reading group summer 24
Marry for What? Caste and Mate Selection in Modern India (2013)

Abhijit Banerjee, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune

August 2024

Outline

Introduction

Model

Setting & Data

Preference Estimation

Introduction

- how preferences for a noneconomic characteristic (e.g., caste) can affect equilibrium patterns of matching
- write a matching model and apply it to a dataset based on surveys conducted in India

Introduction

- since many women don't work in india, marriage becomes a critical economic decision
- caste remains a key determinant of marriage patterns in India
- Hitsch, Hortaçsu, and Ariely (2010) use online dating data to estimate racial preferences in the US: find strong evidence
- preference can be horizontal and/or vertical - mostly the former
- even more striking than the race analogy since this is a classic economic problem

Outline

Introduction

Model

Setting & Data

Preference Estimation

Setup

- simple marriage model (read: heteronormative) with 2-sided matching
- horizontal and vertical caste preferences
- men, women, caste $i \in 1, 2$ where 1 is higher than 2
- man's quality $x \in L, H$, woman's quality $y \in L, H$
- ω_{yi} and μ_{xi} are the number of men and women

Setup

- payoff from caste: $A(i,j) = 1 + \alpha(\beta(2-j) - \gamma(i-j)^2)$
- payoff from quality: $f(x,y)$ with $f_x(x,y) \geq 0$ and $f_y(x,y) \geq 0$ plus symmetry
- utility: $u(i,j,x,y) = A(i,j)f(x,y)$
- note: $A(i,j) > 0$, $\partial^2 A(i,j)/\partial i \partial j > 0$
- $\beta = 0$ implies purely horizontal
- $\gamma = 0$ implies purely vertical
- also assume there exists ν_{yi} and κ_{xi} people who are caste-neutral ($\alpha = 0$)
- there's lots of symmetry here, zero utility from not being matched, NTU

Conditions

- Balance: $\omega_{t1} + \omega_{t2} = \mu_{t1} + \mu_{t2}$ and $\omega_{Hi} + \omega_{Li} = \mu_{Hi} + \mu_{Li}$
- Strong Balance: $\omega_{ri} = \mu_{ri}$ and $\nu_{ri} = \kappa_{ri}$
- Limited Caste Neutrality: The number of CN H1 men is less than the number of caste conscious H2 women, and the number of CN H1 women is less than the number of caste conscious H2 men

Stable Matches

- APC: difference in average match quality of women of same quality who marry within/below caste relative to those who marry above caste
- so APC is zero under complete assortative matching or when everyone is CN
- xic is a person of type x, caste i, c preference where $c \in N, C$

Proposition 1

$$\beta_0 \equiv \frac{1}{\alpha}\left(\frac{f(H,H)}{f(H,L)} + \alpha\gamma - 1\right)$$

$$\beta_1 \equiv \frac{1}{\alpha}\left(\frac{f(H,H)}{f(H,L)}(1 - \alpha\gamma) - 1\right).$$

- All equilibria will include only assortative matches if $\beta \leq \beta_1$. If $\beta > \beta_0$, then:
  - all equilibria must have some non-assortative matching as long as condition LCN holds
  - if there is at least one non-assortative match there must be at least one out-of-caste non-assortative match
  - all out-of-caste non-assortative matches must involve an H type of caste 2 matching with an L type of caste 1
- so when preferences are mostly horizontal, out-of-caste matches will look like in-caste matches on non-caste attributes, i.e., they will be assortative

Outline

Introduction

Model

Setting & Data

Preference Estimation

Setting

- set of all matrimonial ads in sunday edition of ABP from Oct 2002 to March 2003
- parents place an ad - response via post or phone call
- one must be patient - 44 percent of the sample of ad-placers who were interviewed were married or engaged although most had placed only a single ad
- 65 percent met through an ad, the rest met through relatives or, in 20 percent of the cases, on their own

Data

- exclude Christians and Muslims
- exclude ads that list phone number - so no unobserved information
- of the remaining, randomly sampled 784 ads - asked to be interviewed (all but one agreed)
- detailed information about search process was collected - identified ads they responded to, total letters received, considered letters
- randomly sampled these letters and coded them all up
- re-interviewed a year later

Data

Figure 1: Example

Caveats

- the sample is not representative of indian marriage market
- these are mostly middle-class upper-caste Bengalis, well-educated
- almost all men and women (85%) have at least a bachelor's degree

Summary Statistics

TABLE 1—SUMMARY STATISTICS: AD-PLACERS

	Ads placed by females				Ads placed by males			
	Full set ($N = 14{,}172$)		Interviewed ($N = 506$)		Full set ($N = 8{,}038$)		Interviewed ($N = 277$)	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Number of responses, caste			22.67	19.84			82.71	76.10
Brahmin	0.26	0.44	0.27	0.44	0.27	0.44	0.25	0.44
Baidya	0.04	0.20	0.04	0.20	0.03	0.18	0.05	0.21
Kshatriya	0.02	0.13	0.02	0.12	0.02	0.13	0.01	0.12
Kayastha	0.30	0.46	0.35	0.48	0.29	0.45	0.32	0.47
Baisya and others	0.18	0.39	0.19	0.39	0.20	0.40	0.18	0.38
Sagdope and others	0.13	0.34	0.10	0.30	0.13	0.34	0.12	0.33
Other castes	0.02	0.14	0.02	0.13	0.02	0.12	0.03	0.16
Scheduled castes	0.06	0.23	0.03	0.16	0.05	0.21	0.04	0.20
Physical characteristics								
Age	26.68	3.90	26.59	3.65	31.58	4.31	32.14	4.45
Height (meters)	1.56	0.04	1.58	0.04	1.68	0.06	1.70	0.06
Skin tone	2.36	0.84	2.30	0.80				
Very beautiful	0.06	0.24	0.08	0.27				
Beautiful	0.56	0.50	0.44	0.50				
Education and income								
Less than high school	0.03	0.16	0.03	0.16	0.02	0.13	0.01	0.10
High school	0.06	0.24	0.09	0.29	0.08	0.28	0.10	0.30
Post-secondary	0.01	0.11	0.00	0.05	0.04	0.20	0.06	0.23
College	0.51	0.50	0.53	0.50	0.46	0.50	0.43	0.50
Master's	0.32	0.47	0.29	0.45	0.22	0.41	0.18	0.39
PhD	0.07	0.25	0.06	0.23	0.17	0.38	0.22	0.42
log wage	5.55	0.36	5.54	0.35	5.20	0.79	5.61	0.53
log income	9.22	0.83	8.75	0.77	9.46	0.75	9.44	0.67
Geography								
Living in Kolkata	0.51	0.50	0.80	0.40	0.50	0.50	0.76	0.43
Family from West Bengal	0.44	0.50	0.39	0.49	0.45	0.50	0.39	0.49
Demands mentioned								
Only within caste	0.09	0.29	0.10	0.30	0.10	0.30	0.09	0.28
Caste no bar	0.31	0.46	0.33	0.47	0.26	0.44	0.24	0.43
No dowry demanded	0.03	0.16	0.02	0.12	0.12	0.32	0.10	0.31

Notes: Statistics are computed only among individuals reporting a given characteristic. Statistics on the number of ads which omitted given characteristics can be found in online Appendix Table C4.

Outline

Introduction

Model

Setting & Data

Preference Estimation

Empirical Strategy

- value of spouse j to person i

$$U(X_j, X_i) = \alpha X_j + \beta f(X_i, X_j) + \mu_i + \varepsilon_{ij},$$

- authors know the following:
  - whether ad-placer follows up with letter writer
  - ranking of sample letters by ad-placer
  - ads that ad-placers responded to
  - how many responses an ad received
  - a letter writer decided to respond

Empirical Strategy II

- regressions take the form

$$y_{ij} = \alpha X_j + \beta f(X_i, X_j) + \upsilon_i + \varepsilon_{ij},$$

- conditional logit and OLS
- using ad-placer responses to letters
- y_{ij} is a dummy for whether ad-placer i responded to letter-writer j

Empirical Results - hilarious

- both women and men prefer more educated spouses
- women prefer men with higher incomes
- men prefer younger women, and women prefer men their own age
- men prefer women who describe themselves as beautiful or very beautiful
- probability of getting a call back would be higher for a very lightskinned woman without any education than for a dark-skinned woman with a college degree
- caste is important: OLS estimates indicate that a woman is 13 pp more likely to call back a prospective groom if he is from the same caste, controlling for all other attributes; man 17 pp more likely
- example: the probability to be called back is the same for a man from the same caste and no education as that for a man from a different caste with a master's degree
- men are willing to sacrifice three shades of skin tone to marry someone within their caste

Empirical Results - Heterogeneity

- hierarchical binary logit model: allows for the coefficients of our binary choice model to differ across individuals but imposes that they be drawn from normal distribution
- allow heterogeneity to depend on caste, age, height, income
- one-third of the sample appears to have no preference for marrying within caste, fraction is larger among women

Predicting matches - Simulations

- GS - men propose
- construct ordinal preferences for each woman in the sample
- use predicted utilities from the previous equation to generate this ranking

$$U_{ij}^{k} = \hat{\alpha}_k X_i + \hat{\beta}_k f(X_i, X_j)$$

- apply GS algorithm to this ranking until all men have been matched
- it does pretty well!
- the proportion of within-caste marriage falls by a large fraction when preferences are caste-blind
- the correlations in age, height, and education increase as the preferences for caste diminish
- suggests that the equilibrium price of caste ought to be low

Predicting matches: Results

	Without search frictions			With search frictions		
	Mean (1)	2.5 percentile (2)	97.5 percentile (3)	Mean (4)	2.5 percentile (5)	97.5 percentile (6)
Panel A. Simulated						
Age difference	**5.90**	**5.55**	**6.34**	**5.89**	**5.41**	**6.35**
Age correlation	**0.88**	**0.80**	**0.92**	**0.86**	**0.79**	**0.91**
Height difference	**0.11**	**0.11**	**0.12**	**0.11**	**0.11**	**0.12**
Height correlation	0.86	0.81	0.90	0.83	0.78	0.88
Same caste	0.93	0.83	0.99	0.94	0.83	0.99
Caste difference	**0.21**	**−0.02**	**0.64**	**0.19**	**−0.01**	**0.60**
Caste correlation	**0.85**	**0.50**	**1.00**	**0.86**	**0.52**	**1.00**
Same education	**0.56**	**0.23**	**0.79**	**0.55**	**0.21**	**0.78**
Education difference	−0.25	−0.50	0.00	*−0.17*	*−0.45*	*0.15*
Education correlation	**0.39**	**0.18**	**0.57**	**0.40**	**0.17**	**0.60**
Same family origin	1.00	0.99	1.00	1.00	0.99	1.00
Family origin difference	**0.00**	**0.00**	**0.01**	**0.00**	**−0.01**	**0.01**
Family origin correlation	1.00	0.98	1.00	0.99	0.97	1.00
Same residence	**0.77**	**0.30**	**1.00**	**0.72**	**0.00**	**1.00**
Location correlation	**0.24**	**−0.25**	**0.98**	**0.17**	**−0.50**	**1.00**
log wage difference	−0.33	−0.55	−0.12	−0.25	−0.52	−0.03
log wage correlation	**0.13**	**−0.21**	**0.43**	**0.12**	**−0.24**	**0.48**
Income difference	**20,855**	**−10,000**	**115,501**	**19,722**	**−11,300**	**188,000**
Income correlation	**0.29**	**−1.00**	**1.00**	**0.25**	**−1.00**	**1.00**
Quality difference	*0.14*	*0.13*	*0.15*	*0.15*	*0.13*	*0.16*
Quality correlation	**0.27**	**0.10**	**0.44**	**0.26**	**0.02**	**0.45**

Predicting matches: Results

	Considered			Matches		
Panel B. Observed						
Age difference	5.90	5.82	5.99	5.70	5.35	6.05
Age correlation	0.83	0.81	0.85	0.65	0.57	0.73
Height difference	0.12	0.12	0.12	0.12	0.11	0.13
Height correlation	0.38	0.35	0.42	0.39	0.29	0.49
Same caste	0.75	0.73	0.77	0.69	0.64	0.75
Caste difference	−0.09	−0.13	−0.05	0.01	−0.14	0.16
Caste correlation	0.85	0.82	0.87	0.76	0.69	0.83
Same education	0.45	0.43	0.47	0.44	0.38	0.50
Education difference	0.34	0.31	0.38	0.29	0.14	0.44
Education correlation	0.42	0.38	0.46	0.36	0.24	0.47
Same family origin	0.78	0.77	0.80	0.76	0.71	0.82
Family origin difference	0.01	−0.02	0.03	0.04	−0.02	0.11
Family origin correlation	0.54	0.50	0.58	0.51	0.39	0.64
Same residence	0.45	0.43	0.50	0.48	0.38	0.58
Location correlation	0.04	−0.04	0.12	−0.06	−0.22	0.21
log wage difference	0.14	0.08	0.19	0.25	0.13	0.36
log wage correlation	0.07	−0.07	0.20	0.19	−0.13	0.50
Income difference	9,277	−3,842	22,397	28,374	−16	56,764
Income correlation	0.58	0.49	0.81	0.45	0.08	0.81
Quality difference	0.10	0.10	0.11	0.12	0.11	0.13
Quality correlation	0.04	−0.24	0.34	0.20	0.07	0.32

Conclusion

- people marry within caste
- there is no reason to expect that economic growth by itself will undermine caste-based preferences in marriage
- caste-based preferences in marriage are unlikely to be a major constraint on growth (???)
- need not worry about inequality since matching is already quite assortative
- i thought this to be a very silly paper

See ya