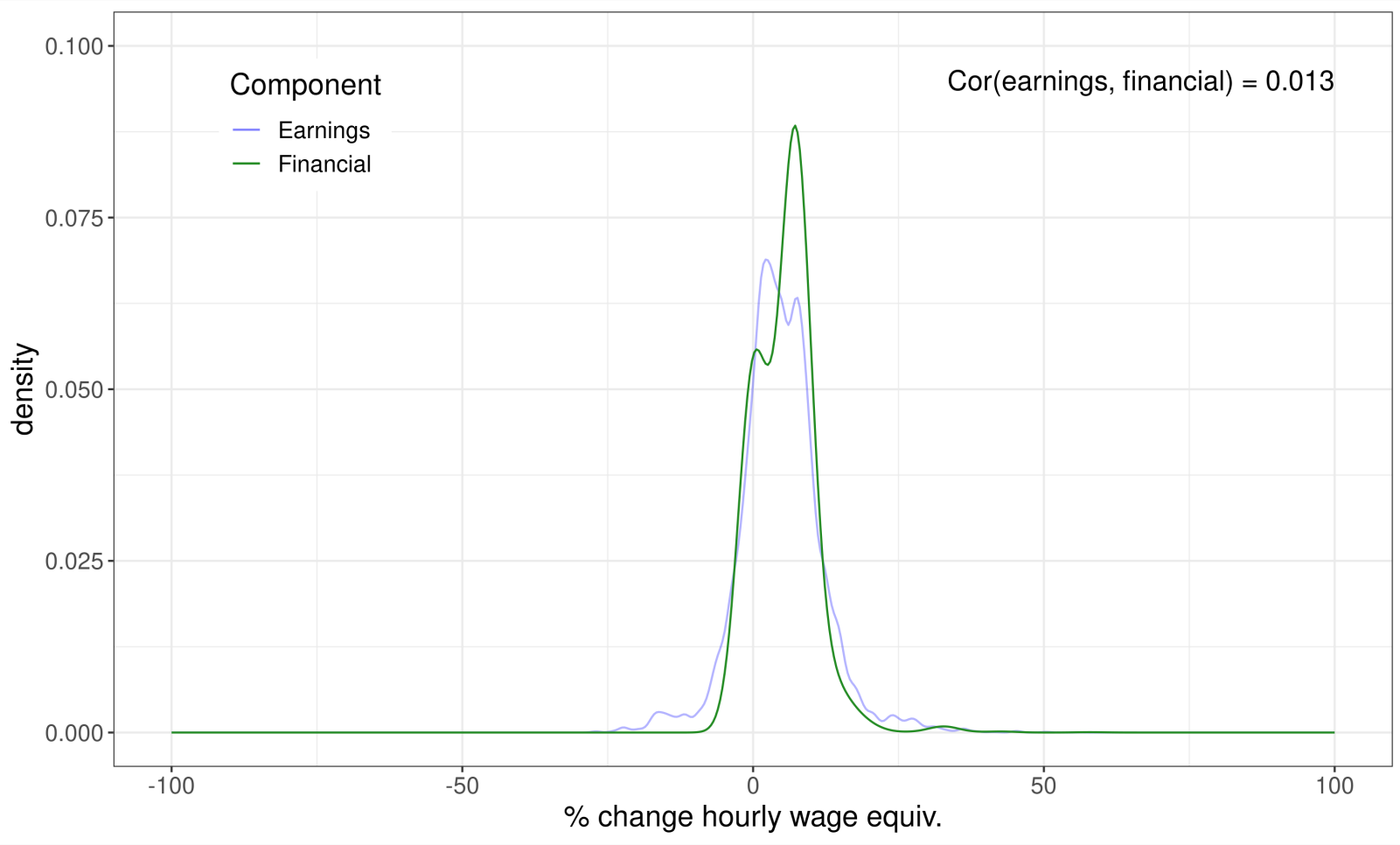
  $\rightarrow$ 61.9% have high-enough non-pec. factors to attend university
- earnings expectations equal to third quartile

  $\rightarrow$ still only 74.8% would attend university

Full sample (1990): splitting θ into financial and other

- Financial factors distribution very similar to earnings expectations
- Mean financial positive (5.4%), mean other now negative (−2.1%)
- SD financial smaller (4.9%) than other (17.8%)

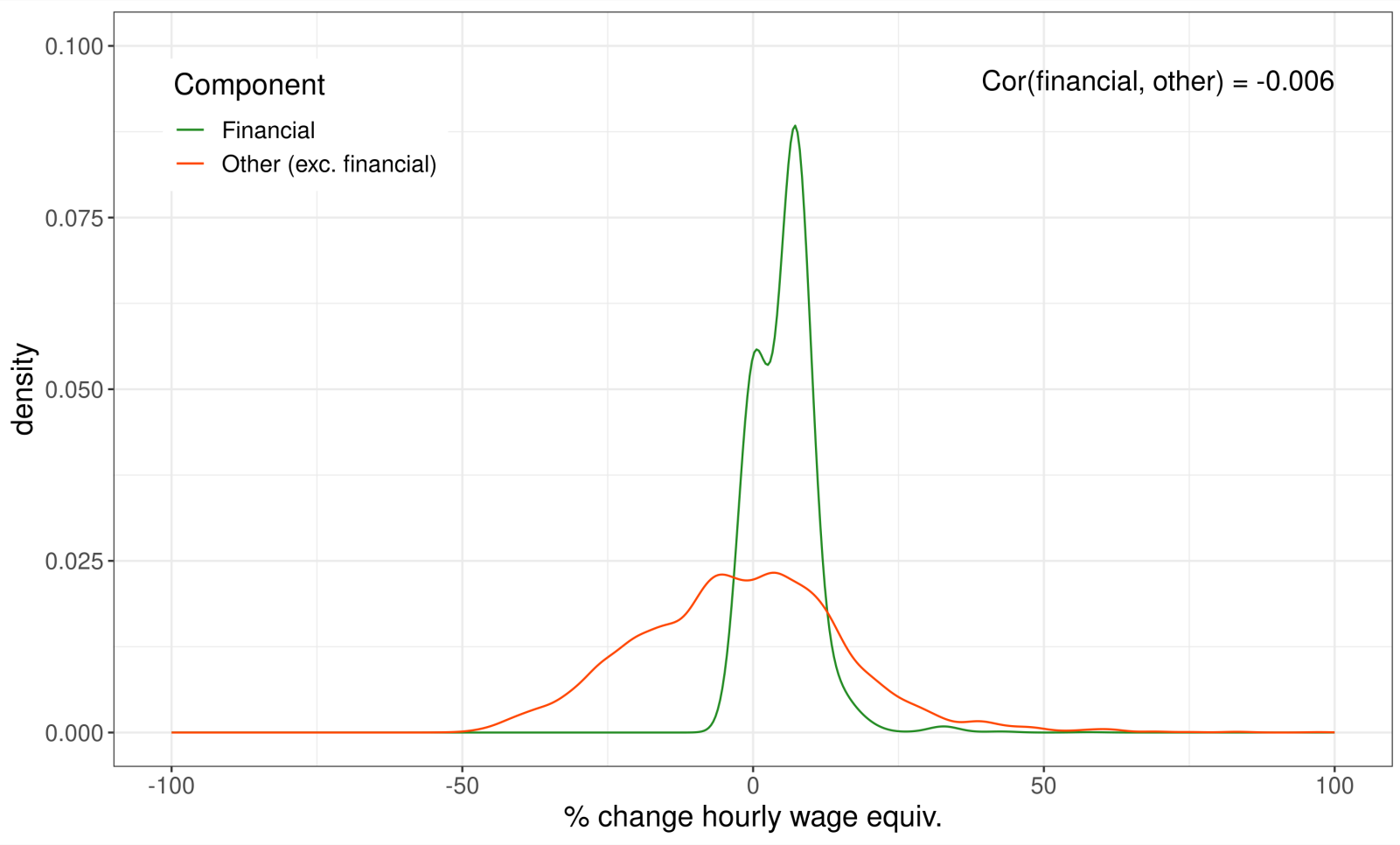
financial factors



Full sample (1990): splitting θ into financial and other

- Financial factors distribution very similar to earnings expectations
- Mean financial positive (5.4%), mean other now negative (-2.1%)
- SD financial smaller (4.9%) than other (17.8%)

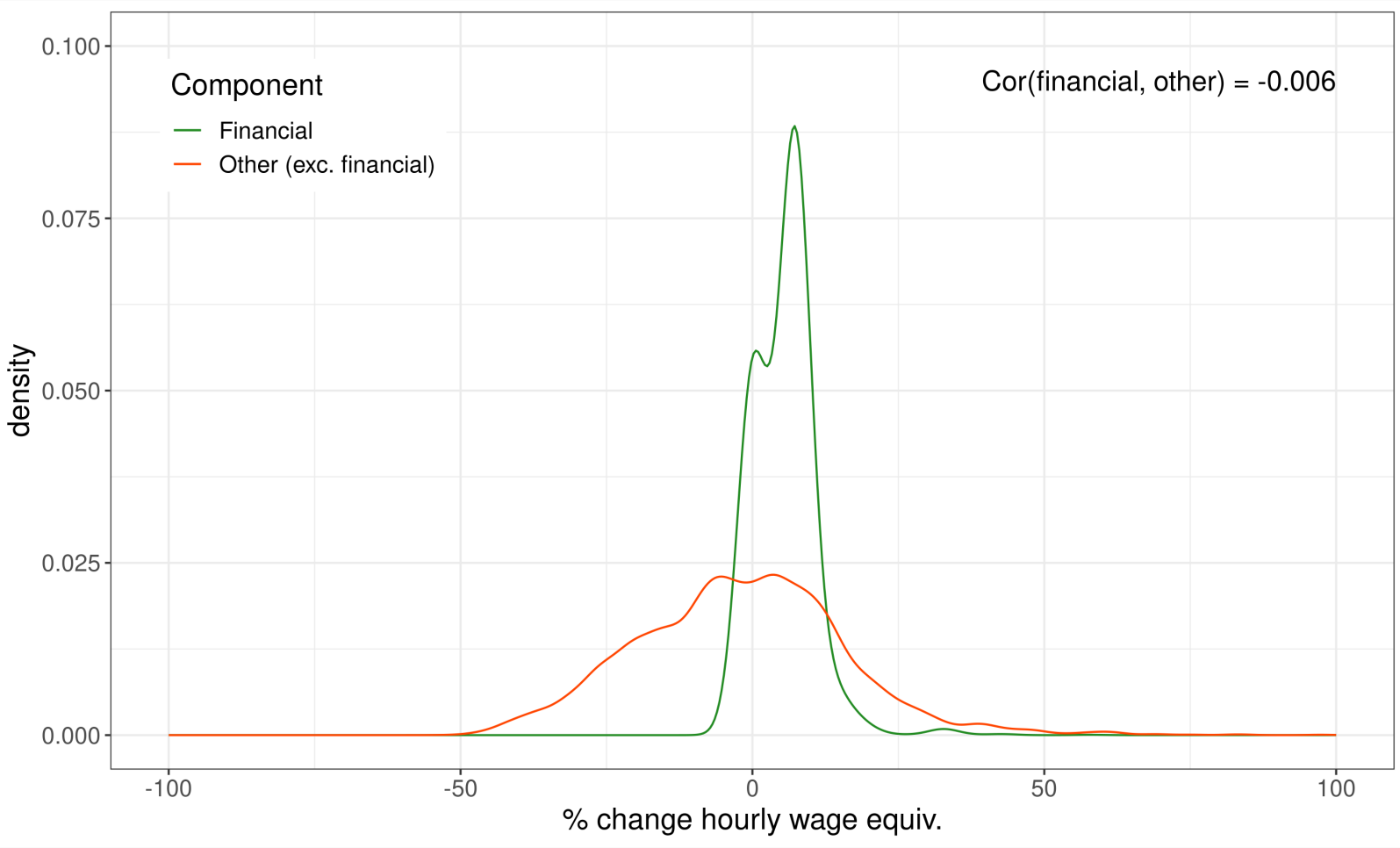
financial factors



Full sample (1990): splitting θ into financial and other

- Financial factors distribution very similar to earnings expectations
- Mean financial positive (5.4%), mean other now negative (−2.1%)
- SD financial smaller (4.9%) than other (17.8%)

financial factors



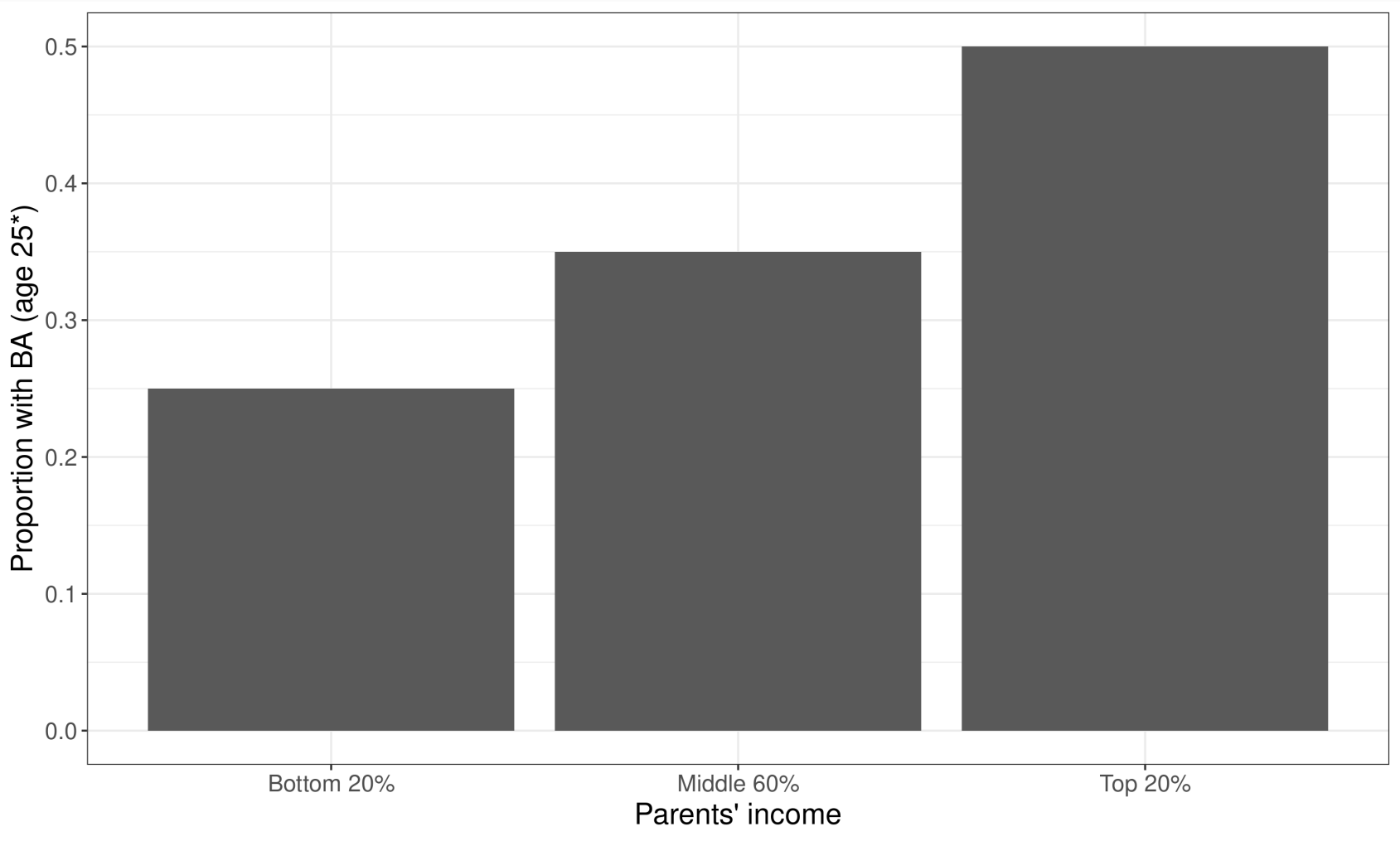
Factors by SES group

- Use same parameter estimates as before
- Split sample by parental earnings at 16 into 3 groups:
  $\rightarrow$ bottom 20%, middle 60%, top 20%
- plot distribution of factors within each group

Why?

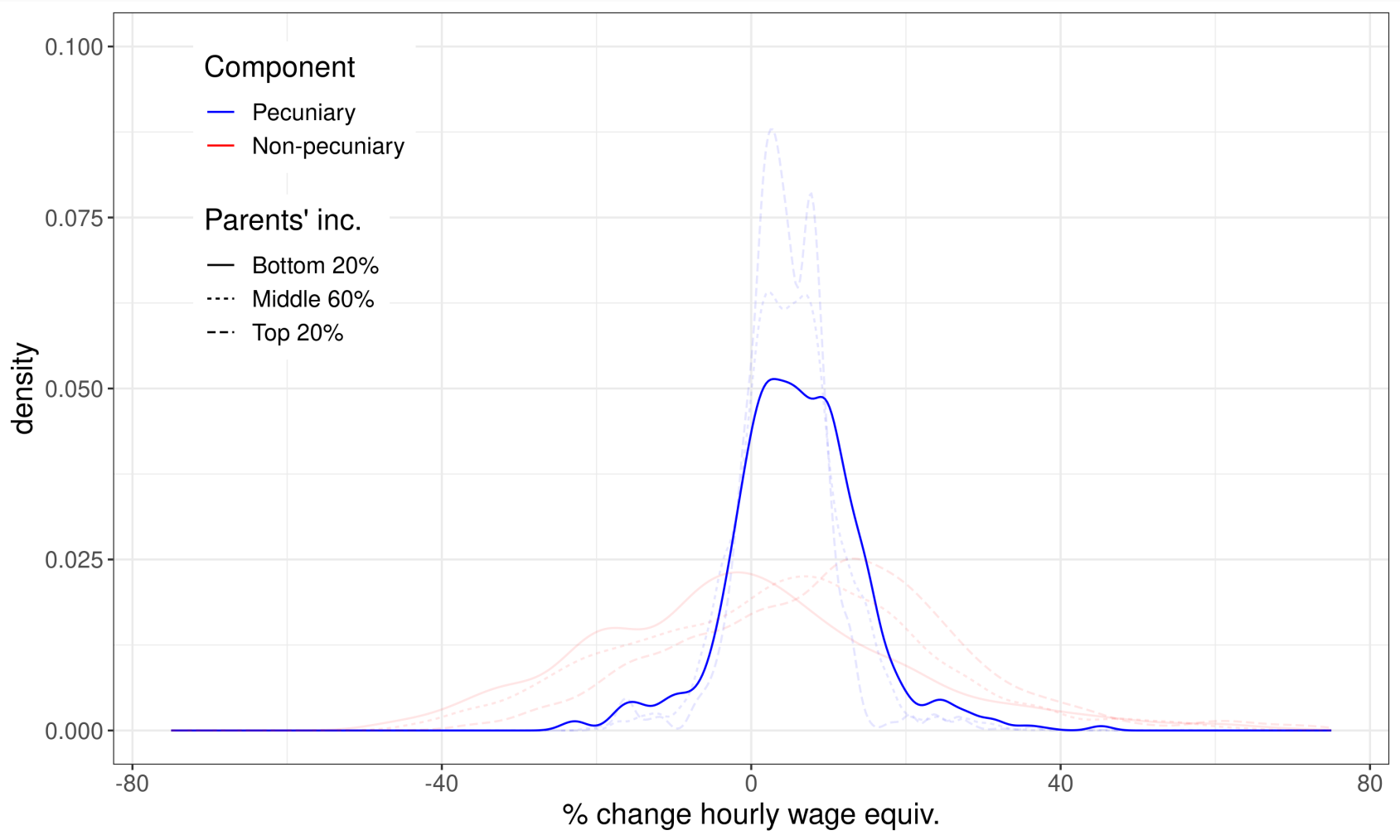
Huge SES-gap in attainment

Children of top 20% by income **twice as likely to hold BA at 25** than children of bottom 20%



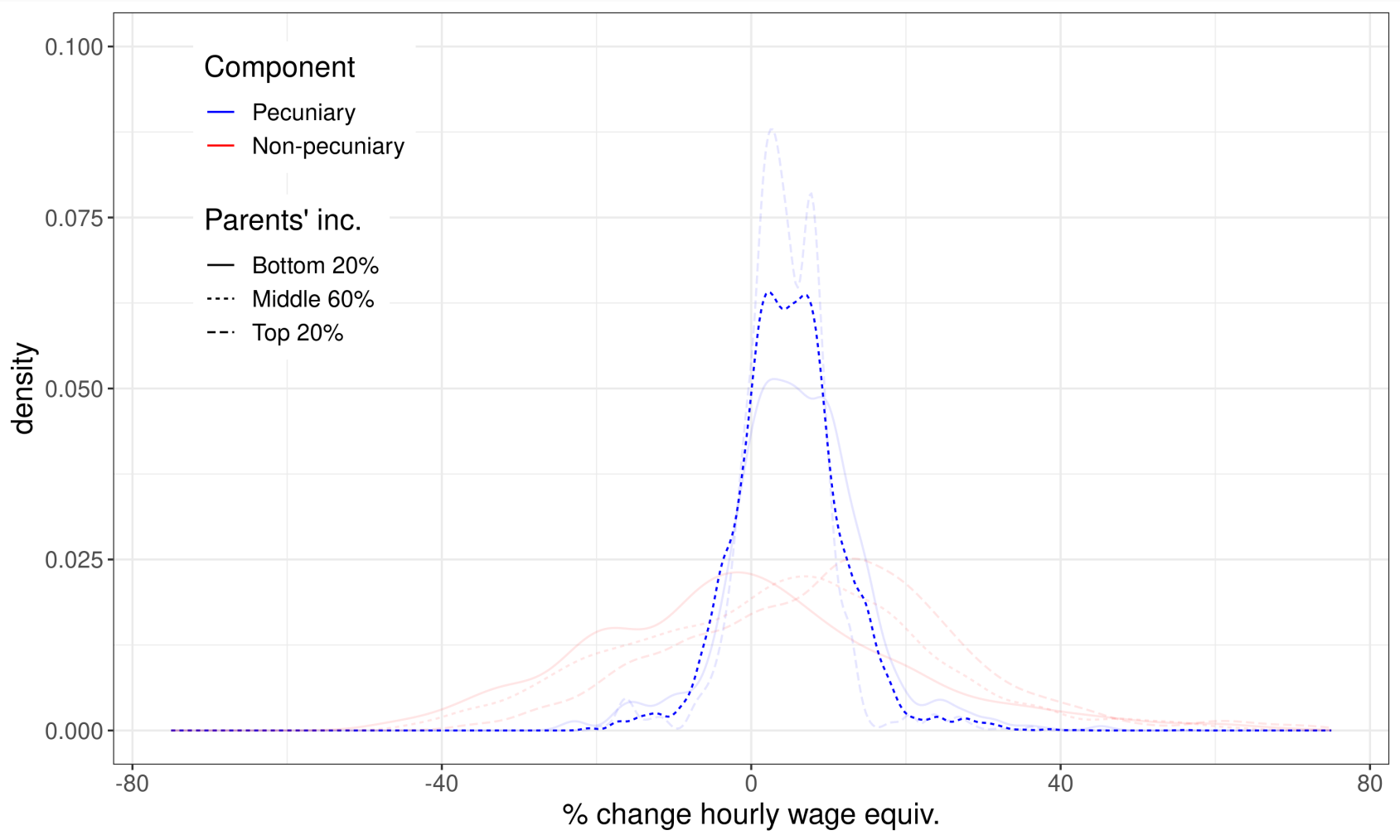
By SES (1990, SES = parents' income at 16)

Parental income	Mean	Variance	Skewness
Bottom 20%	0.147	0.018	0.335
Middle 60%	0.143	0.013	0.271
Top 20%	0.099	0.008	0.650



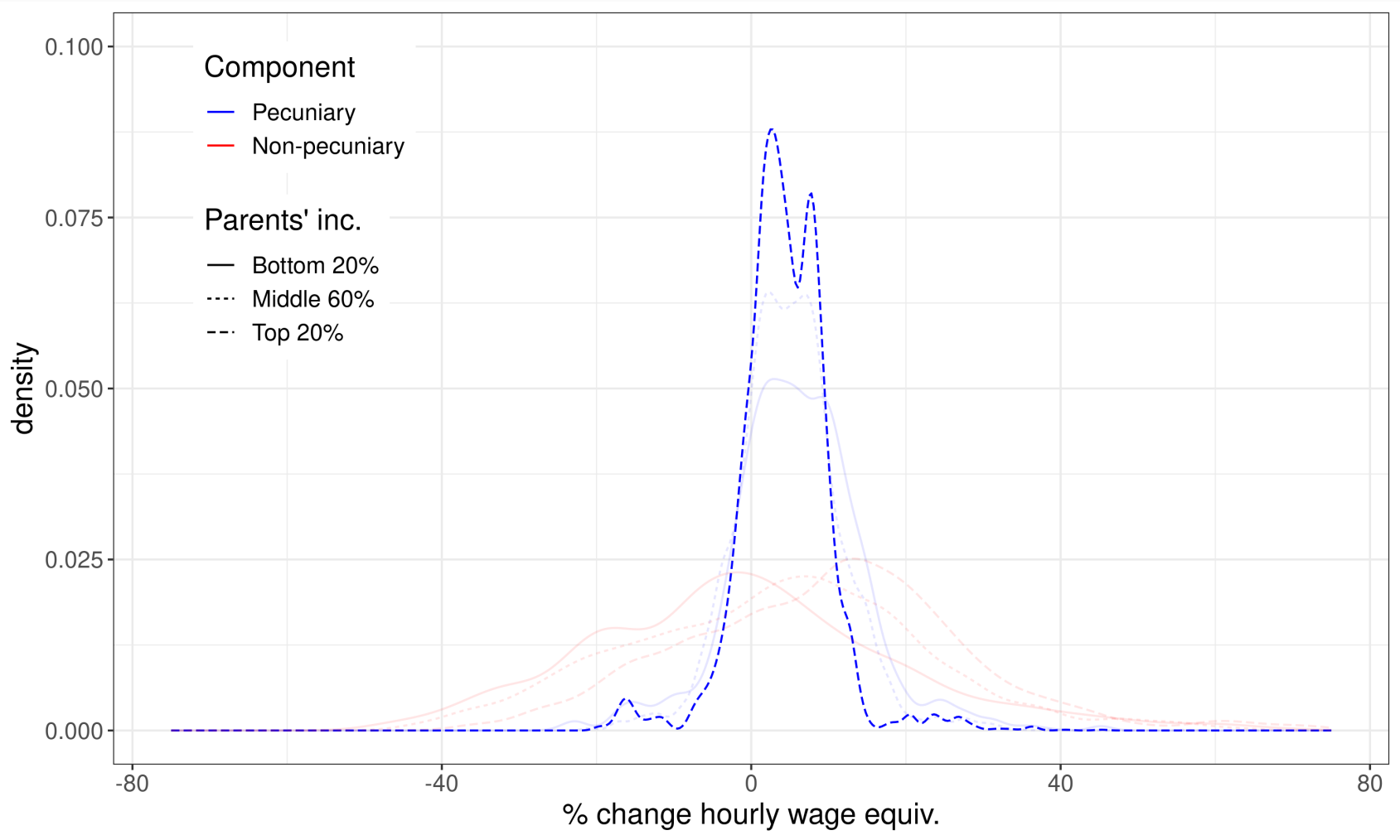
By SES (1990, SES = parents' income at 16)

Parental income	Mean	Variance	Skewness
Bottom 20%	0.147	0.018	0.335
Middle 60%	0.143	0.013	0.271
Top 20%	0.099	0.008	0.650



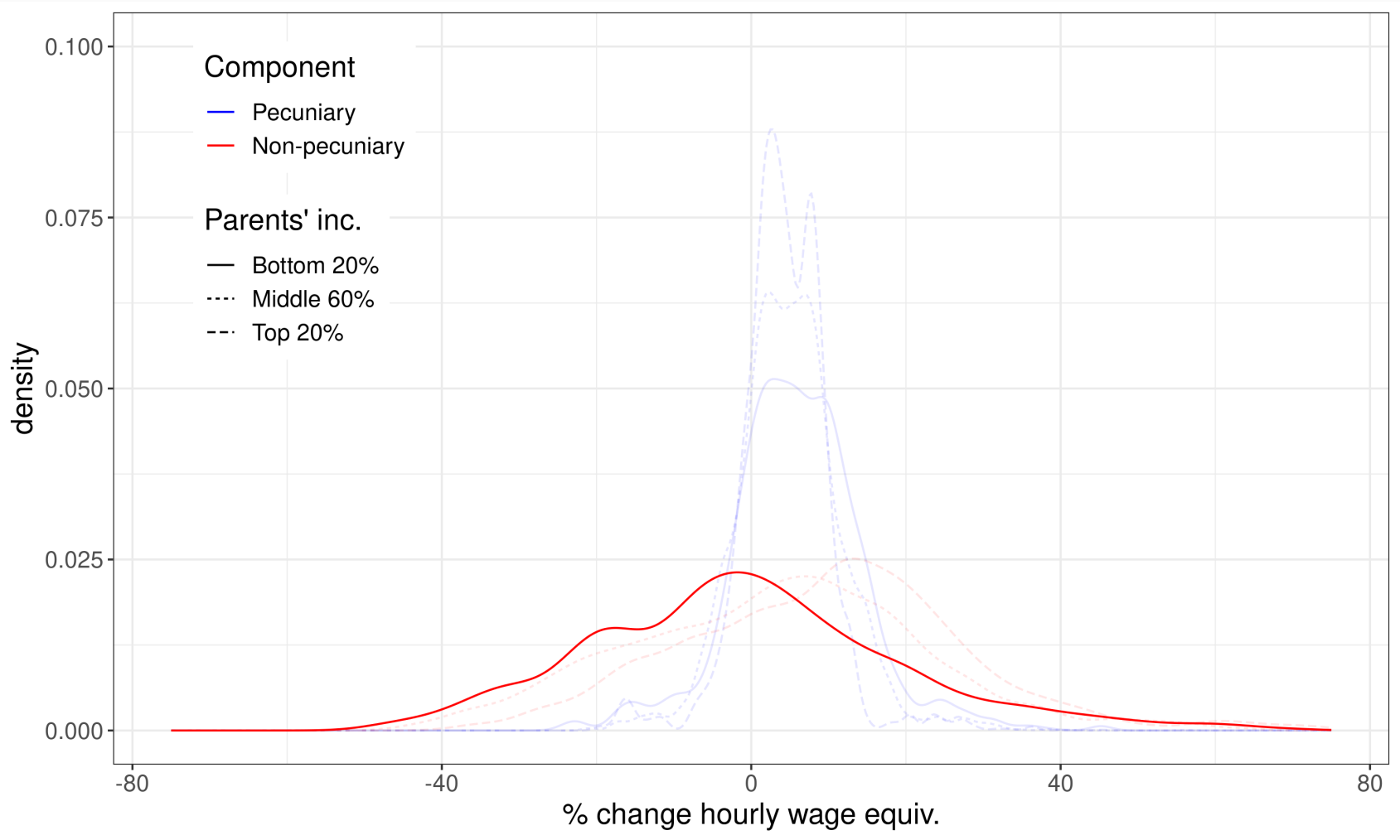
By SES (1990, SES = parents' income at 16)

Parental income	Mean	Variance	Skewness
Bottom 20%	0.147	0.018	0.335
Middle 60%	0.143	0.013	0.271
Top 20%	0.099	0.008	0.650



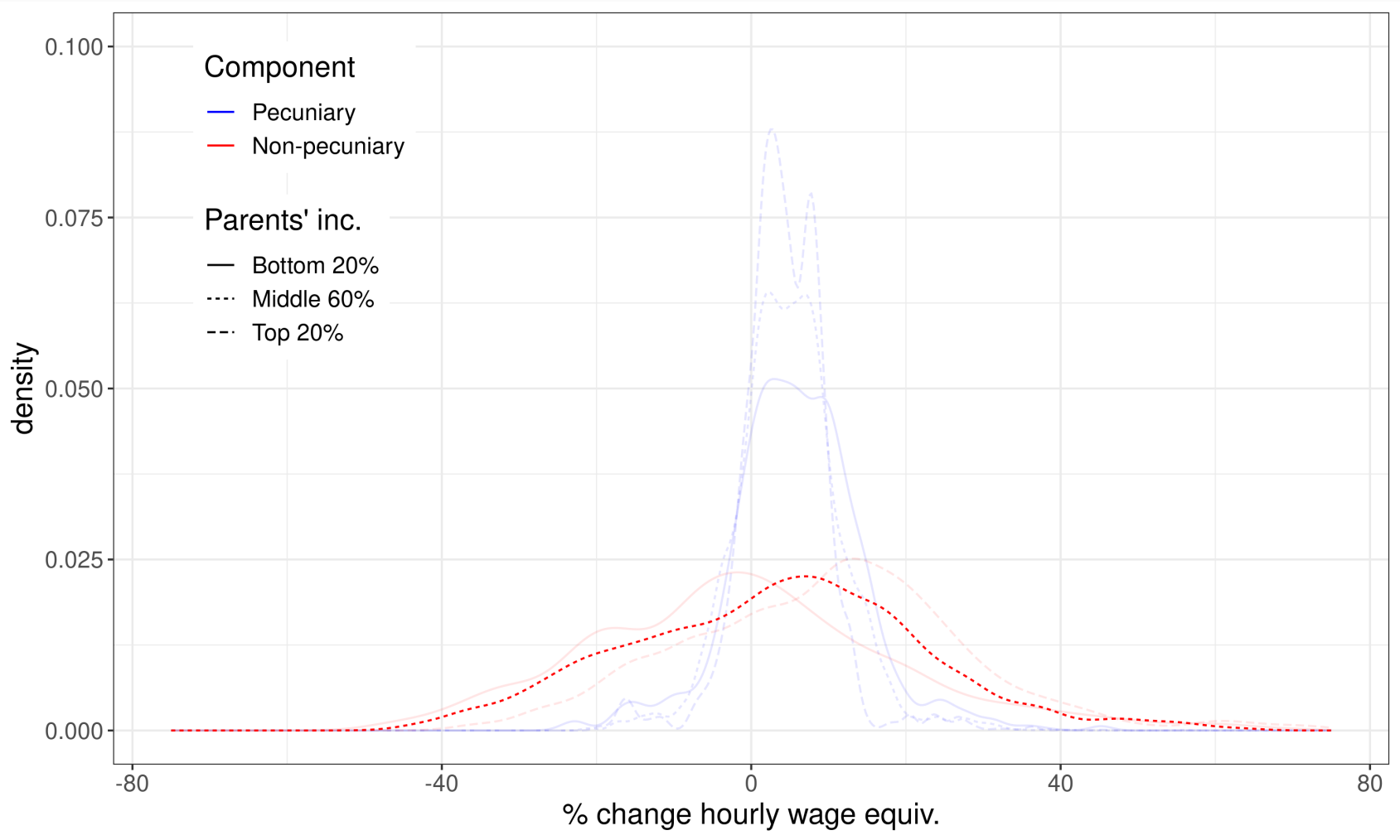
By SES (1990, SES = parents' income at 16)

Parental income	Mean	Variance	Skewness
Bottom 20%	−0.185	0.066	0.520
Middle 60%	−0.183	0.063	0.524
Top 20%	−0.037	0.297	28.57



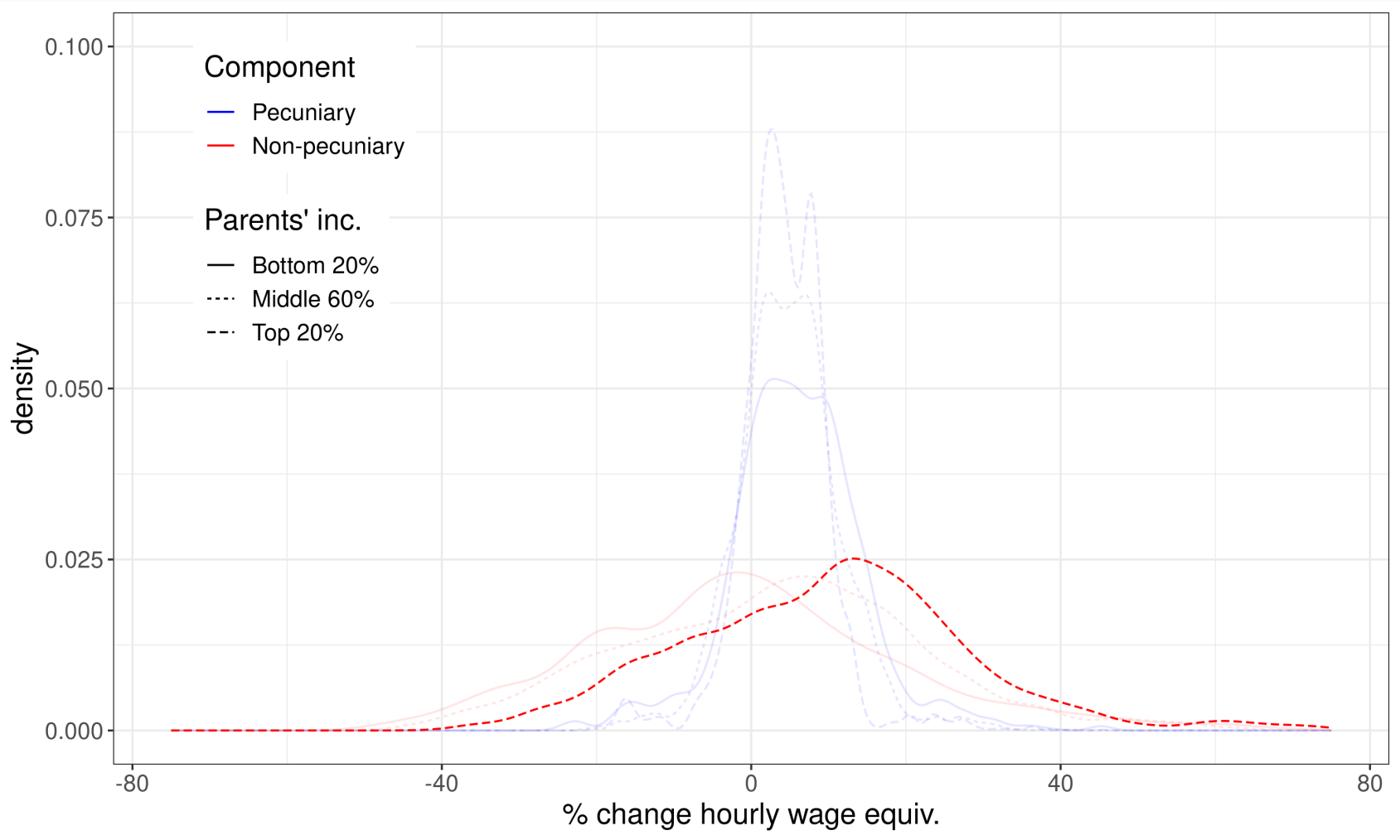
By SES (1990, SES = parents' income at 16)

Parental income	Mean	Variance	Skewness
Bottom 20%	−0.185	0.066	0.520
Middle 60%	−0.183	0.063	0.524
Top 20%	−0.037	0.297	28.57



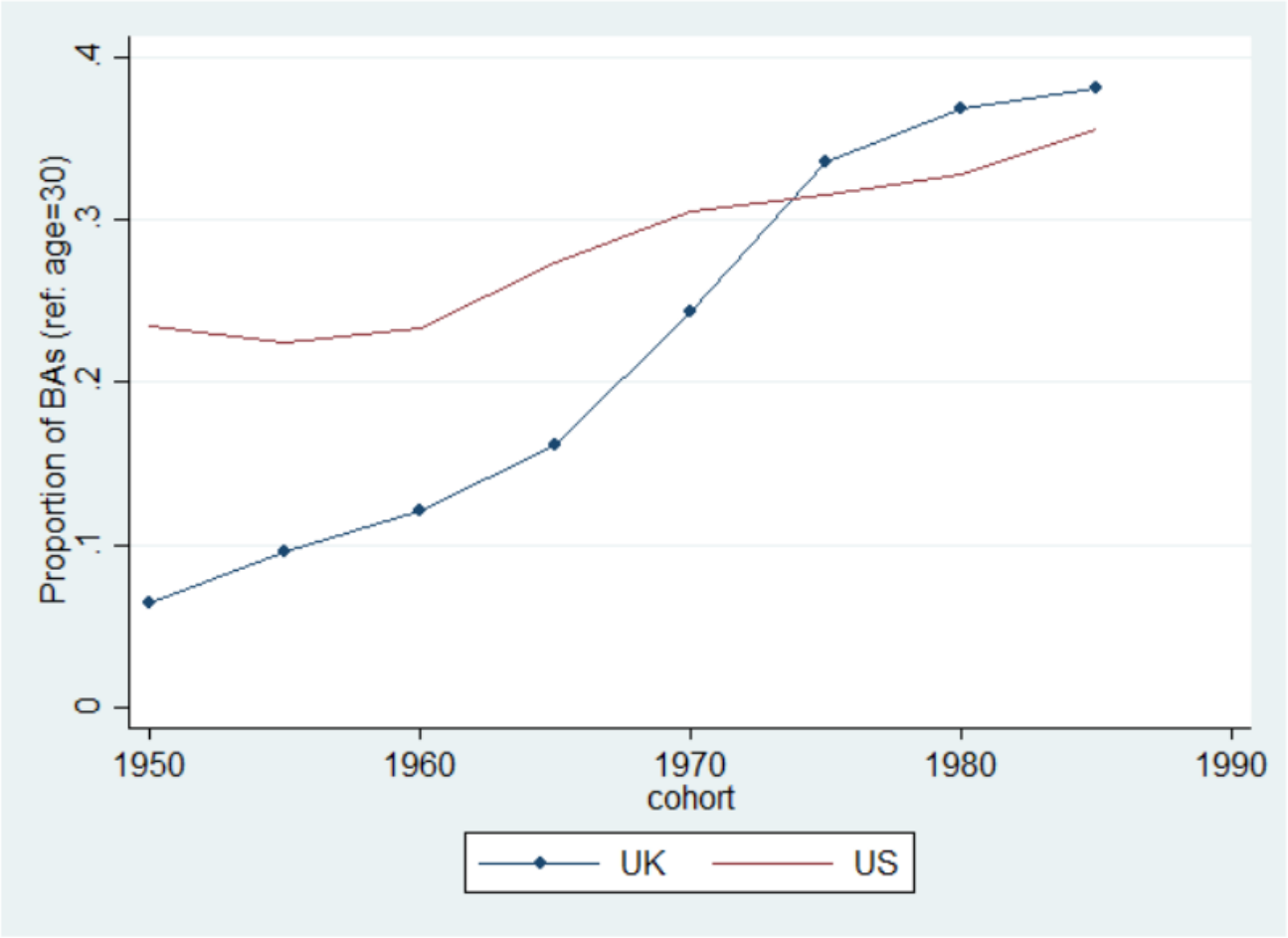
By SES (1990, SES = parents' income at 16)

Parental income	Mean	Variance	Skewness
Bottom 20%	−0.185	0.066	0.520
Middle 60%	−0.183	0.063	0.524
Top 20%	−0.037	0.297	28.57



Huge expansion in HE attainment

Figure 2: Proportion of People with a BA or Higher Education by Cohort, UK and US



Note: sample restricted to age 22-59 and excludes full-time students. Each education-cohort cell has at least 100 observations.

Source: Blundell et al. (2018).

Huge expansion in HE attainment

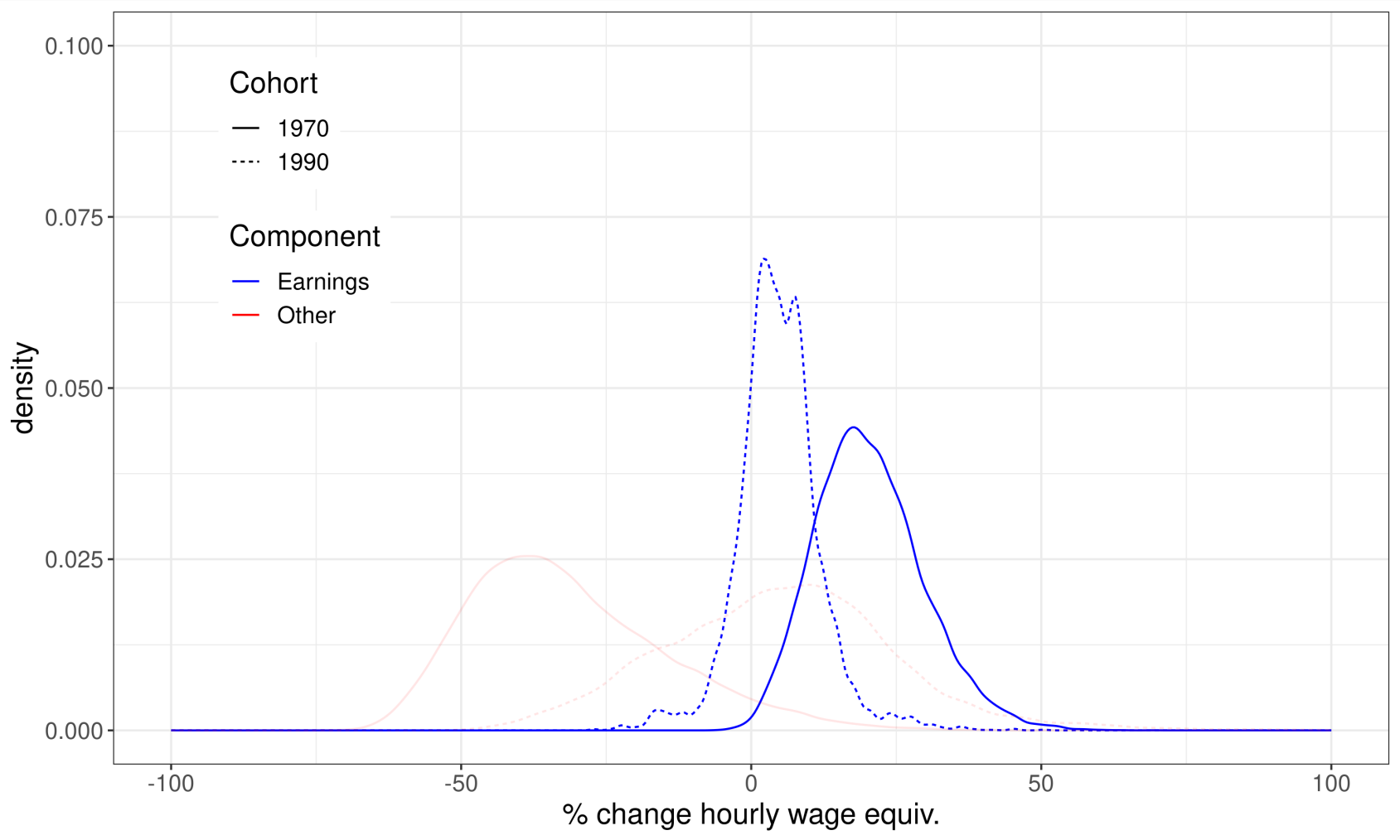
- Re-estimate model on data from similar cohort study born in 1970

  $\Rightarrow$ Compare factors in 1970 to 1990

Changes between 1970 and 1990 cohorts

- earnings exp. fell: 20% $\rightarrow$ 4.8% (mean); 9.3% $\rightarrow$ 7.1% (sd)
- non-pec. benefits increased: $-31\% \rightarrow 4.9\%$ (mean); $17\% \rightarrow 20\%$ (sd)

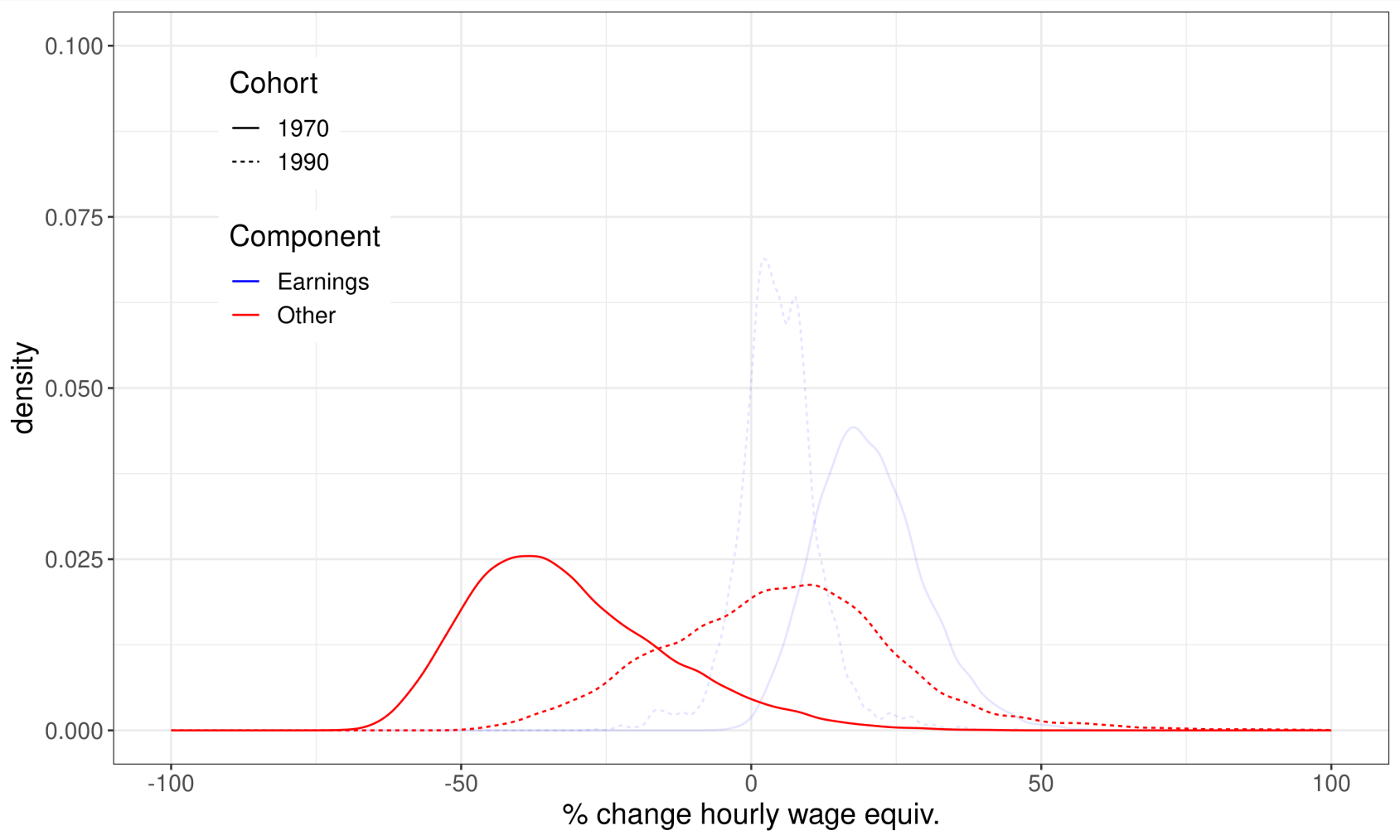
$\Rightarrow$ non-pec. factors driving the large increase in attendance



Changes between 1970 and 1990 cohorts

- earnings exp. fell: 20% → 4.8% (mean); 9.3% → 7.1% (sd)
- non-pec. benefits increased: −31% → 4.9% (mean); 17% → 20% (sd)

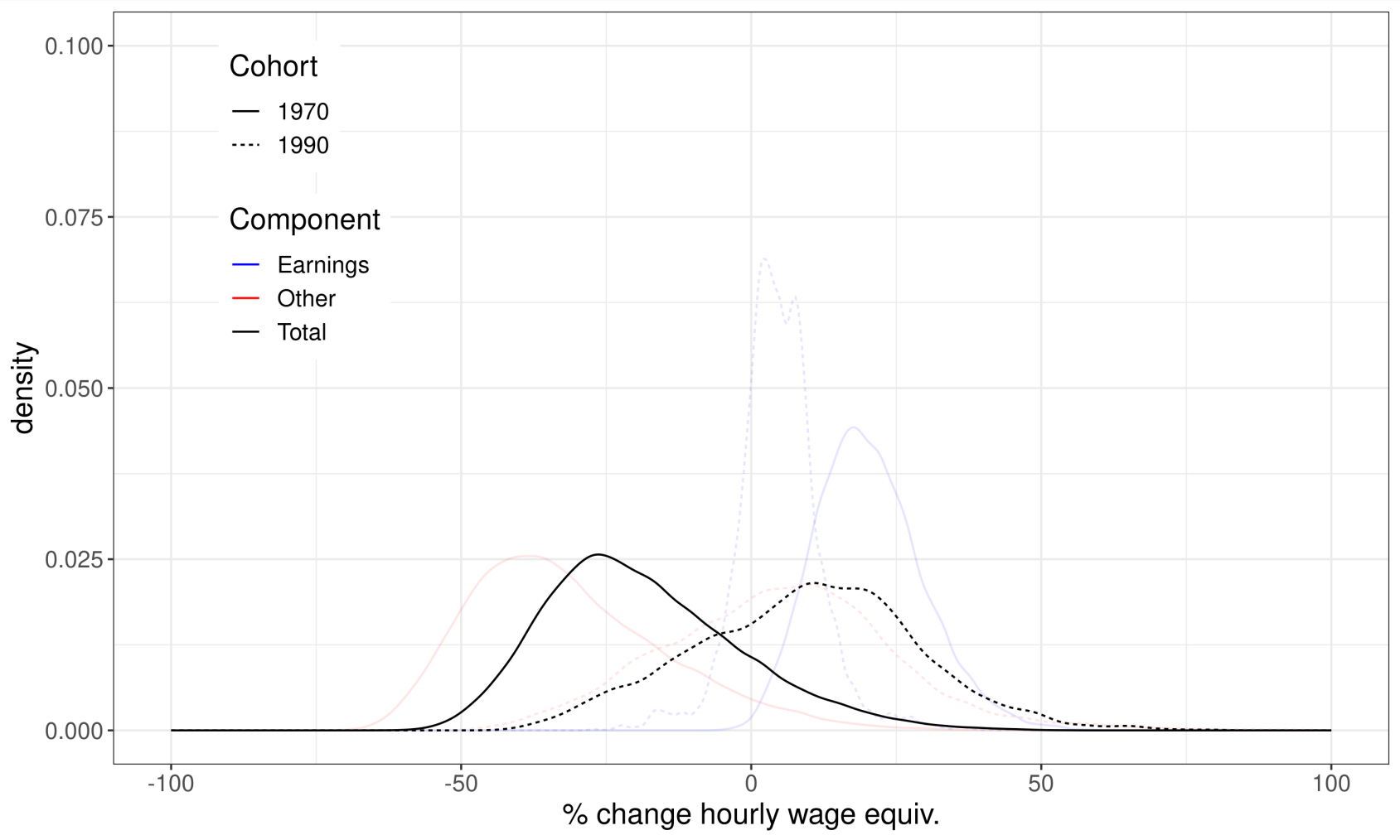
  ⇒ non-pec. factors driving the large increase in attendance



Changes between 1970 and 1990 cohorts

- earnings exp. fell: 20% $\rightarrow$ 4.8% (mean); 9.3% $\rightarrow$ 7.1% (sd)
- non-pec. benefits increased: -31% $\rightarrow$ 4.9% (mean); 17% $\rightarrow$ 20% (sd)

$\Rightarrow$ non-pec. factors driving the large increase in attendance



What’s next?

Conclusion

- Earnings exp. less important than other factors in HE decision
- Earnings exp. similar across SES, other factors driving gap
- Other factors drove large increase in HE attainment since 1970s

Still lots to do:

1. **Continue to decompose other factors into meaningful components**
   - Form natural "groups" of factors: Career; financial (now and future); Social life / environment; Education; Personal development; Time
   - But ind. parameters $\rightarrow$ smaller groups difficult to interpret
2. **Improve the model of earnings expectations**
   - Model lifetime earnings, rather than relying on single point
   - Allow for **unobserved heterogeneity** in earnings [see e.g. Heckman et al. (2006)]

Conclusion

- Earnings exp. less important than other factors in HE decision
- Earnings exp. similar across SES, other factors driving gap
- Other factors drove large increase in HE attainment since 1970s

Still lots to do:

1. **Continue to decompose other factors into meaningful components**
   - Form natural "groups" of factors: Career; financial (now and future); Social life / environment; Education; Personal development; Time
   - But ind. parameters $\rightarrow$ smaller groups difficult to interpret
2. **Improve the model of earnings expectations**
   - Model lifetime earnings, rather than relying on single point
   - Allow for **unobserved heterogeneity** in earnings [see e.g. Heckman et al. (2006)]

Thank you

Thank you :)

References i

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.

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.

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.

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.

References ii

Mincer, J. (1974). *Schooling, experience, and earnings*. Number 2 in Human behavior and social institutions. National Bureau of Economic Research; distributed by Columbia University Press, New York.

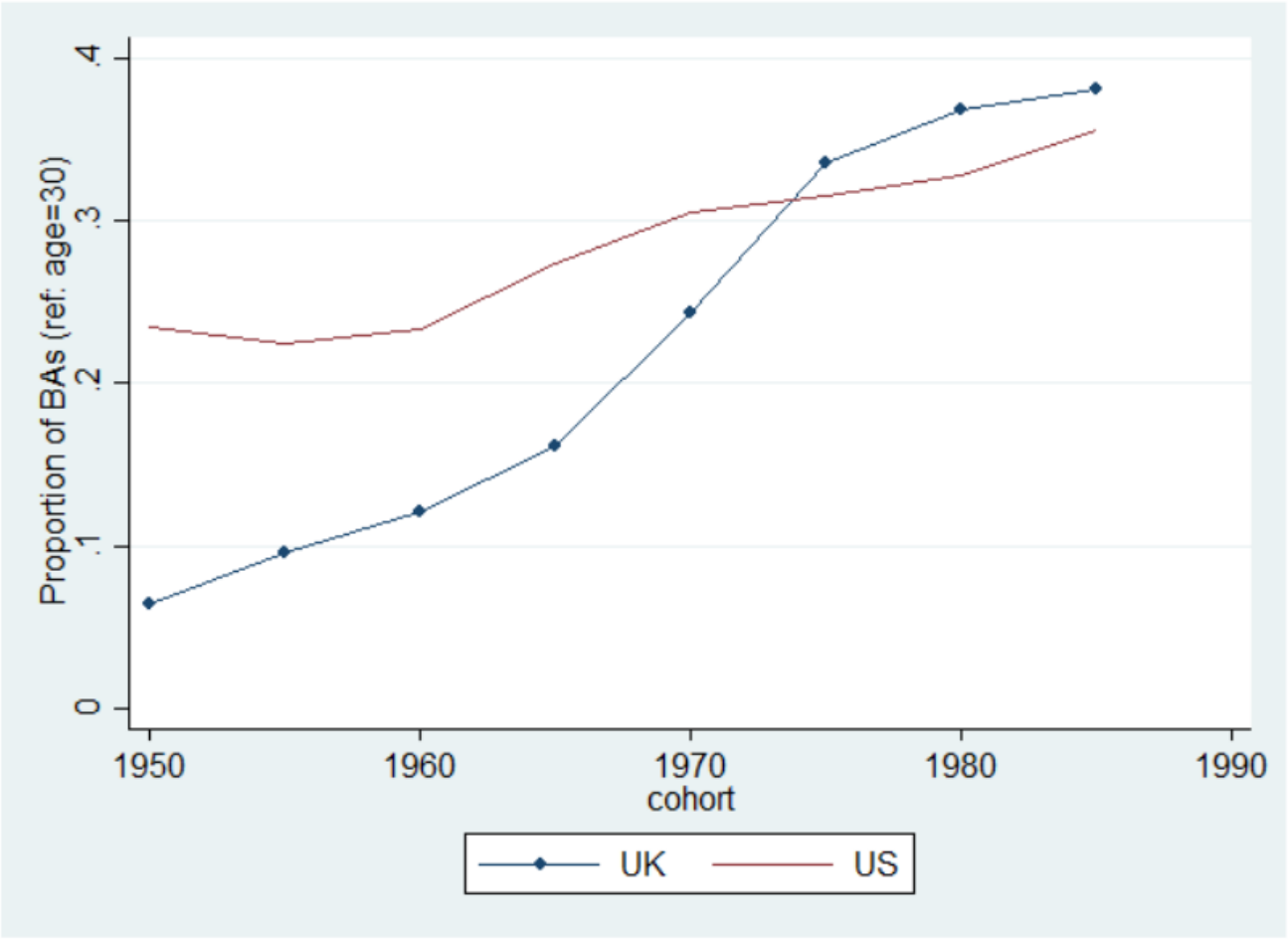
OECD (2018). *Equity in Education: Breaking Down Barriers to Social Mobility*. PISA. OECD.

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.

UK HE Expansion

Figure 2: Proportion of People with a BA or Higher Education by Cohort, UK and US



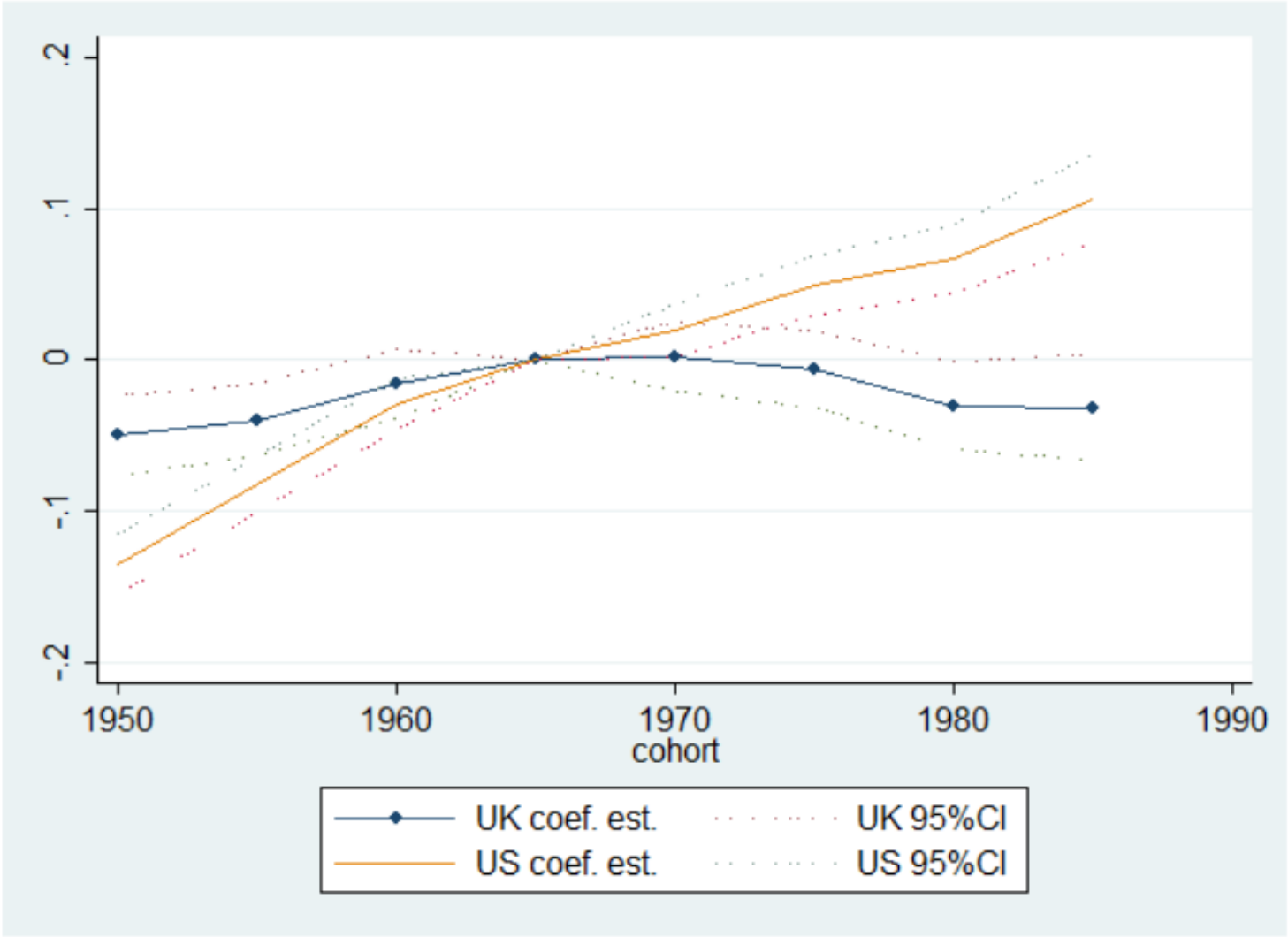
Note: sample restricted to age 22-59 and excludes full-time students. Each education-cohort cell has at least 100 observations.

Source: Blundell et al. (2018).

back

Graduate-wage premium

Figure 3: Ratio of BA median wage to that of high-school graduates, cohort effects

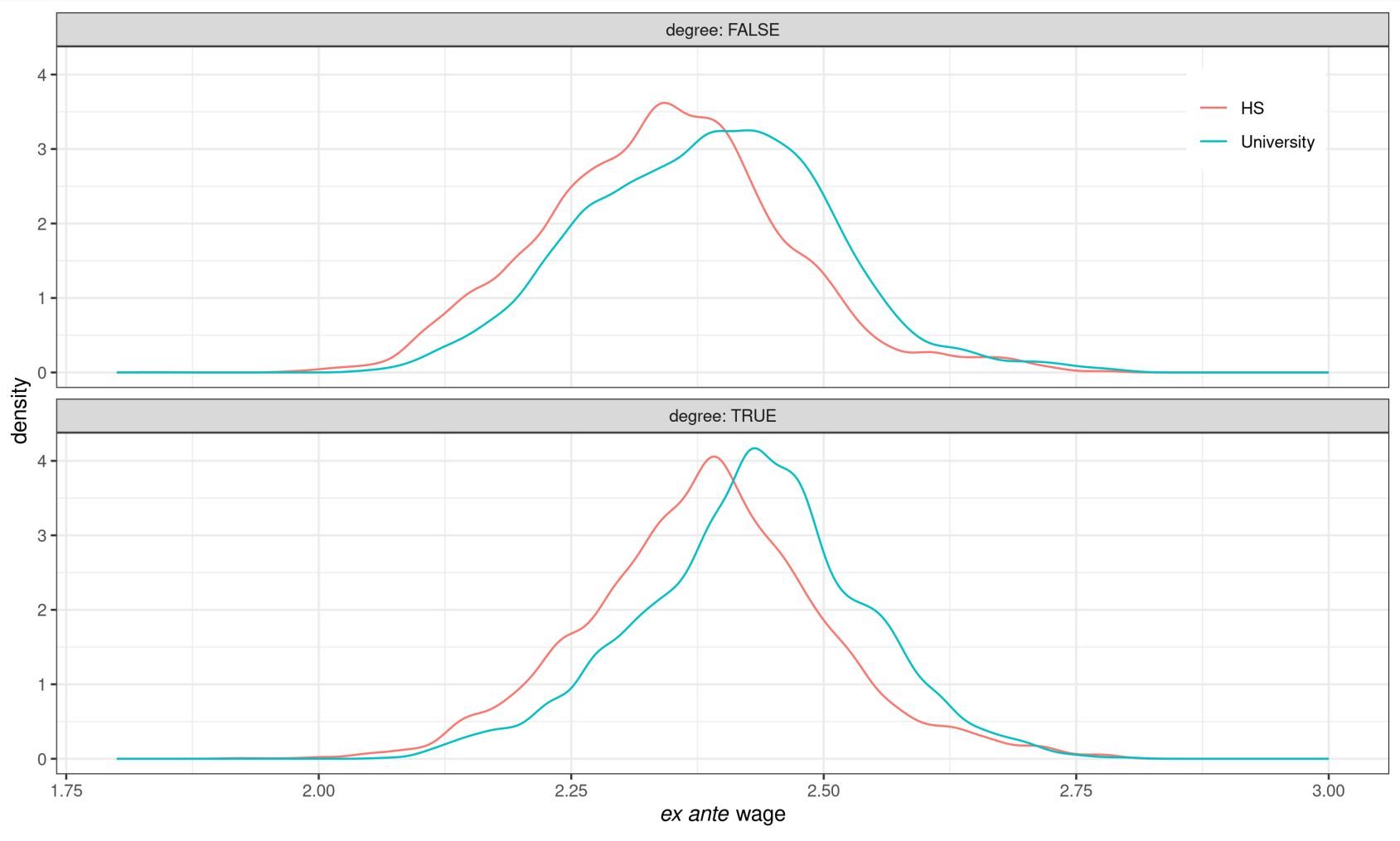


Source: Blundell et al. (2018).

back

Wage distributions (Next Steps)

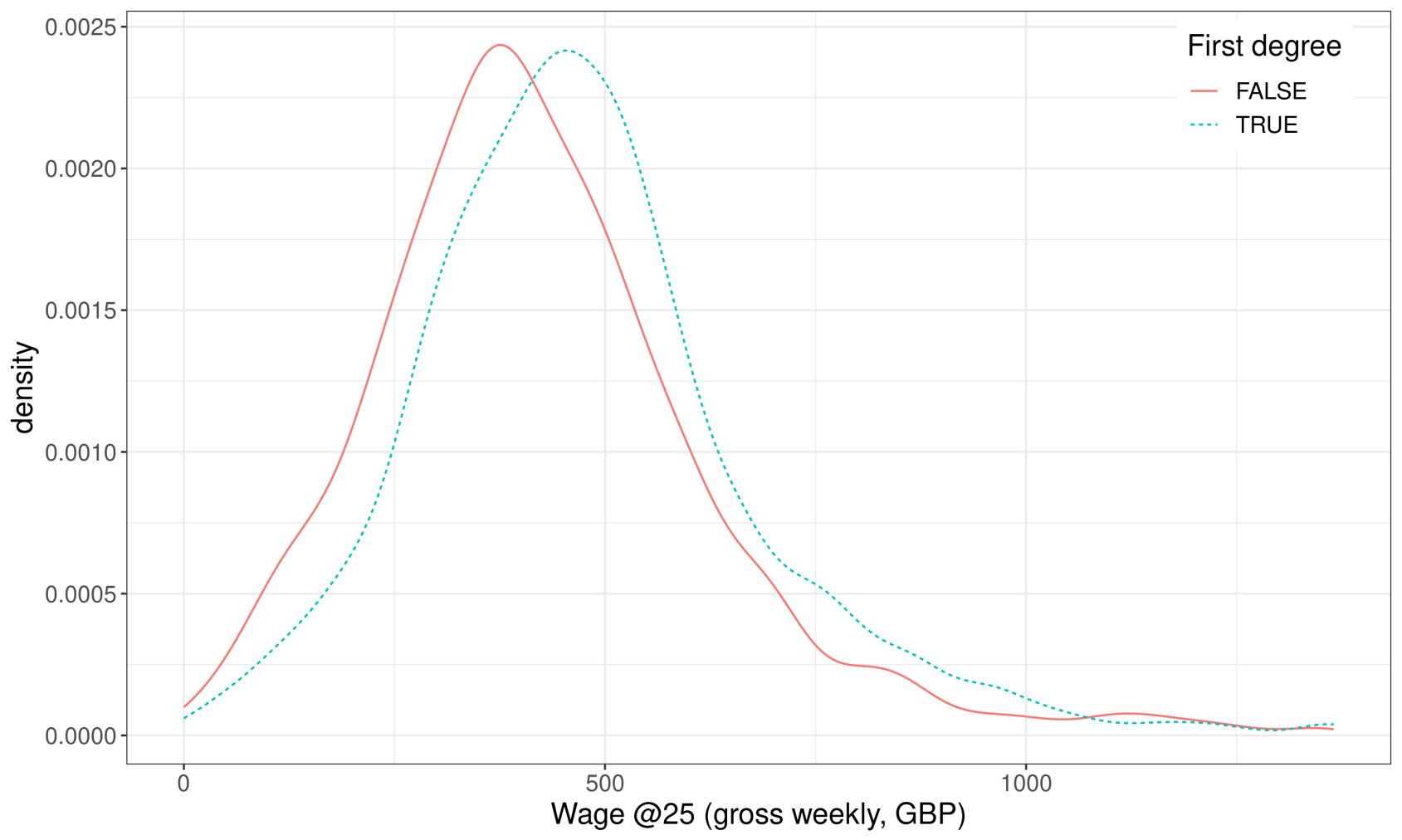
Ex ante (expected) wages, Y_s^{ea}



back

Wage distributions (Next Steps)

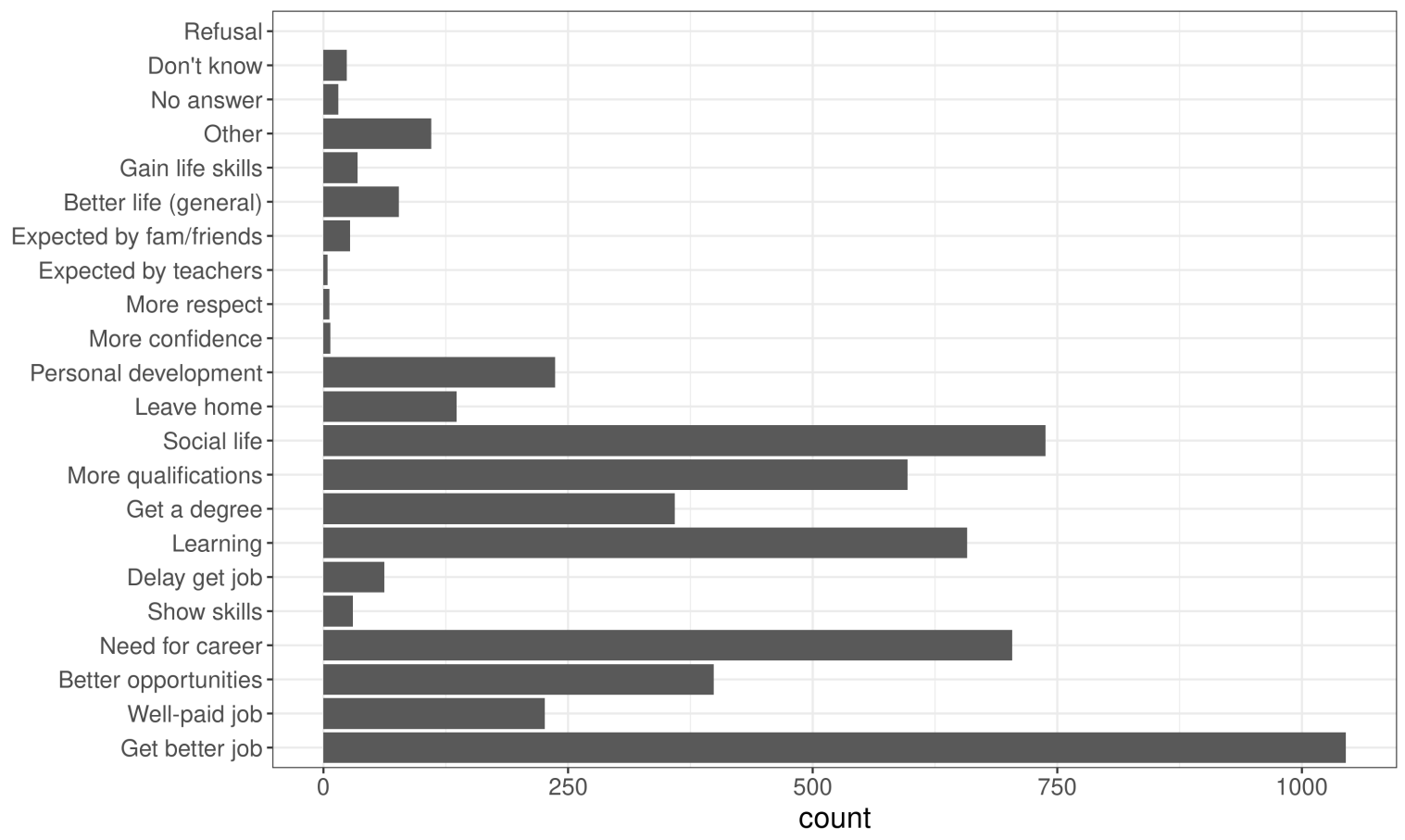
Ex post (realised) wages, Y_s



back

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?

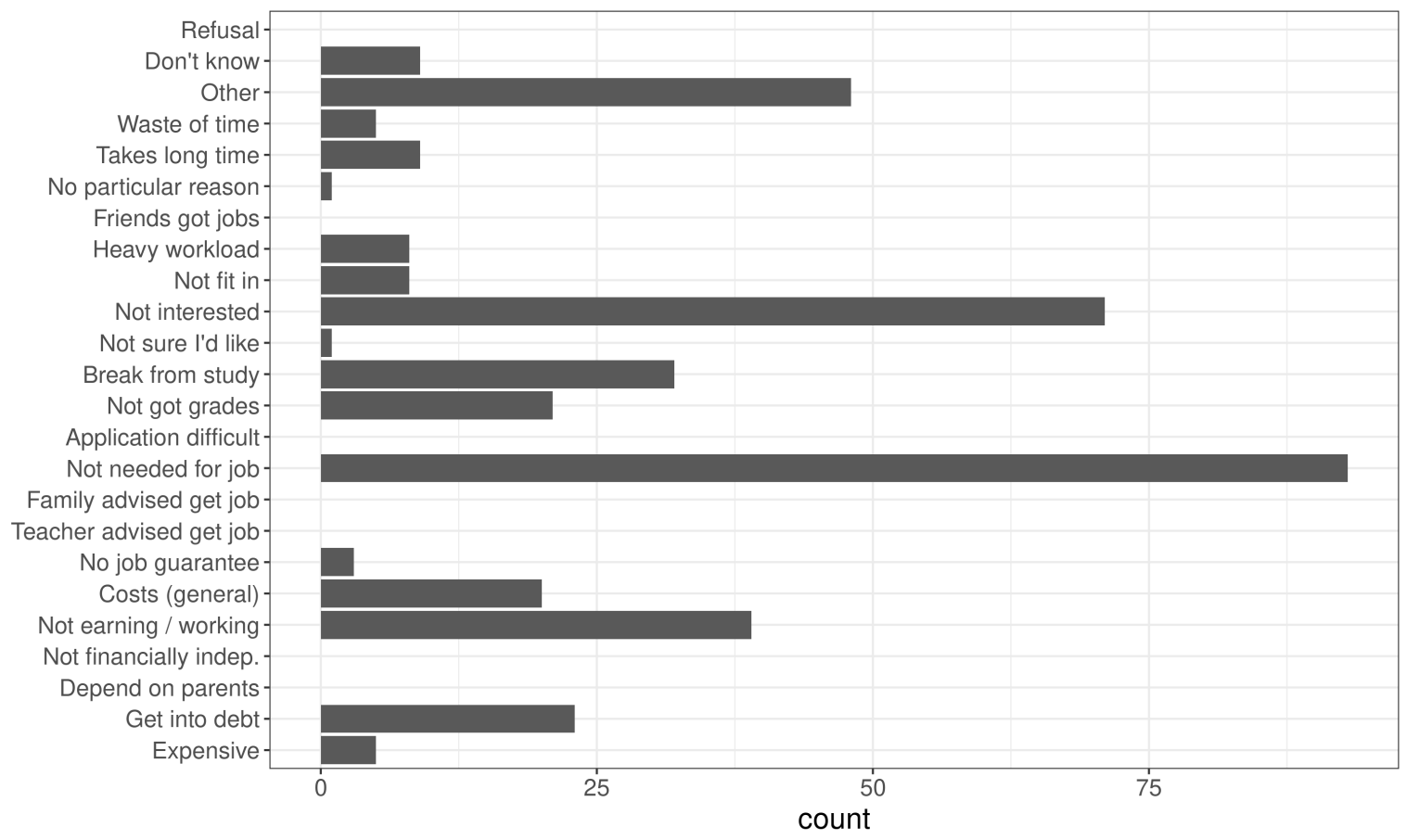


Notes: Students with >5 GCSEs @A*–C and **who plan to apply**.
Open-ended. $N = 4,640$.

back

Main reasons for not applying

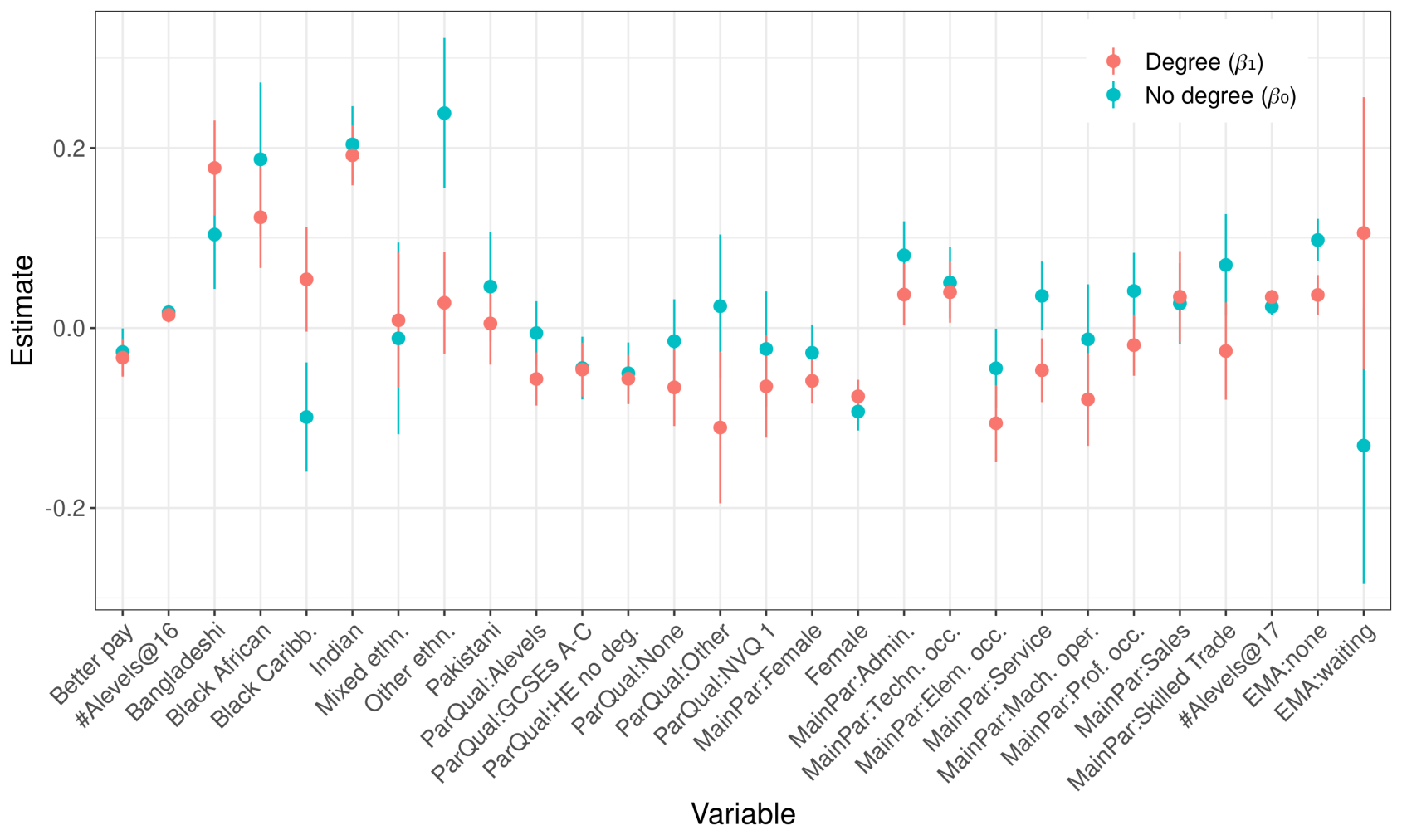
You have said that you are not planning to apply to university/not likely to ever apply to university. What are the MAIN reasons why you decided not to apply for a place at a university?



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

back

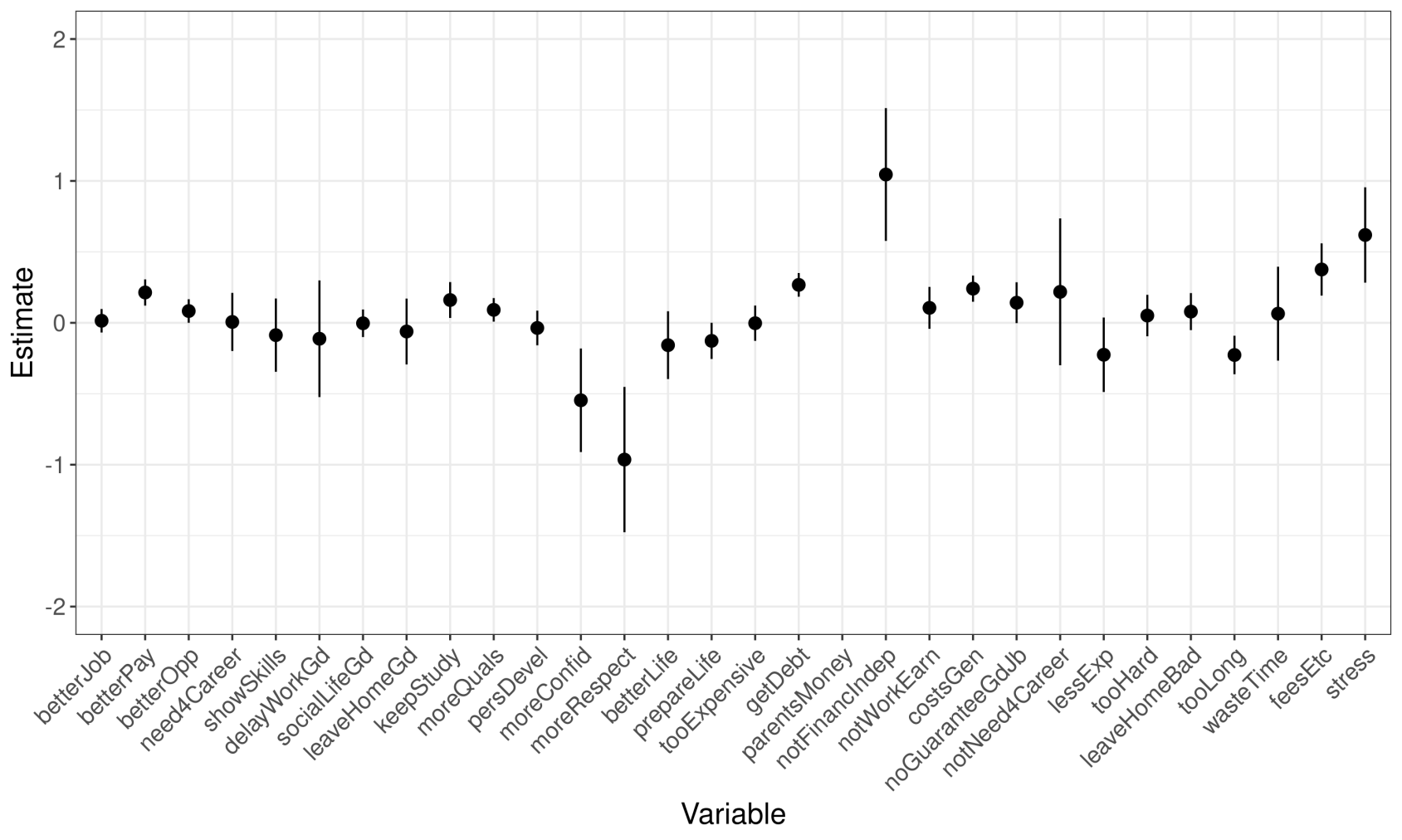
Wage equation parameters, β_s



back

Choice equation parameters, γ

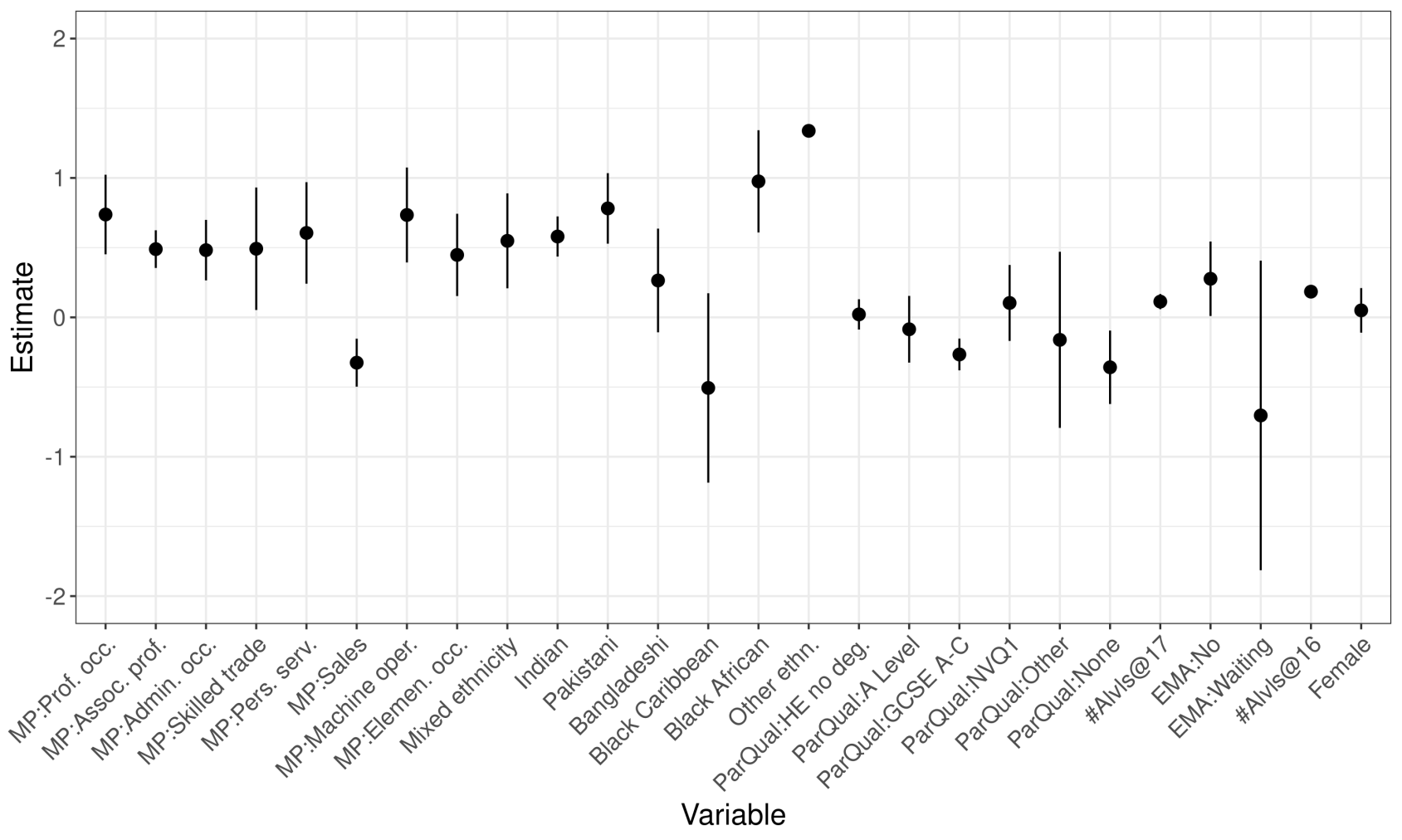
Responses to open-ended questions



back

Choice equation parameters, γ

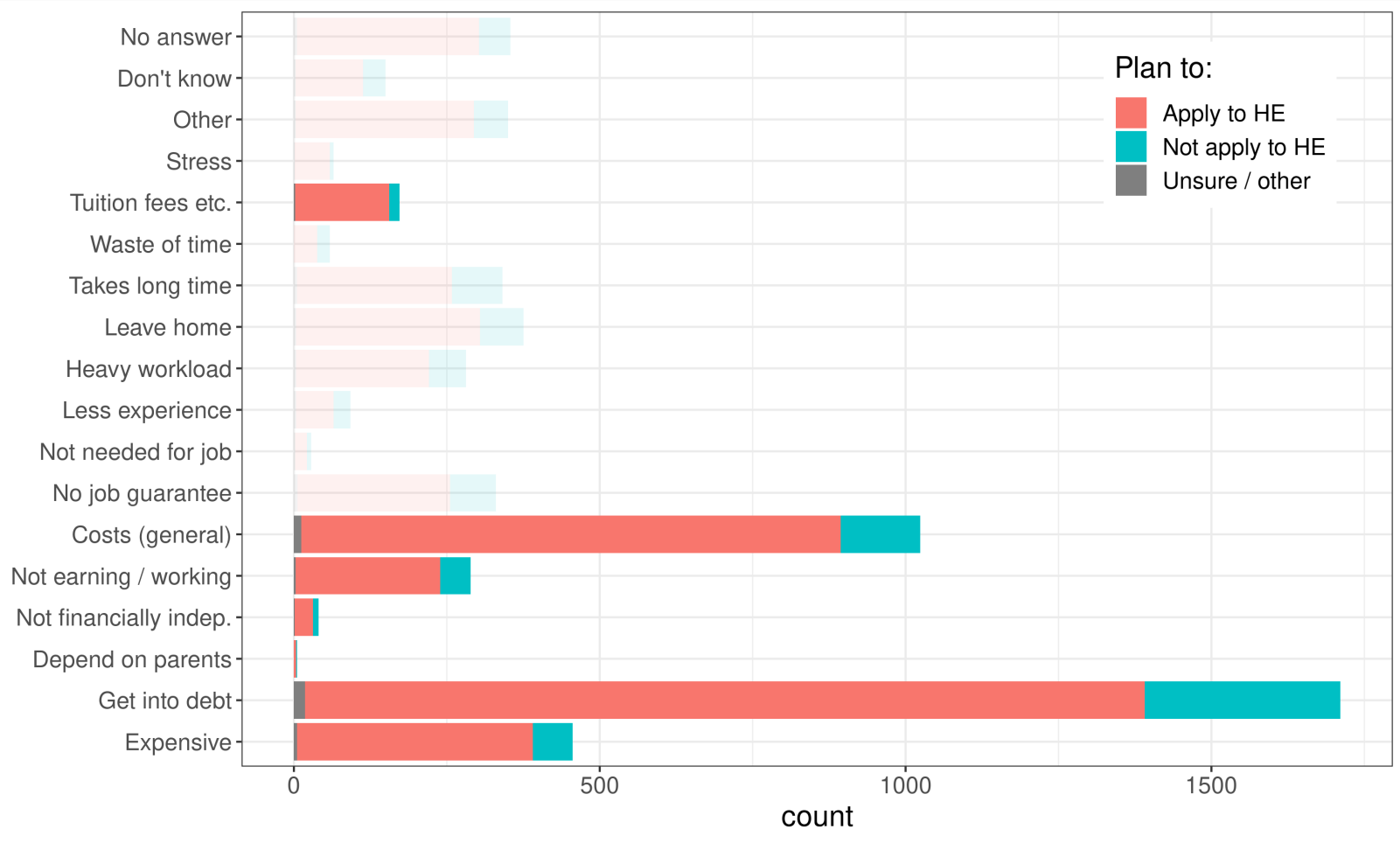
Background characteristics, X



back

Financial factors

Responses to open-ended questions classified as financial factors for decomposition of “psychic costs” (all *disadvantages*)



back