Gender Segregation in Education and Its Implications for Labour Market Outcomes: Evidence from India

Soham Sahoo (Indian Institute of Management Bangalore)
Stephan Klasen (University of Goettingen)

UNU-WIDER Conference, Bangkok, September 2019

Research questions

- Is there gender based educational segregation prevailing at the school level?
  - Focus on stream choice at post-secondary level
  - Identify gender bias in the choice of STEM subjects

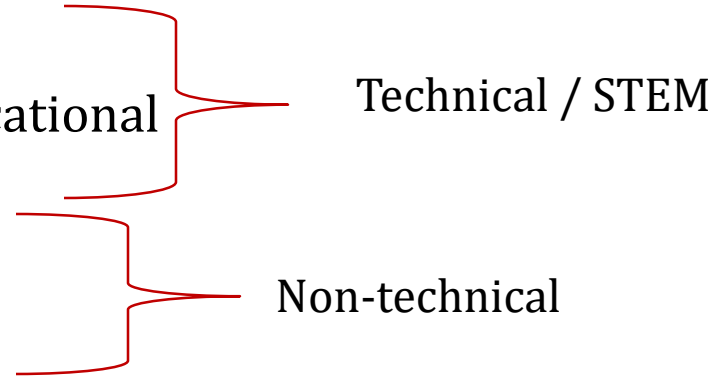
- Does gender segregation in stream choice at the post-secondary level link to labour market outcomes later in life?

Background and motivation

- School enrollment has increased in recent times, gender gap in enrollment rate has disappeared. But gender disparity may prevail in the nature of education choice.
- Economic participation of women has increased over time
  - But female labour force participation is stagnant / declining in India
  - Occupational segregation still persists: major reason behind gender gap in earnings.
- World Development Report 2012: “the seeds of segregation are planted early” and “gender differences in education trajectories shape employment segregation”.
- Considerable literature on developed countries, but not much is known about developing countries
  - India (along with China) has the highest share of STEM graduates in the world

Post-secondary education: Indian context

- Grades 11 and 12: higher-secondary level
  - First time students have to choose a stream
  - Stream choice affects tertiary education
- At the higher-secondary level (10+2), students have to choose a stream:
  - Science
  - Engineering/Vocational
  - Commerce
  - Arts/Humanities
  - Others



Data

- IHDS: household level longitudinal data: 2005 and 2012
- Almost 86 percent of households in 2005 were resurveyed in 2012
- Mainly use: India Human Development Survey-II (IHDS-II), 2011-12 is a nationally representative, multi-topic survey of 42,152 households in 1,503 villages and 971 urban neighborhoods across India.
- Advantage of panel data: use round 1 to control for past characteristics, including cognitive ability

Raw gender gap in different subjects (15-60 age group)

Identifying intra-household gender gap in stream choice

- Define technical education (*Techedu*) as STEM = 1, Arts/Humanities= 0

- A linear probability model for technical stream choice, for children in 15-18 year age group :

$$Prob(Techedu_{ih} = 1) = \alpha + \beta Female_{ih} + \gamma X_{ih} + \theta Ability_{ih} + \phi_h + \varepsilon_{ih}$$

- Importance of household fixed effects:
  - Son preferring, differential stopping behaviour: female children tend to end up in larger families
  - If technical education requires higher investment, this might show gender gap on average just because girls are from larger families where per child (education) investment is less

Accounting for cognitive ability

- Large literature on gender gap in math score
  - Mainly caused by background characteristics, psycho-social factors
  - Is it the difference in cognitive ability that drives enrollment into technical education?

- We use two different proxies for cognitive ability
  - Secondary School Leaving Certificate (SSLC) exam results (1^{st}, 2^{nd}, 3^{rd} divisions)
  - Data on *past test scores* on math, reading, and writing
    - Collected from earlier round of data (2005) on children in 8-11 years age
    - These children will be 15-18 years in second round in 2012, age-group corresponds to post-secondary schooling

Gender gap in technical stream choice

	Chose technical subjects in higher-secondary education			
	(1)	(2)	(3)	(4)
Female	-0.178***	-0.213***	-0.202***	-0.221***
	(0.013)	(0.031)	(0.063)	(0.062)
Age (years)	-0.024***	0.028	-0.069	-0.058
	(0.008)	(0.029)	(0.070)	(0.070)
Birth order	-0.036	0.035	-0.061	-0.065
	(0.025)	(0.054)	(0.122)	(0.122)
Number of siblings	-0.007	0.085	0.184	0.236
	(0.015)	(0.115)	(0.258)	(0.248)
Father's years of education	0.004**	-0.001	-0.031	-0.030
	(0.002)	(0.009)	(0.050)	(0.053)
Mother's years of education	0.022***	0.013	-0.004	0.007
	(0.002)	(0.017)	(0.043)	(0.045)
Secondary result: 1st division		0.240***		0.200
		(0.065)		(0.153)
Secondary result: 2nd division		0.127**		0.072
		(0.057)		(0.128)
Math: number			-0.016	0.011
			(0.225)	(0.188)
Math: subtraction			0.221	0.276
			(0.227)	(0.205)
Math: division			0.501**	0.563**
			(0.250)	(0.222)
Constant	0.897***	-0.157	1.709	1.262
	(0.153)	(0.561)	(1.288)	(1.278)
Observations	5,213	5,207	2,656	2,634
R-squared	0.086	0.129	0.236	0.283
Household fixed effects	No	Yes	Yes	Yes
Number of households (fixed effects)		4,653	2,515	2,496

The results are from a linear probability model taking children in the age-group of 15-18 years. Robust standard errors clustered at the household level are given in parentheses. *** p<0.01, **

Mean of Techedu is 0.5

Hence girls are about 40 percent less likely than boys to choose technical stream as compared to arts/humanities

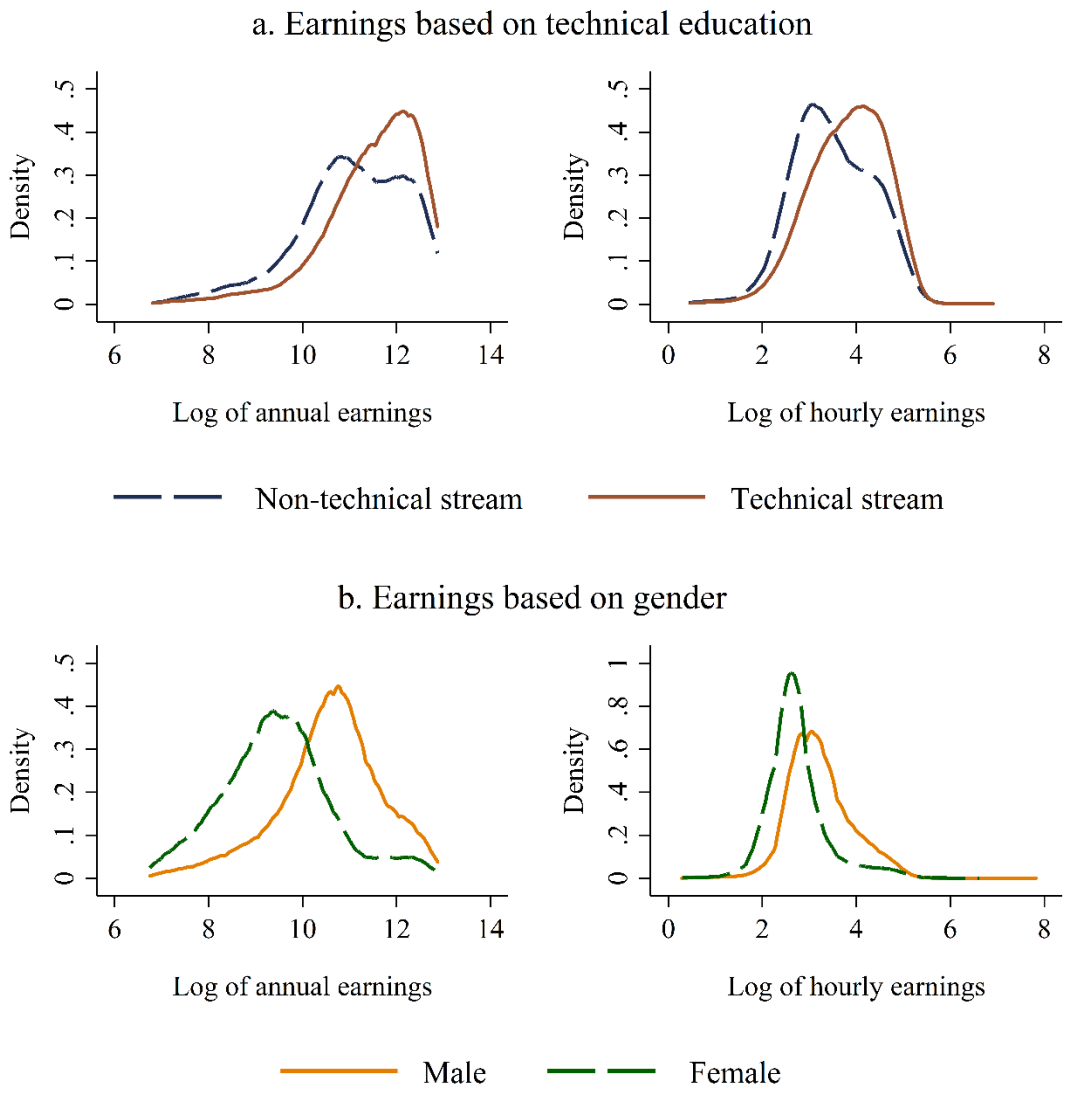
Choice of individual streams (not shown): Girls are least likely to choose science, followed by commerce.

Heterogeneity in gender gap

	Chose technical subjects in higher-secondary education					
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.168***	-0.196***	-0.166***	-0.181***	-0.251***	-0.345***
	(0.018)	(0.074)	(0.014)	(0.066)	(0.018)	(0.085)
Female * Household income per capita (baseline)	-0.001	-0.002				
	(0.001)	(0.004)				
Household income per capita (baseline)	0.003***					
	(0.001)					
Female * Educational parity between parents			0.007***	0.018**		
			(0.003)	(0.009)		
Educational parity between parents			0.018***	0.023		
			(0.002)	(0.047)		
Female * Science/technical colleges in district (number per million population)					0.008***	0.013**
					(0.001)	(0.005)
Science/technical colleges in district (number per million population)					0.003***	-0.006
					(0.001)	(0.016)
Constant	0.827***	1.331	0.896***	1.378	0.707***	-0.379
	(0.162)	(1.274)	(0.152)	(1.257)	(0.157)	(1.311)
Observations	4,622	2,634	5,213	2,634	4,998	2,583
R-squared	0.088	0.284	0.087	0.298	0.112	0.321
Other individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Secondary exam result	No	Yes	No	Yes	No	Yes
Past math score	No	Yes	No	Yes	No	Yes
Past reading & writing score	No	Yes	No	Yes	No	Yes
Household fixed effects	No	Yes	No	Yes	No	Yes
Number of households (fixed effects)		2,496		2,496		2,447

The results are from a linear probability model taking children in the age-group of 15-18 years. Robust standard errors clustered at the household level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Correlation between higher-secondary stream choice and adult-life earnings



Relationship between stream choice and labour market outcomes

- Consider all individuals in 25-60 years age-group
- Regression of labour market outcome:

$$Y_{ih} = \delta + \eta\, Female_{ih} + \lambda\, Female_{ih} * Sec_{ih} * Techedu_{ih} + \psi\, Female_{ih} * Sec_{ih} + \pi\, Sec_{ih} * Techedu_{ih} + \sum_{k} \tau_k Eduyear_{ihk} + \rho\, Ability_{ih} + \zeta X_{ih} + \sigma_h + \epsilon_{ih}$$

- Effect of technical stream choice on outcome of women:

$$E(Y \mid Female = 1, Sec = 1, Techedu = 1) - E(Y \mid Female = 1, Sec = 1, Techedu = 0) = \lambda + \pi$$

- Differential effect between men and women is captured by λ

- Focus is on intra-household differences (household FE model)
- We control for cognitive ability by using SSLC (10^{th} board exam) results

Labour force participation

	Probability of labour force participation		
	(1) Overall	(2) Rural	(3) Urban
Female	-0.586***	-0.566***	-0.643***
	(0.009)	(0.012)	(0.014)
Female * Secondary pass * Technical stream (λ)	0.090***	0.107***	0.088***
	(0.016)	(0.030)	(0.019)
Secondary pass * Technical stream (π)	0.005	-0.013	0.001
	(0.009)	(0.014)	(0.011)
Female * Secondary pass	-0.040***	-0.065***	0.027**
	(0.009)	(0.014)	(0.013)
Secondary result: 1st division	0.024*	-0.012	0.045***
	(0.012)	(0.019)	(0.017)
Secondary result: 2nd division	0.005	-0.014	0.020
	(0.010)	(0.014)	(0.014)
Constant	1.052***	1.082***	0.999***
	(0.018)	(0.023)	(0.030)
Effect of Technical stream on females (λ + π)	0.095***	0.094***	0.089***
	(0.015)	(0.029)	(0.018)
Observations	80,302	51,976	28,326
R-squared	0.665	0.646	0.704
Years of education dummies	Yes	Yes	Yes
Other individual covariates	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes
Number of households (fixed effects)	38,656	25,205	13,451

Technical education may give women access to better quality jobs: this may promote their labour force participation .

Quality of jobs matter for women (social stigma, family status concerns), but not for men's participation.

Hence the effect of technical education is larger for women, closing the gender gap.

Earnings

	Log of annual earnings					
	Without selection correction			With selection correction		
	(1)	(2)	(3)	(4)	(5)	(6)
	Overall	Rural	Urban	Overall	Rural	Urban
Female	-0.774***	-0.785***	-0.793***	-1.100***	-0.983***	-1.485***
	(0.033)	(0.041)	(0.056)	(0.118)	(0.127)	(0.270)
Female * Secondary pass * Technical stream (λ)	0.073	-0.177	0.154*	0.118	-0.137	0.282***
	(0.082)	(0.176)	(0.088)	(0.083)	(0.183)	(0.094)
Secondary pass * Technical stream (π)	0.007	0.045	0.049	-0.004	0.011	0.014
	(0.048)	(0.081)	(0.056)	(0.051)	(0.082)	(0.058)
Female * Secondary pass	0.367***	0.440***	0.406***	0.348***	0.441***	0.470***
	(0.057)	(0.098)	(0.070)	(0.058)	(0.102)	(0.083)
Secondary result: 1st division	0.139**	0.096	0.251***	0.251***	0.150	0.452***
	(0.066)	(0.106)	(0.078)	(0.078)	(0.111)	(0.105)
Secondary result: 2nd division	0.113**	0.075	0.179***	0.155***	0.092	0.266***
	(0.056)	(0.083)	(0.067)	(0.056)	(0.083)	(0.075)
Inverse Mills ratio				-1.945***	-1.305*	-3.328***
				(0.694)	(0.783)	(1.282)
Constant	10.413***	10.240***	10.869***	11.715***	11.143***	12.862***
	(0.075)	(0.093)	(0.136)	(0.457)	(0.548)	(0.776)
Effect of Technical stream on females (λ + π)	0.080	-0.132	0.203**	0.114	-0.126	0.297***
	(0.078)	(0.171)	(0.079)	(0.081)	(0.173)	(0.086)
Observations	36,760	25,382	11,378	36,677	25,350	11,327
R-squared	0.371	0.387	0.352	0.371	0.387	0.355
Years of education dummies	Yes	Yes	Yes	Yes	Yes	Yes
Other individual covariates	Yes	Yes	Yes	Yes	Yes	Yes
Household fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of households (fixed effects)	25,511	17,078	8,433	25,452	17,055	8,397

This regression considers salaried/casual wage employees in the age-group of 25-60 years. Robust standard errors (bootstrapped for columns 4–6) clustered at the household level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Discussion of the results

- Mediators: the effect on income is due to higher participation in
  - Salaried employment
  - Male dominated (hence higher paying) occupations

- Can the effects be interpreted as *causal*?
  - Women who self-select into technical education may have higher ability, hence have better labour market outcomes
  - We try to control for *cognitive ability* by including SSLC exam performance
    - This may not capture all aspects of ability, especially *non-cognitive ability*
    - We try to capture *personality traits* by including women's decision making power within household, and socio-political participation outside household: results remain unchanged

Conclusion

- In recent years, girls are 20 percentage points (= 40%) less likely than boys to choose technical education (STEM).
  - This gender gap is *not* driven by (mathematical) ability.

- Technical education at higher-secondary level has strong connection with women's adult-life labour market outcomes: employment, occupational segregation, and earnings
  - Returns to technical education seem higher for women than men, hence it closes the gender gap in labour market outcomes within households

- Gendered outcomes in adult-life can be determined at an early level, in schools. Policies should enable girls to break gender barrier and choose subjects traditionally dominated by boys.

Thank you