

Plan: The Self-Perpetuating Effect of Caste Discrimination

Andrew Clausen

March 7, 2007

1 Literature Critique

Parente and Prescott (2000) document in “Barriers to Riches” that the average person in a rich country like the United States is about thirty times richer than the average person in a poor country like Nigeria, and ask “why do international incomes vary so much?” They argue that monopolies that have little incentive to adopt new technologies can account for the gap. However, their aggregate data and reduced form models can not explain why some societies have institutions that foster competition and allocate resources to innovation, and others do not. Furthermore, they can not study how these institutions change over time.

Fang (2001) argued in “Social Culture and Economic Performance” that societies with racial or social discrimination provide poor incentives for the lower class to acquire the efficient level of skills. Moreover, affirmative action policies directed at restoring equality can be counterproductive; in Malaysia, policies directed at providing the majority indigenous population with education and employment advantages over the Chinese appear to have only widened the lead of Chinese over Malay wages. Fang (2001) argues that policies giving a social group an “easy option” harms their incentives to acquire skills, and hence their social

group’s reputation. Thus discrimination remains ingrained.

India has a high degree of caste-based discrimination, with each occupation having workers drawn from a small set of subcastes known as *jatis*. Munshi and Rosenzweig (2006) consider if new industries such as information technology and financial services are breaking down the traditional caste hierarchy. Since all university education is in English, children planning a career in these industries must attend a relatively expensive English school. Thus, information about schooling choice is informative for occupation choice.

So, Munshi and Rosenzweig (2006) surveyed almost five thousand students in the Dadar neighbourhood of Bombay, recording their languages of education, and variables relevant to this choice including gender, caste, parents’ education, parents’ income, and schooling costs. They found that English education has increased across all castes and genders, but that the caste-gap for girls has narrowed more quickly than for boys. This leads them to say that “the key question is why the lower-caste boys seemingly fail to take advantage of the new economic opportunities.” (page 5)

Their explanation is that the caste system only allocates male employment, and that the new professions are therefore more attractive to women, who cannot exploit caste networks to gain employment. Their case consists of several claims: job referrals are primarily jati-based (page 4); referrals are primarily for low-paid and unskilled work (page 5); referrals are only available to men (page 9); Marathi education channels students into working class jobs (page 4); returns to English education are caste-invariant (implicit by the lack of qualification throughout – eg pages 11 – 13).

Munshi and Rosenzweig (2006) would like to have estimated the effect of jati networks on boys’ education decision by running the regression

$$\mathbb{P}(E_{ij} = 1) = \alpha R_j + X_{ij}\beta + \omega_j,$$

where $E_{ij} = 1$ if individual i in jati j is schooled in English, R_j is the fraction of the jati that received referrals, X_{ij} is the parents' schooling language, and ω_j is the human capital of individuals in jati j .

However, the level of jati networking R_j , and the jati's human capital ω_j are likely to be correlated. To isolate the referral effect, Munshi and Rosenzweig (2006) exploit the observation that job referrals are only available to men to arrive at the following regression,

$$\mathbb{P}(E_{ij} = 1) = (\alpha - \tilde{\alpha})R_j \cdot B_{ij} + X_{ij}\tilde{\beta} + X_{ij} \cdot B_{ij}(\beta - \tilde{\beta}) + \gamma B_{ij} + f_j,$$

where $B_{ij} = 1$ for boys, $\tilde{\alpha} \approx 0$ is the network effect for girls, $\tilde{\beta}$ the effect of parents' schooling for girls and $f_j = \tilde{\alpha}R_j + \omega_j$.

Assuming $\tilde{\alpha} \approx 0$, they estimate $\hat{\alpha} = -0.398$ with standard error 0.091. They conclude that caste networks are a major reason why parents choose not to send their boys to English language schools.

The claim that referrals are primarily made through caste networks is exploited in the regression through the use of R_j , which is the proportion of fathers in jati j who received referrals. If referrals were primarily through occupation-related connections instead, then R_j would be a poor measure of the child's referral opportunities. The authors do not use any evidence from their data to support this claim, but refer to other literature.

The claim that referrals are for low-paid and unskilled work is not used as far as I can tell.

The claim that referrals are only available to men appears as the assumption that $\tilde{\alpha} = 0$, which is used to separate out the network effect from the parents' education effect. The evidence in the data for this claim is that 85% of fathers received referrals, but only 15% of mothers. Different values of $\tilde{\alpha}$ can be used to get more conservative estimates of the network effect for boys.

While the regression equations described above (copied verbatim) do not include family characteristics such as parents' income, occupation and years of schooling, the regressions they report do include these variables. They clearly should be included, since they affect the opportunities and costs of schooling. Ideally, more factors, such as the number of children should have been included to avoid an omitted variable bias.

However, introducing these variables also introduces endogeneity problems; for example parents might work harder so that they can fund their children's education. In this case, unobserved characteristics such as the jati's ability might be correlated with both income and education, leading to endogeneity bias.

Moreover, the regression does not measure the level of caste discrimination the child might face in a professional occupation. Clearly, this is an important factor for education decisions as more discrimination will decrease low-caste investment in education. Moreover, discrimination is crucial for determining the extent to which the caste system is self-perpetuating. Instead, the paper either implicitly assumes that this effect is absent – that the returns from English schooling are uniform across castes – or they bundle it into the miscellaneous unobserved jati human capital term.

2 Plan

I plan to measure the effect of caste discrimination in emerging professions such as information technology and finance on education decisions. This is an important question because it measures the extent to which India's caste system is self-perpetuating. This is of general interest because it is an case study of the change of a local institution during globalization.

I plan to measure the effect of discrimination on education decisions by:

1. Estimating the returns to English schooling for each caste.
2. Estimating the effect of returns to English schooling in the regression equation.

I have obtained the Munshi and Rosenzweig (2006) survey data through the American Economic Review's data availability policy, so I can replicate their work.

There are two ways I could measure the returns to English schooling. Firstly, I could look at the pay for English educated parents in the Dadar sample. This would be a poor approximation since the sample size is too small (less than 200 of about 5000 observations across 59 jatis), and most of these parents are not employed in the emerging professions.

A better way to measure the returns to English schooling would be to examine a larger sample which includes income, occupation and jati. A cross-section would be sufficient for measuring discrimination effects. I hope that the Indian census will have enough information, but I haven't been able to determine this yet.¹

I expect my model will be similar to Munshi and Rosenzweig (2006). The main difference will be the addition of a term ϕT_j , where T_j is the expected returns to English education in jati j , and ϕ is the effect on the education decision.

Another difference is that I plan to model the inside-caste opportunities R_j by expected wages from job referrals, rather than the probability of obtaining a referral.

So far, I do not have any strategy to overcome the omitted variable bias from not knowing the number of children, and the endogeneity bias from income.

References

Fang, H. (2001). Social culture and economic performance. *The American Economic Review*, 91(4).

Munshi, K. and Rosenzweig, M. (2006). Traditional institutions meet the modern world:

¹The Indian census website only works with Windows, and I don't have access to any Windows-based computers. So I will have to wait until I get back to school.

Caste, gender, and schooling choice in a globalizing economy. *The American Economic Review*, 96(4):1225 – 1252.

Parente, S. and Prescott, E. (2000). *Barriers to Riches*. MIT Press.