

# The Determinants of College Major Choice: Evidence from Iranian University Entrance Exam

Sepehr Ekbatani

University of California, Los Angeles

September 22, 2017

# What Determines Major Choice?

- Labor market outcome:  
To what extent do expectations on future earnings and unemployment affect major choice?
- Other factors:  
Distance from home, school prestige, peer effect, parents, course enjoyment...

# What Does Literature Find?

- Growing literature on the returns to education by field of study: Hastings et al. (2013) (Chile), Kirkeboen et al. (2016) (Norway), Altonji & Zimmerman (2017), Arcidiacono (2004)
- Papers on determinants of major choice and role of information: Arcidiacono et al (2012), Zafar (2008), Hastings et al (2015), Wiswall & Zafar (2014), Baker et al (2017), Montmarquette et al (2002)
- Papers on the role of non-pecuniary reasons in major choice: Beffy et al (2012), Wiswall & Zafar (2015a), Stinebrickner & Stinebrickner (2014)

- Lack of good data. Number of observations are very limited.
- Limited to surveys from single colleges
- Ex-post survey. Student is already enrolled in a major and there exist a possibility of bias in his answers.

# Main Contributions

- Large number of observations
- Ex-ante choices
- Ordered choices
- School choice vs. Major choice

# Higher Education in Iran and Data

# Number of College Students per 100'000

	1975	1985	1995	2000	2005	2010
Iran	428	398	1557	2191	3103	5217
Turkey	690	870	2043	2548	3147	4147
Egypt	1053	1711	-	3779	3309	3337
Brazil	1039	-	-	1629	2503	3158
United States	4730	5158	5442	4735	5905	6673
England	1267	1807	3167	3478	3844	3969

**Table:** Number of College Students per 100'000

Source: UNESCO Education and Literacy Statistics

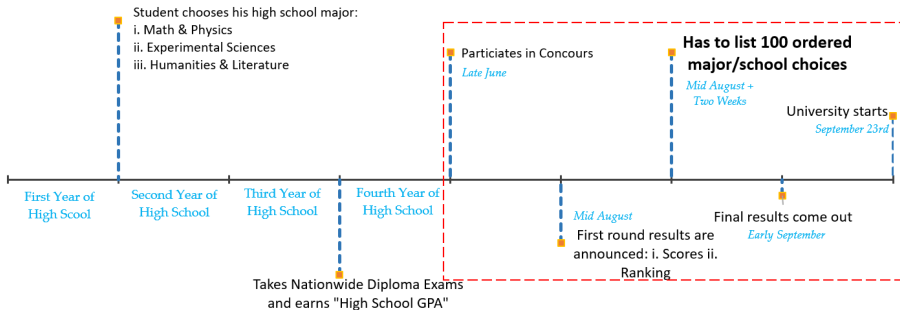
# Higher Education is a Big Deal!

- Every year around 900'000 students (60% women), who have completed four years of high school education, participate in nationwide university entrance exam (henceforth Concours).
- Private school tuition costs for the last year of high school exceeds an average worker's yearly income.
- In Tehran only, there are about 1800 test preparation institutions providing mainly afternoon classes and mock exams.
- Revenue of these institutions is estimated to be around \$2B (Almost 0.5% of Iran GDP).

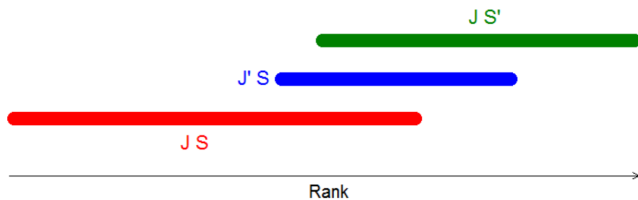




# High School Timeline



# How Students are Assigned to Majors?



- National Organization of Educational Testing (NOET)
  - Rank of student
  - Admitted school/major
  - City of residence
  - Gender
  - Scores on different courses
  - High school GPA
  - **100 ordered choices**

# Summary Statistics

Year	2009	2010	2011	2012	2013	2014	2015	2016
Number of Students	80'652	175'880	205'983	178'820	96'400	102'020	75'013	104'634
Number of Exam IDs	80'652	212'079	245'975	210'424	117'486	125'882	95'984	133'281
Average Number of Exams per Student	1	1.21	1.19	1.18	1.22	1.23	1.28	1.27
Average Number of Listings per Student	67	68	64	60	66	68	73	68
Total Number of Observations	5'401'227	14'367'653	15'663'283	12'576'816	7'719'203	8'604'346	7'010'853	9'048'905

Table: Total Summary Statistics

# Two Additional Data Sets

- Household Income and Expenditure Survey (HIES)
  - Earnings of 105 majors extracted based on Selected Characteristics of Occupations 1988 (SCO88)
- Labor Force Survey (LFS)
  - Unemployment rate for 17 broad major categories obtained based on International Standard Classification of Education 1997 (ISCED97)

# Standard Classifications

---

ISCED97	SCO88
31 Social and Behavioral Sciences	2441 Economics 2442 Sociology 2443 Political Sciences 2445 Psychology
52 Engineering	2132 Computer Engineering 2145 Mechanical Engineering 2143 Electrical Engineering
⋮	⋮

---

Table: Link between ISCED97 and SCO88

## Some Descriptive Facts



# Did Students Get What They Wanted?

		Same University		Total
		No	Yes	
Same Field	No	64.67%	6.79%	71.46%
	Yes	25.15%	3.38%	28.54%
Total		89.82%	10.18%	100%

**Table:** Share of students who got admitted to same field/university as what they listed first.

# Did Students Get What They Wanted?

		Same City		Total
		No	Yes	
Same Field	No	49.98%	21.49%	71.46%
	Yes	19.68%	8.85%	28.54%
Total		69.66%	30.34%	100%

**Table:** Share of students who got admitted to same field/city as what they listed first.

# Did Students Get What They Wanted?

		Same City		Total
		No	Yes	
Same Field	No	49.98%	21.49%	71.46%
	Yes	19.68%	8.85%	28.54%
Total		69.66%	30.34%	100%

**Table:** Share of students who got admitted to same field/city as what they listed first.

- Arcidiacono (2012):

While the coefficient on earnings was positive and significant, it was much smaller in magnitude. We believe this was because of the wording of the question, as close to 20% of individuals reported a most preferred major that was not the major they chose.

# Number of Listings

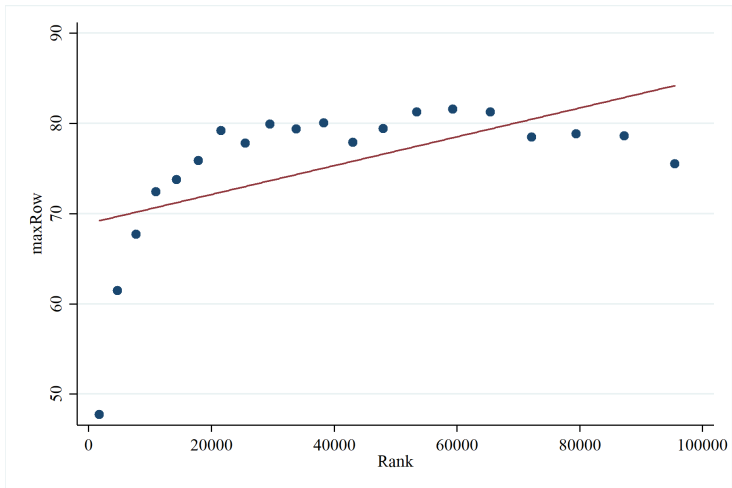


Figure: Average number of listings for different ranks

# Chance of Admittance vs. Row Number

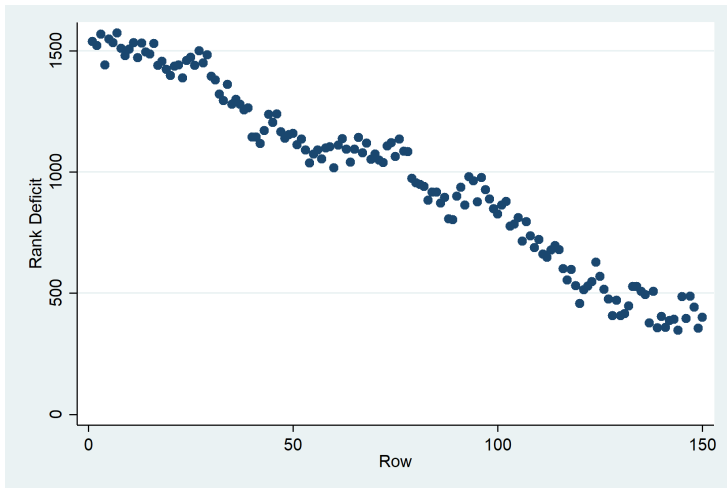


Figure: Ranks needed for admittance vs. Row number in the major choice list

# Empirical Strategy and Results

# Three Approaches

- Basic OLS
- Multinomial Logit
- Rank-ordered Logit

$$\text{Avg Rank}_{jt} = \alpha_1 \log(\text{Salary})_{jt} + \alpha_2 \text{Unemp}_{jt} + \theta_j + \gamma_t + \epsilon_{jt}$$

	Average Rank of Students		
	(1)	(2)	(3)
Log Annual Salary	-5548.4 (3500.1)	-10909.2*** (2478.06)	-895.1 (3718.2)
Unemployment Rate	1014.5*** (260.4)	-85.8 (250.9)	-335.6 (266.6)
Constant	148424.1* (64274.2)	314705.8** (53856.2)	120561.6 (75064.77)
Major Fixed Effect	No	Yes	Yes
Year Fixed Effect	No	No	Yes
Observations	571	571	571
R <sup>2</sup>	0.029	0.71	0.73

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table: Determinants of Major Popularity



# Multinomial Logit

- Following Arcidiacono (2012)
- Probability of choosing major  $j$  by individual  $i$ :

$$Pr(y_i = j|x_i) = \frac{\exp(x_i\beta_j)}{\sum_{k=1}^J \exp(x_i\beta_k)}$$

- Logarithm of relative probabilities:

$$\eta_{ij} = \ln \left( \frac{Pr(y_i = j)}{Pr(y_i = J)} \right) = x_i\beta_j$$

# MLogit Results

	ISCED97			
	Business & Administration	Engineering	Architecture & Building	Health
Female	-0.209 (-0.39)	-2.088*** (-3.78)	-1.458** (-2.73)	-1.762** (-2.89)
Mid Size Cities	-0.482 (-0.84)	-0.124 (-0.21)	-0.944 (-1.64)	0.819 (1.22)
Small Cities & Villages	-1.336* (-2.00)	-0.614 (-0.89)	-1.430* (-2.15)	1.202 (1.56)
High School GPA	0.118 (1.47)	0.405*** (4.71)	0.314*** (3.82)	1.443*** (8.85)
Age	-0.164 (-0.69)	0.00153 (0.01)	-0.0360 (-0.15)	0.671* (2.31)
Log SCO88 Salary	8.785*** (8.35)	16.41*** (13.37)	9.957*** (9.08)	17.33*** (13.67)
Constant	-160.9*** (-7.90)	-311.9*** (-13.04)	-187.2*** (-8.80)	-362.6*** (-14.33)
Observations	728			

*Humanities* is the base category. Other categories are omitted to improve readability.

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Table:** Multinomial Logit Estimation of Major Choice

# Rank Ordered Logit

- Let  $r_i = (r_{i1}, r_{i2}, \dots, r_{i100})'$  be the ordered majors for student  $i$ .  
( $r_{im} = \{j, s\}$ )
- If  $v_{ijs}$  is the indirect utility function of student  $i$  from pursuing major  $j$  in school  $s$ , probability of observing  $r_i$ :

$$Pr(r_i) = Pr(v_{ir_{i1}} > v_{ir_{i2}} \dots > v_{ir_{i100}})$$

- Assuming that unobservable preferences follow a Type I extreme value distribution:

$$Pr(r_i) = \prod_{m=1}^{99} \frac{\exp(v_{ir_{im}})}{\sum_{l=m}^{100} \exp(v_{ir_{il}})}$$

- And the log likelihood to be maximized:

$$L = \sum_{i=1}^N \log[Pr(r_i)] \quad (1)$$

# Rank Ordered Logit Results

	(1) Row	(2) Row	(3) Row
Log Annual Salary SCO88	-0.194*** (0.00147)	-0.194*** (0.00147)	-0.195*** (0.00150)
Unemployment	0.0288*** (0.000121)	0.0288*** (0.000121)	0.0287*** (0.000124)
Own State	-0.490*** (0.00185)	-0.508*** (0.00274)	-0.674*** (0.00304)
Female		0.0389*** (0.00250)	
OwnState × Female		-0.435 (.)	
Mid Size Cities			0.340*** (0.00298)
Small Cities & Villages			0.394*** (0.00384)
Own State × Mid Size Cities			0 (.)
Own State × Small Cities & Villages			-0.0406 (.)
Observations	3713118	3713118	3585197

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table: Rank Ordered Logit Estimation of Major Choice

## Conclusion and Further Work

- Fulfill the gap in the literature with the data on Iranian university entrance exam.
- Estimated a rank ordered logit model of major choice.
- Earnings expectation, chance of getting a job and distance from hometown were found as important factors for choice of major/school.
- Female students and students with lower socioeconomic status are less willing to study in a school outside their own state.

# Further Work

- Think of narrower questions that can be answered using the data. Not necessarily related to major choice.
- Think of policy evaluations that can be done using the data.
- Link to the labor market survey to find the long term effects.

- Link of our survey has been sent to students who have taken Concours in 2010. [▶ Survey](#)
- They have been in labor force for almost 2-3 years now.
- Asked about their job and income and how they did in their study as a bachelor student.
- Takes about 10-15 minutes.
- There is a prize of \$70 for 5 people.



Merci!

# Possible Future Topics

- Causal effect of social media on social variables.
  - eg. divorce, student's performance...
- Gender performance gap evaluation using RCT in online marketing industry. [▶ Link](#)

کدرشته های انتخابی ، رتبه و آخرین رتبه قبولی در سهمیه

ترتیب علاقه	کدرشته انتخابی	رتبه داوطلب	آخرین رتبه قبولی	نسخه
1	1543	62	75	* 3
2	1550	42	60	
3	1546	36	134	
4	1549	30	64	
5	1548	32	142	
6	1545	35	134	
7	1541	11	107	
8	1552	29	272	
9	1551	10	524	
10	1547	27	176	
11	1542	11	507	
12	1540	8	198	
13	2582	6	309	
14	1201	36	104	
15	1375	13	497	
16	1376	8	537	
17	1377	12	538	
18	1214	35	142	
19	1461	32	237	
20	1507	12	532	
21	1574	15	363	
22				
22				