

# Preferences over Higher Education and Jobs: Understanding the Gender Gap

Aparajita Dasgupta    Anisha Sharma

Ashoka University

March 28, 2019

# Question

Can heterogeneous preferences over education and jobs explain the gender earnings gap?

# Motivation

Recent evidence suggests wage gaps across subjects is as large as across school vs college education (Altonji 2014)

# Motivation

- In the US, women are two-thirds as likely as men to study STEM (Gemici and Wiswall 2014).
- In India, women are a third as likely as men to study computer science and engineering and two thirds as likely to study business (MHRD 2018).
- In Indian schools, girls are  $3/4$  as likely as boys to study science and almost equally likely to study business (DISE 2018)

# Motivation

- In the US, women are two-thirds as likely as men to study STEM (Gemici and Wiswall 2014).
- In India, women are a third as likely as men to study computer science and engineering and two thirds as likely to study business (MHRD 2018).
- In Indian schools, girls are  $3/4$  as likely as boys to study science and almost equally likely to study business (DISE 2018)

Can different preferences over subjects and jobs explain the gender gap? Why do relatively few women study lucrative subjects?

# What does this project do

- Estimate the economic return to different subject streams
  - Across STEM, business and the humanities

# What does this project do

- Estimate the economic return to different subject streams
  - Across STEM, business and the humanities
- Identify determinants of subject choice
  - Incorporate measures of ability, non-cognitive traits, preferences, expectations

# What does this project do

- Estimate the economic return to different subject streams
  - Across STEM, business and the humanities
- Identify determinants of subject choice
  - Incorporate measures of ability, non-cognitive traits, preferences, expectations
- Estimate a model of subject choice given preferences over subjects and job characteristics



## Related literature

- Return to a major: Altonji et al. (2012, 2014), Gemici & Wiswall (2014), Sahoo and Klaasen (2018)
- Estimating a model of subject choice by gender:
  - Preferences rather than abilities: Gemici and Wiswall (2014)
  - Preferences for job-related attributes: Zafar (2013), Wiswall and Zafar (2017)
  - Behavioural differences: Reuben et al (2013)
- Non-cognitive skills and education choice: Buser et al (2014)
- Subjective expectations and education choice: Dominitz and Manski (1996), Attanasio and Kaufman (2014)
- Stated preference data and education choice (Weinberger 2004)

- 1 Introduction
- 2 Data
- 3 Returns to majors
- 4 Major choice
- 5 Survey
- 6 Conclusion

## Institutional background

Elite graduate programme at a university near Delhi

# Institutional background

Elite graduate programme at a university near Delhi

- Multiple cohorts (2011-2018) of comparable students in terms of abilities
- Distribution across subjects is similar to national distribution
- Diverse in terms of subject and socioeconomic background
- 73% on financial aid

# Young India Fellowship

## Aims of the programme

- Identify and develop “leadership potential” among elite students
- Study a diverse range of subjects across 8 terms in 48 weeks
- Undertake a 6-month internship project in partnership with a client from the private, public or social sector
- Receive mentoring from business and academic leaders to develop their professional skills and networks

# Administrative data

- Data collected as part of online application
- Academic background (class 10, 12, UG, PG)
- Socioeconomic background (location of upbringing, parent's education and employment status, household income)

# Survey data

## Subject choice module

- Directly collected information on preferences

# Survey data

## Subject choice module

- Directly collected information on preferences

## Labour market module

- Labour market histories,
- Expectations of future labour outcomes



# Survey data

## Subject choice module

- Directly collected information on preferences

## Labour market module

- Labour market histories,
- Expectations of future labour outcomes

## Personality traits module

- Measures of attitudes towards risk and competition, and self-esteem

# Sample

We collect information on 675 jobs of 326 survey respondents across a number of occupations.

- Between 22-35 yrs old
- 54% women
- 50% science, 20% econ/business, 30% humanities majors
- Information includes wages, occupational sector, duration of work, location

# Sample

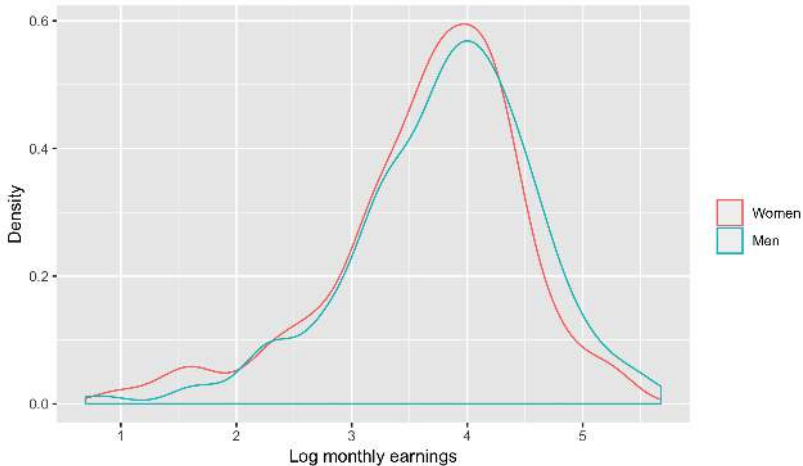
We collect information on 675 jobs of 326 survey respondents across a number of occupations.

- Between 22-35 yrs old
- 54% women
- 50% science, 20% econ/business, 30% humanities majors
- Information includes wages, occupational sector, duration of work, location

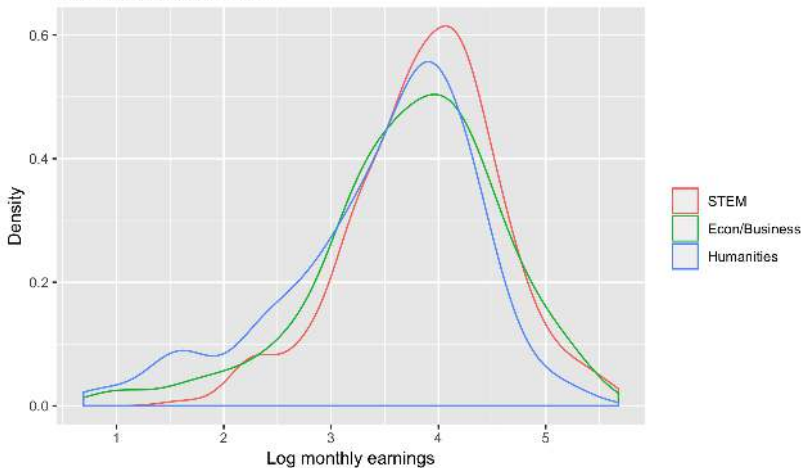
Comparison with all-India data:

- Women account for 47% of enrolment.
- **All India:** 56% of all students enrol in arts/business in college, 41% in STEM (MHRD 2018)

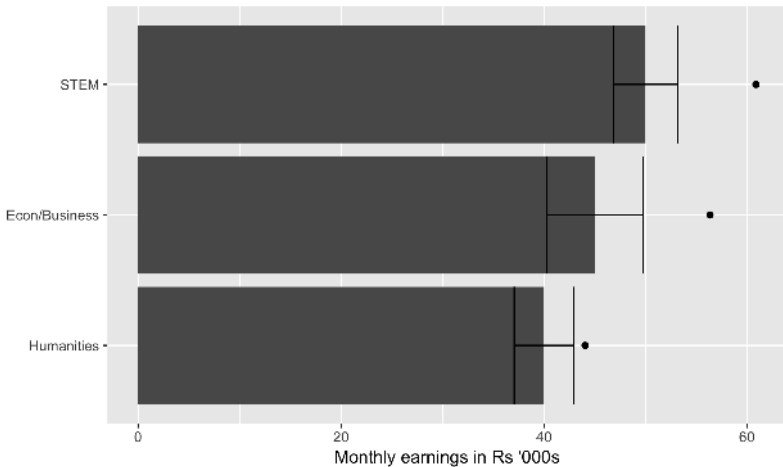
### Kernel density estimates



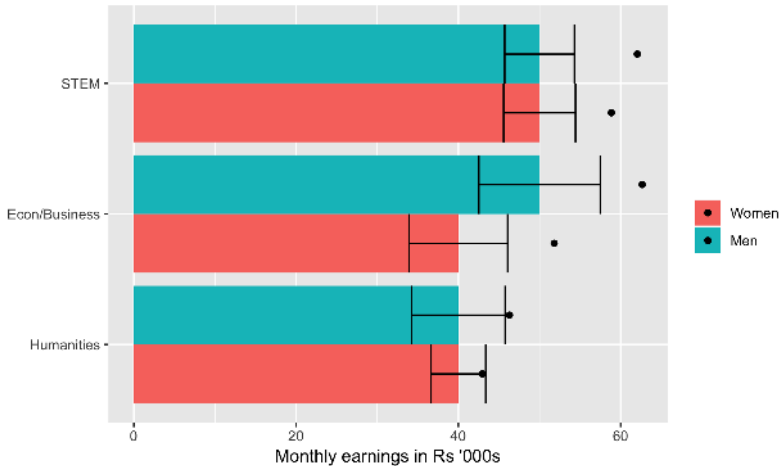
### Kernel density estimates



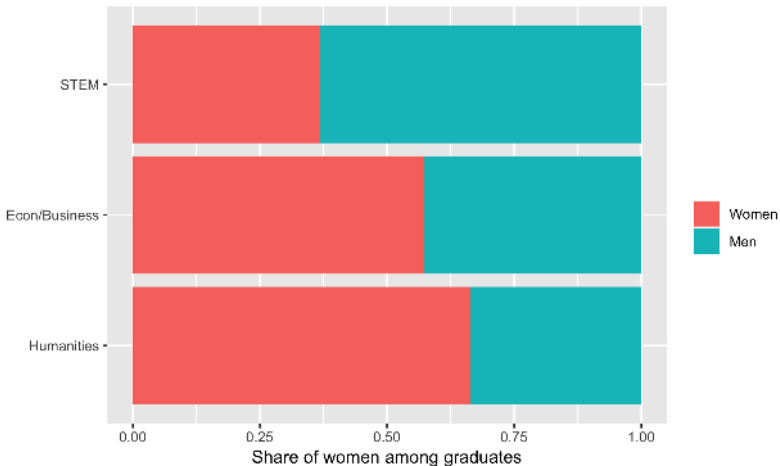
### Earnings by stream



## Earnings by gender and stream

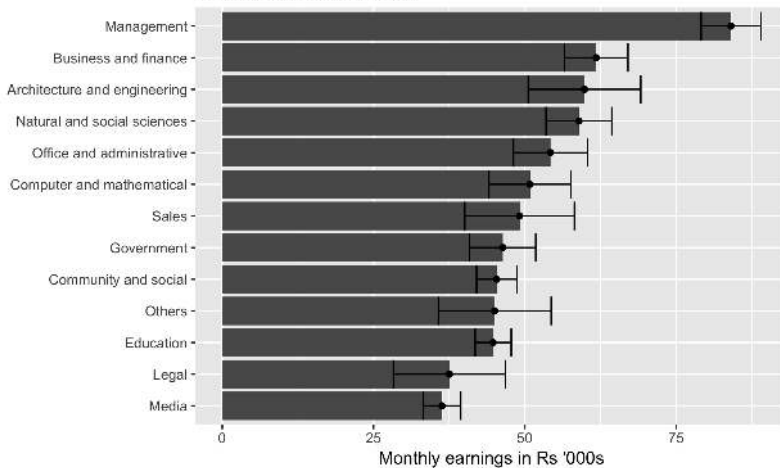


### Share of women across majors

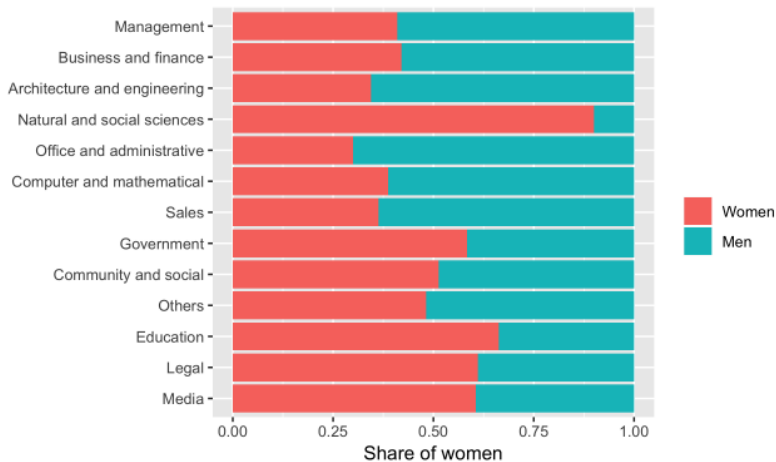




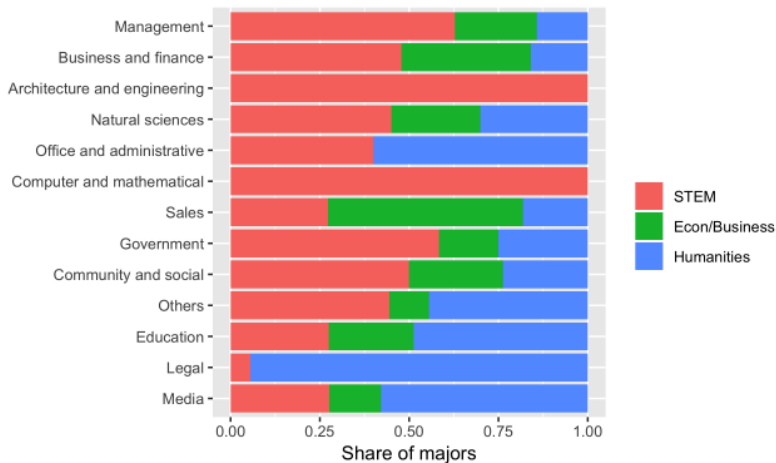
## Earnings by occupation



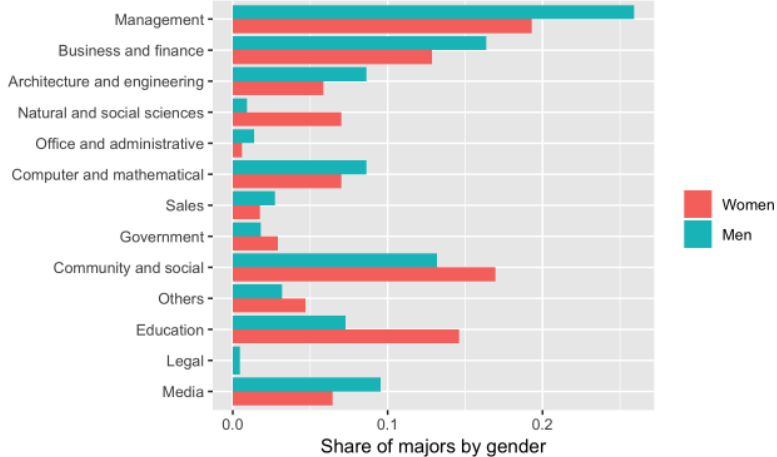
## Share of women among graduates



## Share of majors within occupations



## Share of STEM/Econ graduates across occupations



# Earnings

$$y_{imjt} = \beta_m M_i + \beta_g \text{Male}_i + \gamma \mathbf{X}_i + \psi_j + \delta_t + \epsilon_{imjt}$$

- $y_{imjt}$  are the earnings of person  $i$  in occupation  $j$  at time  $t$ , after studying major  $m$
- $M_i$  is the major studied by person  $i$
- $X_i$  contains measures of ability, non-cognitive traits and socio-economic background
- Standard errors clustered by individual

# Earnings

	Dependent variable: log income in '000s of Rs				
	1	2	3	4	5
Econ & Business	0.108 (0.097)	0.064 (0.101)	0.040 (0.105)	0.055 (0.104)	0.120 (0.098)
Arts & soc sci	-0.218** (0.084)	-0.212** (0.096)	-0.187** (0.091)	-0.222** (0.090)	0.021 (0.092)
Male	0.136* (0.073)	0.175** (0.076)	0.231*** (0.079)	0.222*** (0.075)	0.152** (0.067)
Work Exp (yrs)	0.302*** (0.066)	0.159*** (0.021)	0.126*** (0.024)	0.127*** (0.024)	0.120*** (0.022)
Grade 10 testscore		0.180** (0.074)	0.167** (0.080)	0.175** (0.082)	0.107 (0.077)
English fluency		0.226** (0.087)	0.109 (0.083)	0.106 (0.082)	0.121 (0.076)
Competitiveness			-0.034 (0.047)	-0.034 (0.045)	-0.004 (0.044)
Self-esteem			0.015* (0.008)	0.015* (0.008)	0.010 (0.008)
Admissions score			0.076** (0.038)	0.083** (0.037)	0.085** (0.033)
Risk measure				0.002 (0.012)	-0.005 (0.011)
Year effects	Yes	Yes	Yes	Yes	Yes
Occupation controls	No	No	No	No	Yes
HH controls	No	No	No	Yes	No
Observations	675	670	611	611	611
$R^2$	0.25	0.26	0.32	0.32	0.41

## Explaining major choice

- Ability
- Non-cognitive measures (attitudes towards risk, competition, self-esteem)
- Preferences
- Expectations

# Explaining major choice

- Ability
- Non-cognitive measures (attitudes towards risk, competition, self-esteem)
- Expectations
- **Preferences**



# Subject choice model

Student  $i$  faces a choice across  $M$  different majors derives utility from choosing the  $m$ th major of

$$U_{im} = u(Y_{im}, Z_{im}, X_i) + \epsilon_{im} \quad (1)$$

- $Y_{1im}, Y_{2im}, \dots, Y_{Kim}$  are major-related characteristics
- $Z_{1im}, Z_{2im}, \dots, Z_{Jim}$  are job-related characteristics
- $X_{1i}, X_{2i}, \dots, X_{Li}$  are individual characteristics
- $\epsilon_{im}$  is any unobserved preference of student  $i$  for major  $m$

# Subject choice model

- **Major-related characteristics:** love for the subject, expected grades, training for a specific career
- **Job-related characteristics:** expected wages, work-life balance
- **Individual characteristics:** desire to please parents, peer pressure, influence of teachers and mentors

A student will select major  $m$  over major  $m'$  if  $EU_{im} > EU_{im'}$  for all  $m \neq m'$

# Estimating a choice model

Approach in the literature so far:

- Make assumptions about expectations and use observed choices to estimate model parameters
- Elicit expectations and estimate model parameters
- Collect stated preferences

## Survey questions on preferences

How important were each of the following factors in your decision to study [chosen major]?

- To learn more about subjects I enjoyed studying or was interested in
- To get high grades on the coursework
- ...etc...

Respondents choose from a scale of 0 to 3 where 0 is “not at all important” and 3 is “extremely important”.

Do stated preferences vary by gender?

## Preferences by gender



# Preferences by gender

Use an ordered probit to identify predictors of how important each factor is:

- Gender
- Ability
- Personality
- Socioeconomic background

# Preferences

Dependent variable:

0=Not important, 1=Somewhat important, 2=Moderately important, 3=Very important

	Love for subject	Grades	Wages	Career	Work-life balance	Parents	Peers	Teachers	Option value
Male	-0.613*** (0.138)	-0.450*** (0.128)	0.101 (0.128)	-0.298** (0.131)	-0.130 (0.126)	0.225* (0.134)	0.380*** (0.133)	0.103 (0.127)	-0.210* (0.127)
Grade 10 score	-0.032*** (0.012)	-0.001 (0.013)	0.033*** (0.012)	0.016 (0.013)	0.004 (0.011)	0.043*** (0.012)	0.045*** (0.013)	0.004 (0.012)	0.034*** (0.013)
Risk measure	-0.031 (0.025)	-0.008 (0.022)	0.022 (0.023)	0.005 (0.023)	0.024 (0.021)	-0.001 (0.022)	-0.001 (0.024)	0.021 (0.023)	0.019 (0.023)
Competitiveness	0.223** (0.092)	0.291*** (0.090)	0.189** (0.093)	0.237*** (0.091)	0.187* (0.096)	0.090 (0.088)	0.076 (0.096)	0.309*** (0.088)	0.247*** (0.090)
Self-esteem	0.019 (0.014)	-0.014 (0.013)	-0.022 (0.014)	0.012 (0.014)	-0.027** (0.013)	-0.025* (0.014)	-0.017 (0.015)	-0.005 (0.014)	-0.015 (0.014)
Mother works	-0.286** (0.138)	-0.251** (0.123)	-0.299** (0.127)	-0.333*** (0.125)	-0.168 (0.128)	0.040 (0.128)	-0.013 (0.131)	-0.026 (0.130)	0.005 (0.131)
Mother has UG degree	0.230 (0.186)	0.021 (0.183)	-0.142 (0.156)	-0.163 (0.173)	-0.229 (0.163)	0.140 (0.187)	0.094 (0.189)	0.183 (0.189)	0.247 (0.180)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	322	322	322	322	322	322	322	322	322
Log pseudolikelihood	-347.36	-419.23	-422.30	-401.40	-428.23	-431.95	-378.51	-421.76	-406.99

## Preferences vary by gender

**Women were more likely** to state the following factors were “very important”

- Love for the subject: 24 pp
- Expected grades: 8.5 pp
- Career path: 12.2 pp
- Option value: 8.0 pp

**Men were more likely** to state the following factors were “very important”

- Desire to please parents: 6.6 pp
- Peer pressure: 5.8 pp



# Sensitivity of choice to parameters

Do changes in the reported preferences predict changes in choice of major?

# Sensitivity of choice to parameters

Do changes in the reported preferences predict changes in choice of major?

Estimate a multinomial logit model of subject choice as a function of the stated preferences separately for men and women and find that different preferences are driving selection into different majors.

# University preferences

	Science			Business			Humanities		
	(1) All	(2) Men	(3) Women	(4) All	(5) Men	(6) Women	(7) All	(8) Men	(9) Women
Love of subject	-0.12*** (0.03)	-0.09 (0.05)	-0.11* (0.05)	-0.07* (0.03)	-0.06 (0.04)	-0.09* (0.04)	0.19*** (0.03)	0.15*** (0.04)	0.20*** (0.05)
Grades	-0.07* (0.03)	-0.05 (0.04)	-0.06 (0.04)	0.03 (0.03)	0.02 (0.04)	0.03 (0.04)	0.04 (0.02)	0.03 (0.03)	0.04 (0.04)
Wages	0.04 (0.03)	-0.04 (0.04)	0.10* (0.04)	0.05 (0.03)	0.05 (0.03)	0.07 (0.04)	-0.10*** (0.03)	-0.01 (0.03)	-0.17*** (0.04)
Career	-0.03 (0.03)	-0.01 (0.04)	-0.04 (0.05)	0.04 (0.03)	0.05 (0.03)	0.01 (0.04)	-0.01 (0.03)	-0.04 (0.03)	0.03 (0.04)
Work-life balance	0.03 (0.03)	0.06 (0.04)	0.03 (0.04)	-0.03 (0.03)	-0.04 (0.03)	-0.03 (0.04)	-0.01 (0.02)	-0.02 (0.03)	0.01 (0.03)
Parents	0.04 (0.03)	0.04 (0.03)	0.03 (0.03)	-0.03 (0.02)	-0.00 (0.03)	-0.05 (0.03)	-0.01 (0.02)	-0.04 (0.02)	0.02 (0.04)
Teacher	-0.03 (0.03)	-0.07 (0.04)	-0.02 (0.04)	-0.03 (0.03)	0.01 (0.04)	-0.05 (0.03)	0.06*** (0.02)	0.06** (0.02)	0.08** (0.03)
Peers	0.10** (0.03)	0.17*** (0.04)	0.05 (0.04)	-0.02 (0.03)	-0.08 (0.04)	0.02 (0.03)	-0.07** (0.03)	-0.09** (0.03)	-0.07 (0.04)
Option value	-0.01 (0.02)	0.01 (0.03)	-0.03 (0.04)	0.05* (0.02)	0.01 (0.03)	0.10** (0.04)	-0.04* (0.02)	-0.01 (0.02)	-0.07* (0.03)
Observations	326	150	176	326	150	176	326	150	176

## Preferences vary by gender

**For women**, rating the following as “important” vs “unimportant” predicts:

- Love for the subject: 11 pp less likely to study science
- Expected wages: 10 pp more likely to study science

**For men**, rating the following as “important” vs “unimportant” predicts:

- Peer pressure: 17 pp more likely to study science

# Eliciting subject-specific expectations

Target two groups of undergraduate students – those currently in the process of choosing a major and those in the process of applying for jobs

- Expected earnings conditional on major
- Expected grades conditional on major
- Expected occupation conditional on major

# Eliciting subject-specific expectations

For example:

- What is your expected annual income immediately after graduating college if you major in STEM?
- What is your expected GPA if you major in STEM?
- What is the probability of your working in X sector immediately after graduating college if you major in STEM?

# Eliciting subject-specific expectations

Students applying for jobs apply across a number of sectors:

- Consulting/Finance/Management
- Science/Information Technology
- Development/Policy/Government/not-for-profit
- Media/Publishing

Ask students about labour market expectations conditional on working in a particular sector and to rank their preferred sectors

## Next steps

- Identify preference parameters by gender for college and labour outcomes associated with different majors
- Identify potential biases in expected grades and labour market outcomes that can be addressed through information interventions
- Identify the perceived characteristics associated with jobs in different sectors that are driving the subject choices



# Results: summary

## Gender matters

- There are large gender gaps in wages
- These emerge from both occupational choice and major choice
  - Women select into majors associated with lower wages
  - Women in high-earnings majors select into lower paying occupations
- No evidence these are driven by differences in cognitive abilities or non-cognitive personality measures

# Results:summary

## **Preferences for majors are heterogeneous by gender**

- Women prioritise undergraduate benefits (love of subject, higher grades)
- Men are particularly influenced by parents, peers

## **Shifts in preference parameters predict shifts in major choice**

- Higher weight on expected wages shifts women into STEM and economics/business
- Higher weight on interest in subject shifts women into the humanities
- Lower weight on preferences of parents and peers shifts men into the humanities