Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy

By KAIVAN MUNSHI AND MARK ROSENZWEIG*

This paper addresses the question of how traditional institutions interact with the forces of globalization to shape the economic mobility and welfare of particular groups of individuals in the new economy. We explore the role of one such traditional institution—the caste system—in shaping career choices by gender in Bombay using new survey data on school enrollment and income over the past 20 years. We find that male working-class—lower-caste—networks continue to channel boys into local language schools that lead to the traditional occupation, despite the fact that returns to nontraditional white-collar occupations rose substantially in the 1990s, suggesting the possibility of a dynamic inefficiency. In contrast, lower-caste girls, who historically had low labor market participation rates and so did not benefit from the network, are taking full advantage of the opportunities that became available in the new economy by switching rapidly to English schools. (*JEL* I21, J16, O15, Z13)

The collapse of the former Soviet Union, followed by the economic and financial liberalization of the 1990s, has restructured and "globalized" many economies throughout the world. One consequence of this restructuring, which has been widely observed, is that some groups have taken advantage of the new benefits afforded by globalization, while others appear to have been left behind. This paper addresses the question of whether and how old institutions clash with the forces of globalization in shaping the response of particular groups of individuals to the new economy. Traditional institutions, such as community networks, are generally believed to play an important role in low-income countries by facilitating economic activity when markets function imperfectly. Less well understood is how traditional institutions affect the transformation of economies undergoing change, affecting in turn the distribution of benefits from macroeconomic structural reform.

We explore the role of one such traditional institution—the caste system—in shaping career choices by gender in a dynamic urban context, using new data on schooling choices and income covering the past 20 years in Bombay city, the industrial and financial center of the Indian economy. Bombay is a useful and important setting in which to study the role of institutional rigidities in a dynamic context, as the Bombay labor market was historically organized along caste lines, with individual subcastes or *jatis* controlling particular occupational niches over the course of many generations.[1] A particularly important feature of these caste networks is that they were most active in

* Munshi: Department of Economics, Brown University, Box B/64, Waterman St., Providence, RI 02912 (e-mail: munshi@brown.edu); Rosenzweig: Department of Economics, Yale University, 27 Hillhouse Ave., New Haven, CT 06520 (e-mail: mark.rosenzweig@yale.edu). Leena Abraham and Padma Velaskar of the Tata Institute of Social Sciences collaborated in the design of the survey instrument, carried out the survey, and provided many valuable insights. We are also very grateful to Suma Chitnis for her support and advice at every stage of the project. We received helpful comments from Abhijit Banerjee, David Card, Jan Eeckhout, Lakshmi Iyer, Duncan Thomas, three anonymous referees, and seminar participants at Columbia University, Harvard University, the Massachusetts Institute of Technology, IZA, University of Pennsylvania, Princeton University, UCLA, the University of Southern California, Washington University, and the World Bank. Research support from the Mellon Foundation at the University of Pennsylvania and the National Science Foundation (grant SES-0431827) is gratefully acknowledged. We are responsible for any errors that may remain.

[1] Rajnarayan Chandavarkar (1994 pp. 122, 223), for instance, describes how "[caste] clusters formed within particular trades and occupations ... [this] occupational distribution reflected neither [traditional rural] caste vocation nor the inheritance of special skills. It was produced partly by exclusionary practices by which social groups, once they

working-class occupations dominated by *lower caste men.* Women historically did not participate in Bombay's labor market and so did not benefit from the caste networks, but both men and women scrupulously adhered to the social rule of endogamous marriage within the *jati.*

Although Bombay was a predominantly industrial city for a hundred years beginning in the last quarter of the nineteenth century, in the early 1990s the liberalization of the Indian economy saw a shift in the city's economy toward the corporate and financial sectors. We study how members of different *jatis,* by gender, responded to these changes in the returns to different occupations, and we will show that the historical pattern of networking within the *jati* continues to shape gender-specific, individual responses to these new opportunities in ways that will importantly affect the future distributions of incomes, independent of pre-schooling human capital effects or liquidity constraints.[2]

Our strategy in this paper is to assess how schooling choice, measured by the language of instruction, varied across *jatis,* across boys and girls within *jatis,* and over time. We focus on schooling choice because most adults were already locked into their occupations when the unexpected economic changes occurred. Schooling choice is an important determinant of future occupational outcomes in the Bombay economy and thus reflects the contemporaneous perceptions of expected occupational returns. University education in Bombay is entirely in English, but children choose between English and Marathi (the local language) as the language of instruction at the time they enter school. Schooling in Marathi channels the child into working-class jobs, while more expensive English education significantly increases the likelihood of obtaining a coveted white-collar job. If the economic liberalization of the 1990s effectively increased white-collar incomes, and by extension the returns to English education, then (future) occupational mobility can be identified from changes in the choice of the language of instruction made by parents of school-age children. Examination of the changing patterns of schooling choice by *jati* and gender thus permits an assessment of the interactions between traditional institutions and the new realities of globalization.

Our empirical analysis is based on a survey of 4,900 households belonging to the Maharashtrian community residing in Bombay's Dadar area and a survey of the schools in that locale that we conducted in 2001–2002. Secondary schools in Bombay run from grade 1 to grade 10. The household survey was based on a stratified random sample of students who entered 28 of the 29 schools in Dadar (in the first grade), over a 20-year period, 1982–2001.[3] English is the language of instruction in ten schools in Dadar, while Marathi is the language of instruction in the remaining 18 schools.

The survey data suggest that the returns to English education, for given years of schooling, increased in the 1990s. Based on retrospective information on the annual earnings of the parents of the sampled children, we estimated the returns to English and the returns to years of schooling at five points in time from 1980 through 2000 for working adults between the age of 30 and 55.[4] Figures 1 and 2 provide the estimated returns to schooling attainment and schooling language, for men and women, respectively, in each time period. As can be seen, the returns to years of schooling increased only mildly over time for both men and women. In contrast, the English premium increased sharply from the 1980s to the 1990s for both sexes, rising from 15 percent in 1980 to 24 percent in 2000 for men and from approximately 0 percent in 1980 to 27 percent in 2000 for women. The returns to English for men increase from the mid-1980s, which is most likely due to the decline around

obtained a foothold in a particular occupation, would not admit an outsider."

[2] A recent literature has shown that historical institutions have long-run consequences for growth in low-income countries (Daron Acemoglu et al., 2001; Abhijit Banerjee and Lakshmi Iyer, 2005). These empirical findings, however, do not provide insight into the mechanisms underlying such persistence.

[3] One school refused to provide us with information on its students and will be ignored in all the discussion that follows.

[4] The details of the estimation procedure and the estimates of the returns to English and the returns to schooling (with standard errors) are provided in Munshi and Rosenzweig (2003).

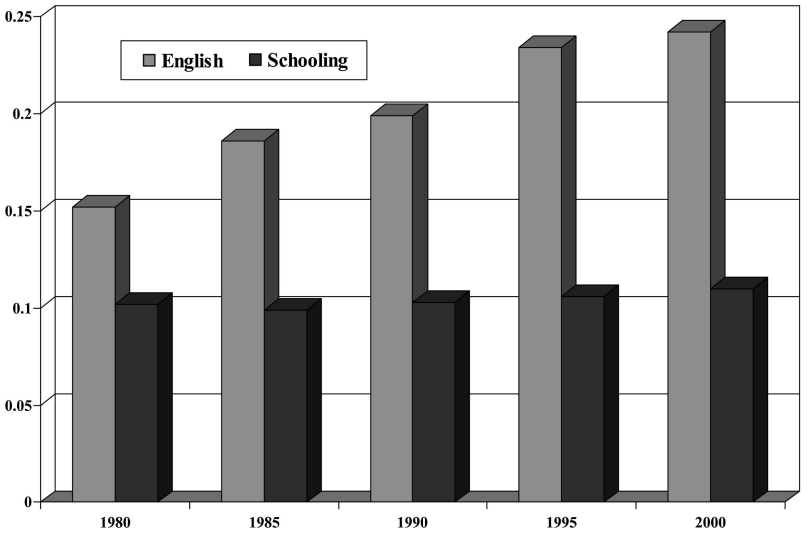

FIGURE 1. RETURNS TO ENGLISH AND SCHOOLING BY YEAR, 1980–2000: MEN AGE 30–55

that time in manufacturing jobs in Bombay (Darryl D'Monte, 2002), but continue to rise through the 1990s.

The survey collected information on schooling choice for 20 cohorts of students who entered the 28 neighborhood schools (in the first grade) over the 1982–2001 period. The time-series data on enrollments in English- and Marathi-medium schools suggest that the changes in the returns to English significantly affected schooling choice for both boys and girls in the sample, across castes and over time. Figure 3 and Figure 4 display the changing proportions of students enrolled in English schools for the 20 entering cohorts from 1982 (cohort = 1) to 2001 (cohort = 20) for three caste group-

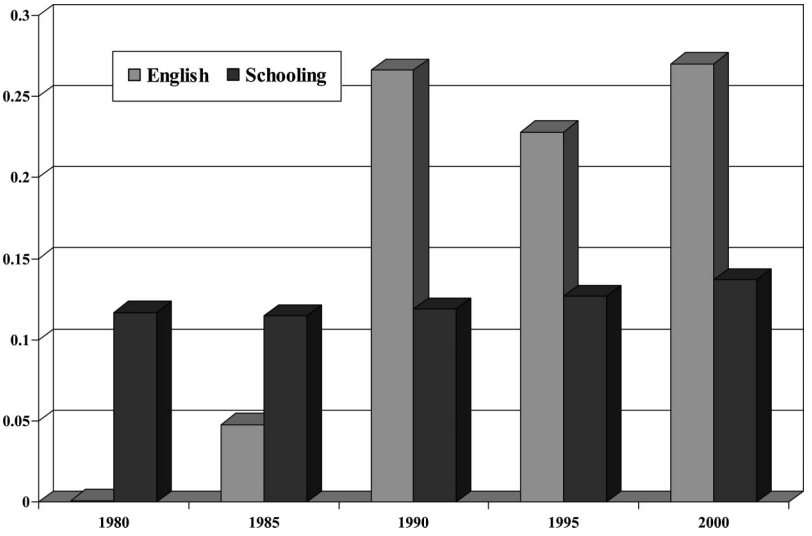

FIGURE 2. RETURNS TO ENGLISH AND SCHOOLING BY YEAR, 1980–2000: WOMEN AGE 30–55

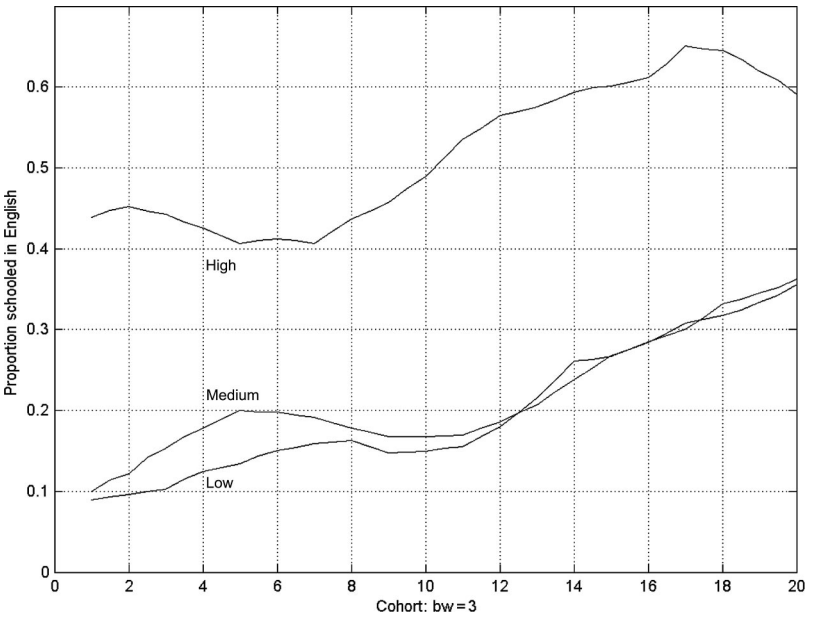

FIGURE 3. ENGLISH SCHOOLING: NET PARENTAL EDUCATION EFFECT—BOYS

ings—low, medium, and high—and by gender.[5] The figures were constructed using the Epanechnikov kernel function to nonparametrically regress schooling choice (1 = English medium; 0 = Marathi medium) on the cohort variable for each caste group, taking into account the strong intergenerational state-dependence with respect to the language of instruction within the family.[6] Although *jatis* define the relevant boundary for the labor-market networks and form the relevant social unit in our analysis,[7] we aggregate the 59 subcastes in our data for expositional convenience in these figures.

Figures 3 and 4 show that enrollment rates in English-medium schools have grown substantially over time for both boys and girls and for all castes.[8] The trajectory is much steeper, however, for the ten most recent cohorts, who would have entered school in the post-reform 1990s. Thus, the increase in the returns to English observed in Figure 1 and Figure 2 appears to have shifted schooling choice toward English education. The figures also indicate substantial differences in English schooling between castes at the beginning of the sample period, reflecting in part the circumstances of the colonial regime. The high castes gained access to clerical and adminis-

[5] Children enter first grade at the age of 6 and complete tenth grade at the age of 15, so the current age of the students in our sample, with only a few exceptions, ranges from 6 to 25. Students in Bombay typically do not change the language of instruction midstream or switch schools after they enter first grade. High castes include all the Brahmin *jatis,* as well as a few other elite *jatis* (CKP and Pathare Prabhus). Low castes include Scheduled Castes, Scheduled Tribes, and Other Backward Castes, as defined by the government of India. Medium castes are drawn mostly from the cultivator *jatis,* such as the Marathas and the Kunbis, as well as other traditional vocations that were not considered to be ritually impure.

[6] If both parents have been schooled in English, it is very unlikely that the child would be sent to a Marathi school, and all the regressions that we later report will also account for such state dependence at the level of the family. Details of the nonparametric estimation procedure and parametric estimates of the schooling regression (with standard errors) are provided in Munshi and Rosenzweig (2003).

[7] As Morris David Morris (1965 p. 76) emphasizes in his historical account of the Bombay labor market, "for any analysis of labor recruitment [in Bombay] ... it is entirely inappropriate to lump into larger groups because of similarity of name, function, social status, or region-of-origin subcastes that are not endogamous."

[8] Details of the nonparametric estimation procedure used to generate these figures are provided in Munshi and Rosenzweig (2003).

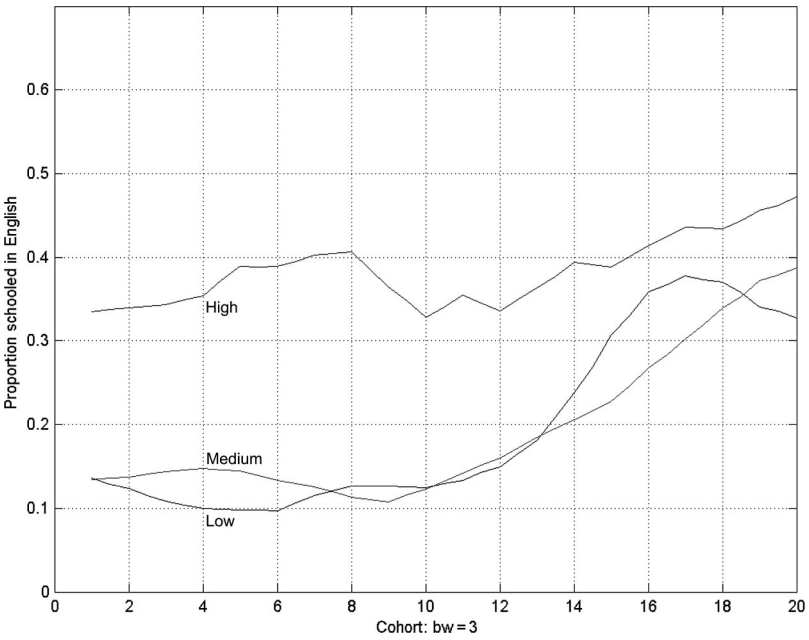

FIGURE 4. ENGLISH SCHOOLING: NET PARENTAL EDUCATION EFFECT—GIRLS

trative jobs under the British, while the lower castes were confined for the most part to working-class jobs. Consistent with the view that Marathi education channels students into working-class jobs, and that English education increases the likelihood of obtaining a white-collar job, we see in Figure 3 and Figure 4 that high-caste boys and girls currently 25 years old (the oldest cohort) were much more likely to have been schooled in English, and that this caste difference in schooling persists over the next ten cohorts. But although the caste gap narrows dramatically for the girls in the 1990s, there is no convergence for the boys. Thus, it appears that caste continues to play a role in shaping schooling choices in the new economy of the 1990s, but only for boys. The key question is why the lower-caste boys seemingly fail to take advantage of the new economic opportunities.

The gender-specific explanation for the observed pattern in Figure 3 that we pursue in this paper is based on network externalities. Numerous studies document higher levels of networking in blue-collar occupations, possibly because the information and enforcement problems that give rise to networks are more acute in those jobs.[9] These studies focus on men, the primary occupants of blue-collar jobs. And among the household heads in Dadar, 68 percent of the men in working-class jobs found employment through a relative or a member of the community, while the corresponding statistic for white-collar workers was 44 percent. Once the (working-class) network is in place, there is a positive externality associated with participation in the network, and hence with the traditional occupational choice in the *jati.* This externality could give rise to intergenerational occupational persistence *at the level of the jati,* with labor market networks channeling boys into particular (traditionally male) occupations and hence toward particular schooling choices.

[9] For example, Albert Rees (1966) found that informal sources accounted for 80 percent of all hires in eight blue-collar occupations versus 50 percent of all hires in four white-collar occupations in an early study set in Chicago. Similarly, 68 percent of blue-collar workers and 38 percent of white-collar workers reported having received help finding a job in M. S. Gore's (1970) study of migrants in Bombay.

Once the returns to the white-collar occupation grow, however, schooling choice must ultimately converge across castes. The explanation for the absence of convergence in Figure 3 that we put forward in this paper is based on the idea that the caste networks might place tacit restrictions on the occupational mobility of their members to preserve the integrity of the network. We will show that although these restrictions might have been welfare-enhancing and indeed equalizing when they were first put in place, such restrictions could result in dynamic inefficiencies when the structure of the economy changes.

The results in this paper provide empirical support for the view that historical occupation patterns kept in place by caste-based networks continue to shape occupational choice, and hence schooling choice, for the boys in the new economy. In contrast, the lower-caste girls who historically kept away from the labor market, and so have no network ties to constrain them, take full advantage of the opportunities that become available in the new economy. The growing disparities in school choices between boys and girls within the traditional *jatis* not only suggest a new balance of economic opportunities by gender, but also could threaten the long-run stability of the caste system, which is based on endogamous marriages within the subcaste. Thus, a complete understanding of the development process must not only take account of the initial conditions and the role of existing institutions in shaping the response to modernization and globalization, but must also consider how these traditional institutions are shaped in turn by the forces of change.

I. The Institutional Setting

A. *Bombay's Labor Market*

Bombay's industrial economy in the late nineteenth century and through the first half of the twentieth century was characterized by wide fluctuations in the demand for labor (Chandavarkar, 1994). It is well known that such frequent job turnover can give rise to labor market networks, particularly when the quality of a freshly hired worker is difficult to assess and performance-contingent wage contracts cannot be implemented. The presence of such recruitment networks has indeed been documented by numerous historians studying Bombay's economy prior to independence in 1947 (Chandavarkar, 1994; Morris, 1965; Alexander R. Burnett-Hurst, 1925). These networks appear to have been organized around the jobber, a foreman who was in charge of a work gang in the mill, factory, dockyard, or construction site, and more importantly also in charge of labor recruitment.

Given the information and enforcement problems associated with the recruitment of short-term labor, it is not surprising that the "jobber had to lean on social connections outside his workplace such as his kinship and neighborhood connections" (Chandavarkar, 1994, p. 107). Here the endogamous subcaste or *jati* served as a natural social unit from which to recruit labor, because marriage ties strengthen information flows and improve enforcement. This widespread use of caste-based networks thus led to a fragmentation of the Bombay labor market along social lines.[10] Other studies also suggest that these patterns tended to persist over many generations. For example, Hemalata C. Dandekar (1986) traces the evolution of a network of Jadhavs (a particular subcaste) belonging to one village in interior Maharashtra. In 1942, 67 percent of the Jadhav migrants from that village were working in the textile mills and 4 percent in other factories. Thirty-five years later, in 1977, 58 percent were still employed in textile mills, while 10 percent were in other manufacturing industries.

A noticeable feature of historical descriptions of caste-based networks in Bombay is that they were restricted to working-class jobs. This is not surprising, because the information and enforcement problems that give rise to such networks tend to be more acute in those occupations. Further, most studies of caste-based networks in Bombay focus on male workers. Women were conspicuously absent from Bombay's labor

[10] The presence of caste clusters has been historically documented in the mills (R. G. Gokhale, 1957), among dock workers (R. P. Cholia, 1941), construction workers, and in the railway workshops (Burnett-Hurst, 1925), in the leather and dyeing industries, and in the Bombay Municipal Corporation and the Bombay Electric Supply and Transportation Company (Chandavarkar, 1994). More recently, Kunj M. Patel (1963) surveyed 500 mill workers in the Parel area, close to the site of our study, in 1961–1962 and found that 81 percent of the workers had relatives or members of their *jati* in the textile industry. Sixty-six percent of the workers got jobs in the mills through the influence of their relatives and friends.

TABLE 1—SCHOOL CHARACTERISTICS AND STUDENT PERFORMANCE

	English medium	Marathi medium
School type	(1)	(2)
Panel A. School characteristics		
Student-teacher ratio	36.71	35.76
	(2.40)	(2.17)
Class size	61.90	62.28
	(3.69)	(3.15)
Students per desk	2.40	2.36
	(0.10)	(0.11)
Proportion of teachers with B.Ed.	0.72	0.70
	(0.07)	(0.05)
Proportion of teachers with higher degree	0.08	0.10
	(0.03)	(0.03)
Computers per student	0.02	0.02
	(0.004)	(0.005)
Student enrollment in secondary section	1528.40	1059.00
	(360.64)	(175.73)
Panel B. School expenses		
Fees	0.48*	0.20*
	(0.01)	(0.01)
Other expenses	1.10*	0.71*
	(0.04)	(0.01)
Panel C. SSC school-leaving exam results (1997–2001)		
Percentage passed	92.59*	51.62*
	(2.04)	(5.95)
Percentage first class among passed	36.2*	24.23*
	(1.69)	(3.35)
Percentage distinction among passed	23.94*	6.90*
	(3.92)	(1.87)
Number of schools	10	18

Notes: Standard errors are in parentheses. Panels A and C use data from the school survey; panel B uses data from the household survey. School characteristics are based on the secondary section (grades 5 through 10). School expenses are measured for 2000–2001 in thousands of 1980 Rupees. To convert to 2000 Rupees, multiply by 4.44. Other school expenses include transportation, coaching classes, textbooks, uniforms, and stationery. Scores above 35 percent are required to pass SSC; scores above 60 percent are required for first class, and above 75 percent for distinction.

* Denotes rejection of the equality of means for the two types of schools with greater than 95-percent confidence.

force, particularly in the working-class jobs (Morris, 1965). These historical patterns of labor force participation by gender will later help explain the schooling choice dynamics, for boys and girls, that we saw in Figure 3 and Figure 4.

B. *The Schools in Dadar*

Our analysis highlights the medium of instruction as the salient feature of schooling choice. It is possible that the choice of the language of instruction merely proxies for school quality. In parallel with the household survey, we carried out a survey of schools based on a questionnaire filled out by school principals. This questionnaire elicited information on a variety of school characteristics, as well as recent student performance on the standardized school leaving examination (common to both Marathi and English schools), which allows us to compare the two types of schools as well as the students across the schools. Table 1, panel A, describes school infrastructure and faculty qualifications in the English and Marathi schools. The average student-teacher ratio, class size, number of students per desk, computers per student, and the proportion of teachers with Bachelor of Education degrees and higher (postgraduate) degrees, are each very similar and statistically indistinguishable for the two types of schools.[11]

[11] A regression of the language of instruction on the set of school characteristics in Table 1, panel A, indicates that

Despite the increase in the demand for English education in the last ten years, as seen in Figures 3 and 4, no new schools were added in this period in Dadar.[12] The English-language schools accommodated this increased demand by adding divisions in each grade, increasing the number of desks in each classroom, and doubling students on each desk. Because the supply of schools was effectively fixed, we would expect the English schools to extract some economic rents from this increased demand through higher fees and schooling costs in general. In contrast, fees in the Marathi schools are subsidized by the state government. Our household survey collected information on school fees and other expenses (transportation, coaching classes, textbooks, uniforms, and stationary) in the last year. Table 1, panel B shows that school fees (in 1980 Rupees) are currently significantly higher in the English schools (480 versus 200 Rupees), as are other expenses (1,100 versus 710 Rupees).

One other difference between the schools is in the performance of the students on the Secondary School Certificate (SSC) school-leaving examination. Table 1, panel C reports student performance on this exam over a five-year period, 1997–2001. Students in the English schools perform much better on this standardized test in terms of the percentage that pass and receive a first class and a distinction.[13] Although these substantial differences in test performance can be explained by differences in school quality, they can also be explained by differential selection by ability into English and Marathi schools, as implied by our network model of school choice, and we will provide evidence supporting this implication of the model below.

In addition, our survey provides direct evidence that the medium of instruction and its implications for children's future role in society, rather than differences in school quality, dominate the schooling choice of parents. The survey elicited from parents the reasons for their choice of school for their child. The percentage of parents reporting that the "quality of education" was a factor in their choice was relatively low and did not differ substantially across parents choosing English-medium schools and Marathi schools—43.7 percent versus 35.2 percent, respectively. In contrast, almost 87 percent of parents who chose English as the medium of instruction for their child reported that career opportunities was a factor in choosing that school. And over 62 percent of parents who chose a Marathi-language school listed closer community ties as a reason.

II. A Simple Model of Schooling Choice

Our first objective in this section is to show how exogenous, historically determined occupational differences across otherwise identical *jatis* can persist when network externalities are present. Because occupational choice translates into schooling choice, this explains the initial caste gap that we observe for the boys in Figure 3 (the model that we lay out in this section applies to the boys, as we will see later that labor market networks are most active among the men). We will show, however, that *jatis* should start to converge once the returns to English grow sufficiently large, which is inconsistent with what we observed in that figure. Our second objective in this section will consequently be to show how network externalities could give rise to endogenous social restrictions on occupational mobility, and by extension schooling choice, preventing convergence across social groups in a changing economic environment.

A. *Population, Community Structure, and Market Structure*

Consider a population with a continuum of individuals. Each individual i is endowed with a level of ability $\omega_i \in \{0, \frac{1}{2}, 1\}$. Note that ability in this section, and throughout the paper, refers to pre-schooling human capital rather than genetic ability. He lives for three periods, studying in the first period and working in the remaining periods. Schooling choice is restricted to instruction in English or Marathi, the local lan-

the joint set of characteristics is not significantly different across the school types.

[12] The average establishment year for the 18 Marathi schools is 1947 and the corresponding year for the 10 English schools is 1959. All schools in the area have now been operating for many decades.

[13] Scores above 35 percent are required to pass, scores above 60 percent are required for a first class, and scores above 75 percent are required for a distinction. The same test is administered to all schools, with the questions translated into English and Marathi.

guage. Occupational choice is restricted to white-collar and working-class jobs. Education in English is required to obtain a white-collar job, but is more expensive than Marathi education, which is assumed for simplicity to be costless. Occupational choice is based on the wage that the individual will receive in the white-collar and the working-class job, net of the pecuniary cost of schooling. Each individual then makes his schooling decision based on the type of job that he (correctly) anticipates he will occupy in the subsequent period. If he prefers to hold a white-collar job, then he will study in English, if not he will study in Marathi, which is less costly.

Each individual is born into a community or *jati.* There is a large number of communities in this economy, and we normalize so that the measure of individuals in each cohort or generation of a *jati* is equal to one. To simplify and highlight the role of network externalities in intergenerational occupational persistence, we assume that the distribution of pre-schooling human capital does not vary over generations or across *jatis.*[14] Within each *jati*-generation there is a measure P_L of low types (with ability $\omega = 0$) and a measure P_M of medium types (with ability $\omega = \frac{1}{2}$).

On the demand side of this labor market, firms operate competitively in both the working-class and the white-collar sectors. We noted earlier that working-class jobs generally tend to be more heavily networked. For the purpose of this simple model, we assume that the white-collar worker's ability, and hence his productivity, can be observed perfectly and so the white-collar wage (net of schooling costs) is specified to be $\theta\omega_i$. Here θ represents the returns to ability in the white-collar job, which in our setup also reflects the returns to English education. In contrast, the nature of the production technology prevents working-class firms from directly observing their employees' ability before they commence work. We take it that the firm is unable to specify a performance-contingent wage contract, and so will use referrals from its incumbent workers to hire new employees, generating a role for the network in the working-class jobs alone. Munshi (2003) provides evidence that experienced workers contribute disproportionately to labor market networks in the United States, and we would expect this pattern to hold up in other economies as well. In our model, the expected working-class wage over the individual's working life is thus specified to be P, the proportion of the previous generation (three-year-olds) in the *jati* that will be employed in the working-class job when he enters the labor force.

[14] We will relax this assumption in the empirical work by allowing for heterogeneity in ability across *jatis.*

B. *The Schooling Equilibrium*

We now proceed to derive the different occupational distributions, and hence schooling equilibria, that can be sustained across *jatis* with the same ability distribution in this setup. Each individual chooses the occupation, and hence the language of instruction, that maximizes his net return. This return depends on his own ability, as well as the proportion of his *jati* in the previous generation employed in the working-class occupation, as described above.

Under conditions that we specify below, with three levels of ability, three distinct schooling equilibria can be sustained within *jatis:* (a) only low types choose Marathi education; (b) low and medium types choose Marathi education; (c) everyone in the *jati* chooses Marathi education.[15]

CONDITION 1: $P_L < \theta/2$.

CONDITION 2: $\theta/2 < P_L + P_M < \theta$.

CONDITION 3: $\theta < 1$.

It is easy to verify that once a *jati* is exogenously assigned a particular occupational distribution, this distribution will persist unchanged over many generations when the conditions above are satisfied.[16] This intergenerational state dependence is a consequence of the network externality associated with the working-class occupation. It implies, in turn, that the probability that any

[15] Munshi and Rosenzweig (2003) consider the general case with N types and N equilibria, without altering the results that we present below.

[16] It is merely necessary to show that no individual wishes to deviate from the occupation, and hence schooling choice, assigned to his type in his *jati* in the previous generation, for each of the schooling equilibria.

individual i drawn randomly from *jati* j will be schooled in English ($E_{ij} = 1$) is related to the proportion of men in the previous generation employed in the working-class job, P_j:

$$\Pr(E_{ij} = 1) = 1 - P_j. \tag{1}$$

This expression will serve as the starting point for the empirical analysis described in Section IV, where we will examine the relationship between schooling choice in the current generation and the occupational distribution in the previous generation, to identify the presence of an underlying network organized around the *jati.*

C. *Schooling Choice as the Returns to English Grow*

The state dependence at the level of the *jati* derived above is obtained under the assumption that the parameters of the model, P_L, P_M, θ, remain stable over time. To explore the effect of the increase in the returns to English (θ) in the 1990s, we now allow for multiple cohorts of unit measure within each generation.

If θ remains constant within a generation, the results derived above follow through without modification for all cohorts. If, however, θ increases across successive cohorts, holding P_j constant, then schooling choice within a *jati* could change over the course of a single generation. When θ just crosses one, high-ability boys belonging to *jatis* that were traditionally in equilibrium 3 switch to English. When θ subsequently reaches $2(P_L + P_M)$, medium-ability boys in *jatis* that were traditionally in equilibrium 2 or equilibrium 3 switch to English, at which point schooling choice across all *jatis* will converge.

Although the network externality described above can explain the persistence of traditional occupational patterns within the *jati* over many generations, and hence the initial caste gap observed in Figure 3, it cannot by itself explain the absence of convergence over the 1990s as the returns to English grew. To explain this absence of convergence, we consider the possibility that heavily networked (working-class) *jatis* might have put restrictions on occupational mobility, and hence schooling choice, in place to preserve the viability of the community network.[17]

To understand why restrictions on mobility might emerge, define a social welfare function that places equal weight on all members of the *jati.* Now the welfare in a *jati* situated in equilibrium 3, in which everyone studies Marathi, is simply the unweighted average of all the payoffs from the working-class occupation, $W = 1$. When θ just crosses one, in a given cohort, all high types in the *jati* can expect to earn more in the white-collar sector than in the *jati's* "traditional" working-class occupation and will thus switch to English schooling. Welfare from that cohort onward is then $W = (P_L + P_M)^2 + (1 - P_L - P_M)$. The new welfare level is a weighted average of $P_L + P_M < 1$ and 1, and so *jati*-level welfare must unambiguously decline when schooling choice, and hence the occupational distribution, shifts. Historically there was intense competition for scarce working-class jobs in Bombay, as noted in Section I. Because larger numbers improve the *jati*'s competitiveness, and increase the working class wage in general, it is easy to see why social restrictions on occupational mobility could emerge endogenously. Moreover, the fact that the lower-caste girls in our sample do not display a similar resistance to change can be attributed to the gender-specific nature of these job networks.

Social restrictions on occupational mobility can be welfare-enhancing for small and medium changes in θ, as noted above. But they could give rise to substantial inefficiencies if they continue to persist when θ grows large. For example, it is easy to verify that the social restrictions described above for equilibrium 3 will be inefficient once θ reaches $1 + (P_L + P_M)$, although a welfare calculation that identifies the presence of such a dynamic inefficiency is beyond the scope of this paper.

[17] Restrictions on mobility do not have to be associated with explicit punishment. Preferences for schooling or future career choices could be determined endogenously, for example, by placing symbolic value on the traditional occupation in the *jati.* Social interactions within the *jati* could also lead individuals to make similar schooling choices and career choices across generations. We do not attempt to distinguish between preferences that are complementary to the network, and the network itself, in this paper.

While we conjecture that restrictions on occupational mobility might be in place in the heavily networked *jatis,* no direct evidence of their presence in Bombay is available. We can, however, test one important implication that is consistent with the presence of these restrictions: the relationship between schooling choice E_{ij} and the occupational distribution within the *jati* in the previous generation P_j must not weaken over successive cohorts in the current generation, even as the returns to English grow. This stability in intergenerational state dependence would then explain the wedge between high-caste and lower-caste schooling choices for boys that was observed through the 1990s in Figure 3.

D. *Selection into Schools*

The model of schooling choice as laid out in this section also has implications for selection, by ability, into English and Marathi schools. Within any *jati,* the average pre-schooling human capital of the English students must be greater than that of the Marathi students. Taking the average across all *jatis,* this implies that average ability must be greater among the English students at any point in time. This observation is consistent with the significantly higher test scores obtained by students in the English schools (Table 1), despite the fact that English and Marathi schools appear to be similar in terms of the resources available per student and the qualifications of the teachers. But how does the ability distribution *within* the English and Marathi schools change across successive cohorts in the current generation as the returns to English grow? Without social restrictions, deviation to English education is ordered by ability, so as θ grows there is a steadily worsening pool of Marathi students. *Jatis* that begin with a greater proportion of their members in working-class jobs have higher ability among the Marathi students, but their shift into English, and hence the decline in ability, must also be more rapid, because all *jatis* ultimately converge. With social restrictions, heavily networked *jatis* continue to begin with a superior ability distribution within Marathi schools, but now there might be no convergence in ability among Marathi students across *jatis.*

Average ability among the English students is greater than average ability among the Marathi students at any point in time, but among the Marathi students it is the group with the highest ability that deviate as θ grows. Thus, while the quality of the Marathi students unambiguously declines over time as the returns to English increase, the change in the quality of the pool of English students is ambiguous.[18]

III. The Household Data

A. *The Survey*

To examine empirically the role of caste networks in shaping mobility during a period of change, we carried out a household survey based on a random sample of students, stratified by caste, who entered the 28 secondary schools in Dadar (in the first grade) over a 20-year period, 1982–2001. This design provides information for the periods before and after the major Indian economic reforms. We obtained a complete list of all students enrolled in grades 1 to 10 in 2001 (the year of the survey), as well as a list of students who were enrolled in grade 10 from 1991 to 2000. Ignoring dropouts, this leaves us with 20 cohorts of students who entered school over the 1982–2001 period. A total of 101,567 students were enrolled in the schools in 2001 or studied in grade 10 over the previous ten years. We drew the roll numbers of 20,596 students randomly from these 20 cohorts, and recovered their names and addresses from the school records. Restricting attention to Maharashtrians residing in Dadar and the immediately adjacent neighborhoods, we were left with 8,092 eligible students to serve as the sampling frame for the survey. The student's name is

[18] For example, in the three-type case with no social restrictions, some *jatis* (in equilibrium 2) have only high-ability children in English schools, while other *jatis* (in equilibrium 1) have both medium- and high-ability children in English schools to begin with. This implies that the quality of the English pool must improve when θ reaches one, because only the high types from *jatis* in equilibrium 3 deviate at that point. But average ability drops below its initial level when θ reaches $2(P_L + P_M)$, because medium and high types in all *jatis* will have switched into English schools by that point. With social restrictions, the change in ability within the English schools becomes even more difficult to characterize.

typically a good indicator of the caste, and we wanted close to 1,000 upper castes in the sample, so all 1,082 students from this population who appeared to be upper castes were selected for the survey. We drew randomly from the remaining students in the sampling frame until the target sample size was reached. The upper castes account for 17.5 percent of the final sample of 4,945 observations, which is slightly higher than the 13.4 percent that we began with in the sampling frame.

The research team interviewed the parents of the selected students at their residences. The survey instrument elicited detailed subcaste information from the respondents and included sections on grandparents' education and occupation, parents' education and occupational and income histories (at five-year intervals from 1980 to 2000), as well as the student's and siblings' subsequent education, occupation, income, and marriage outcomes (where relevant).[19] Information on transfers, assistance in finding jobs, and ties to the community was also collected.

Of the eligible households, 82.5 percent provided completed schedules. This is a relatively high response rate, especially given that some of our addresses were 20 years old. But we might still have obtained a selective sample of households, for a number of different reasons. First, households residing in Dadar who sent their children to study outside the area would be missing from the sample. Second, households who moved out of the area would be among the 17.5 percent of the respondents who did not complete the survey. And third, students from the first ten cohorts who did not reach the tenth grade, and current students who have dropped out, would be missing from the sample. In Section V, we will discuss how our identification strategy is unlikely to be undermined by these potential sources of bias.

B. *Descriptive Statistics: Caste, Occupational Networks and Schooling*

The data provide empirical support for three features of the model of schooling choice laid out in Section II. First, the occupational distribution, a product of historical circumstances, varies by caste, and persists across generations, particularly among the men. Second, working-class jobs are associated with a higher level of referrals (networking). And third, working-class jobs are associated with lower levels of English schooling.

The survey elicited information on parental occupations at five-year intervals from 1980 to 2000. For the grandparents, we simply asked for the main occupation over the individual's working life. The 90 occupations in the data were divided by roughly increasing levels of human capital into seven aggregate categories: unskilled manual, skilled manual, organized blue-collar, petty trade, clerical, business, and professional. We further classified unskilled manual, skilled manual, and organized blue-collar as working-class occupations. Clerical, business, and professional were classified as white-collar occupations. Petty trade is treated as an intermediate unclassified occupation.

Table 2, panel A, describes the occupational distribution across broad caste categories (low, medium, high), separately for the employed fathers, based on information in 1995, and the paternal grandfathers of the students in the sample. Columns 1 to 3 of the panel indicate that lower-caste fathers are much more likely to be employed in working-class occupations (54 percent and 43 percent) as compared with high-caste fathers (18 percent).[20] The same cross-caste pattern is obtained for individual occupations within the working-class and white-collar classifications, with the exception of clerical jobs. The comparison of the fathers in columns 1 to 3 with the grandfathers in columns 4 to 6 also indicates that there has been little change in the basic occupational distribution, as well as the percentage of working-class

[19] The name is usually a good indicator of the individual's community and caste. For example, 98.7 percent of the respondents, whom we had selected on the basis of their names from the school records, said that Marathi was their mother tongue, indicating that they were indeed Maharashtrian. The caste classification is potentially more problematic, however, because lower castes could in some cases change their names or misreport their caste affiliation. Note that such misreporting will not undermine the fixed effects estimation strategy, described below, as long as it does not vary by the gender of the child, within the *jati.*

[20] Note that we use only working-class and white-collar occupations when computing this statistic.

TABLE 2—OCCUPATION, EDUCATION, AND INCOME BY CASTE ACROSS GENERATIONS

Relationship to student	Parent			Grandparent		
Caste	Low	Medium	High	Low	Medium	High
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Fathers and grandfathers						
Employment (%)	97.37	97.31	99.06	98.87	98.86	99.28
Occupational distribution (%)						
Unskilled manual	11.09	7.84	4.41	9.00	3.63	2.10
Skilled manual	17.35	13.70	10.21	11.67	6.72	8.42
Organized blue-collar	22.87	19.22	2.90	22.89	24.23	7.67
Petty trade	4.00	4.51	2.52	3.11	3.20	3.34
Clerical	28.09	36.64	20.81	22.22	23.79	28.84
Business	7.95	8.79	15.51	6.11	4.72	13.00
Professional	8.30	8.79	43.51	5.56	6.18	33.66
Farming	0.35	0.51	0.13	19.44	27.53	2.97
Percent working class	53.64	42.91	18.01	56.24	49.92	19.42
	(1.23)	(1.21)	(1.38)	(1.33)	(1.40)	(1.44)
Years of schooling	9.63	10.22	13.82	—	—	—
	(0.07)	(0.07)	(0.10)			
Monthly income	1.92	1.99	4.61	—	—	—
	(0.04)	(0.04)	(0.25)			
Total number of observations	1,860	1,774	793	1,866	1,934	839
Panel B. Mothers and grandmothers						
Employment (%)	20.56	20.01	51.23	19.31	18.59	15.57
Occupational distribution (%)						
Unskilled manual	29.95	16.94	2.36	24.65	7.18	3.13
Skilled manual	8.82	8.47	6.15	1.70	1.44	3.13
Organized blue-collar	4.01	4.92	0.47	8.50	4.31	0.78
Petty trade	3.74	3.83	1.18	1.13	0.57	0.00
Clerical	31.55	40.71	46.34	4.25	2.30	19.53
Business	4.55	2.46	3.78	2.27	1.44	3.91
Professional	17.38	22.68	39.72	5.10	8.62	67.97
Farming	0.00	0.00	0.00	52.41	74.14	1.56
Percent working class	44.44	31.53	9.09	75.00	51.14	7.14
	(2.62)	(2.48)	(1.41)	(3.39)	(5.36)	(2.30)
Years of schooling	8.03	8.73	13.49	—	—	—
	(0.09)	(0.09)	(0.10)			
Monthly income	0.23	0.30	1.37	—	—	—
	(0.02)	(0.02)	(0.07)			
Total number of observations	1,887	1,954	857	1,885	1,953	854

Notes: Occupational distribution within each caste group is computed using employed individuals only. Employment for fathers and mothers is computed as of 1995. Statistics in columns 4–6 are reported for paternal grandfathers and maternal grandmothers. Working class = 1 if unskilled manual, skilled manual, organized blue-collar; 0 if clerical, business, professional. Standard errors in parentheses. Schooling and income statistics are computed using all parents in the sample, regardless of whether they are employed. Monthly income is measured in thousands of 1980 Rupees in the year closest to the year in which the child entered school.

Occupational categories

Unskilled manual: daily wage labor, deliveryman, servant, hotel worker, helper, cleaner/sweeper, porter, assistant watchman, fisherman, gardener, barber, cobbler (chambhar), unskilled laborer, seaman.

Skilled manual: machine operator, plumber, welder, technician, electrician, mechanic, carpenter, fitter/turner, tailor, painter, film developer, goldsmith, artist, priest, lab assistant, skilled worker, traditional healer (vaidhya), computer operator.

Organized blue collar: mill worker, factory worker, peon, Bombay Port Trust (BPT) worker, Bombay Electric Supply and Transportation (BEST) worker, Bombay Municipal Corporation (BMC) worker.

Petty trade: hawker, storeman (storekeeper), salesman, agent, shopkeeper.

Clerical: supervisor, driver, police, clerk, conductor, stenographer, postmaster, receptionist, foreman/draftsman, secretary.

Business: self business, medical representative, transporter, marketing, consultant, employer, contractor, politician (social worker/leader), merchant.

Professional: tutor, teacher, programmer, engineer, officer, manager, doctor, lawyer, nurse, lecturer, vice-chancellor, librarian, superintendant, director, principal, architect, salaried employee (service), chartered accountant, big businessman.

Farming: farmer, agricultural laborer.

jobs, across the generations within broad caste categories.[21]

Although most men are employed, we see that labor force participation (which includes part-time work) for the women in Table 2, panel B, is relatively low but is growing. Only 15 percent of high-caste grandmothers worked; whereas just over half of high-caste mothers entered the labor force (based on their employment status in 1995). Among the lower castes, the percentage employed remains stable at 20 percent across the generations, but notice that farming is listed as the primary occupation for a large number of working grandmothers. This suggests that urban employment must have increased sharply for the lower-caste women as well.

The occupational distribution across castes for the mothers in panel B, columns 1 to 3, displays a pattern similar to that for the fathers. Lower-caste women are much more likely to be employed in working-class occupations (44 percent and 32 percent) as compared with high-caste women (9 percent). There is an important difference, however, between men and women—although the large difference within the working-class occupations for the men was in access to blue-collar jobs for the lower castes, for women the major difference is in access to unskilled manual jobs; many of the lower-caste women work as sweepers and domestic servants.

Columns 1 to 3 and columns 4 to 6 in panel B suggest that there has been, in contrast to the men, significant intergenerational change in occupational patterns for women within castes. The urban occupations that show the greatest increase are skilled manual, clerical, and professional (with the exception of the high castes).[22] The decline in the percentage of working-class jobs among the lower-caste women, across a single generation, is particularly dramatic. This contrasts with the stability of the occupational distribution for the men, for all castes, that we noted earlier, consistent with the view that labor networks are weak among the women.

Together with the occupational distribution, Table 2 reports the mean years of schooling and monthly income separately by caste for men and women.[23] As expected, high-caste mothers and fathers have significantly more years of schooling and significantly higher incomes. Although the model in Section II assumes that the distribution of pre-schooling human capital is the same across castes (*jatis*), children in a wealthy, educated *jati* that has had access to white-collar jobs for many generations will be nurtured very differently from children in a *jati* that was historically confined to manual jobs. This suggests that pre-schooling human capital could vary in practice across broad caste categories, and across *jatis,* as well. When estimating the effect of the historical occupational distribution on the child's schooling choice, we will consequently take account of the possibility that the occupational distribution could be correlated with the ability distribution in the *jati.*

Table 3 indicates that, as assumed in the model, working-class occupations are associated with higher levels of networking (referrals).[24] Column 1 shows that 68 percent of the working-class men received help from a relative or member of the community in finding their first job (or starting their first business if self-employed), which is significantly higher

[21] The exception is farming, which is listed as the primary occupation for a large proportion of lower-caste grandfathers. This implies, in turn, that roughly one-quarter of the lower-caste fathers are first-generation migrants. Migrants are by definition newcomers in the labor market, and so will be more susceptible to the information problems that generate a need for the caste networks.

[22] The decline in the proportion of high-caste women in professional jobs is most likely because only the highest-ability women of the older generation (grandmothers) entered the labor force.

[23] Recall that income information was collected from each parent at five equal points in time from 1980 to 2000. We use the income (in 1980 Rupees) that coincides as closely as possible with the year in which the child entered school. Thus, the income in 2000 is used for students age 6 to 10, the income in 1995 for students 11 to 15, the income in 1990 for students 16 to 20, and the income in 1985 for students 21 to 25. The same income statistic is used later in the schooling regressions.

[24] The parents of the selected students were asked how they learned about their first job: through a childhood friend, through a college friend, through a relative, through a member of the community (*jati*), or by some other means (which was left open-ended in the questionnaire). This open-ended category included cases in which no help was received, or in which the job was found through newspaper advertisements, campus interviews, and other impersonal information channels. A binary referral variable was then constructed, taking the value of one if the parent learned about the first job from a relative or member of the community, and zero otherwise.

TABLE 3—REFERRALS AND SCHOOLING BY OCCUPATION

Relationship to student	Father		Mother	
Outcomes and choices	Percentage that received referrals	Percentage that studied in English	Percentage that received referrals	Percentage that studied in English
	(1)	(2)	(3)	(4)
Occupation				
Unskilled manual	65.95	0.80	61.29	0.00
Skilled manual	60.13	2.24	45.56	5.56
Organized blue-collar	76.43	0.91	69.44	5.56
All working class	68.44	1.36	57.69	2.24
(standard error)	(1.11)	(0.28)	(2.80)	(0.84)
Petty trade	57.89	1.75	61.76	2.94
Clerical	47.41	2.89	30.56	7.26
Business	49.29	8.53	41.86	9.30
Professional	32.77	11.38	29.25	14.47
All white-collar	43.76	6.20	30.64	10.13
(standard error)	(1.02)	(0.49)	(1.60)	(1.05)
Number of observations	4,515	4,513	1,215	1,215

Notes: Statistics are computed using employed individuals only. Farmers are excluded. A parent is said to have received a referral if a relative or member of the community found him/her a job. A parent is said to have studied in English if he/she studied in that language in secondary school. Occupational categories are defined in Table 2.

than the 44 percent of men in the white-collar jobs who received a referral. The corresponding statistics for the women in column 3 reveal essentially the same pattern, although the level of referrals for the women is generally lower than that for the men, perhaps because the networks for female jobs are less developed.

The model also assumes that Marathi schooling channels the student into a working-class job, while English schooling leads to the white-collar occupation. The survey elicited information on the language of instruction (English versus Marathi) for fathers and mothers, in secondary school. Columns 2 and 4 of Table 3 show that there is a clear distinction between working-class and white-collar jobs with respect to the language of instruction in secondary school. The percentage of men in working-class jobs that attended secondary school in English is just over 1 percent, compared with the 6 percent of men in white-collar jobs. In column 4, a similar pattern is obtained for the women. We have described the relationship between the broad occupational categories (working-class versus white-collar), the level of referrals, and English schooling. But inspection of Table 3 indicates that the level of referrals and English schooling vary systematically within these categories as well. Later, we will take advantage of this finer relationship between particular occupations, the level of referrals, and English schooling, to characterize the occupational distribution in the *jati.*

Table 2 suggests that lower-caste men and women are much more likely to hold working-class jobs. Combining these cross-caste patterns with the results in Table 3, it is not surprising that a much higher proportion of lower-caste men received referrals (60 percent versus 37 percent), and that these men are also much less likely to have been schooled in English (2 percent versus 12 percent). In contrast, although lower-caste women are also much less likely to be schooled in English, the level of referrals is statistically indistinguishable across castes.[25] The level of referrals is low in any case (13 percent for the lower castes and 19 percent for the high castes), especially when compared with the corresponding level for the men, and we will later establish that labor

[25] Although we noted earlier that lower-caste women who work are more likely to hold working-class jobs, which are associated with more referrals, we also saw that lower-caste women are less likely to enter the labor force. These two opposing effects appear to cancel each other, leaving little variation in the level of referrals across castes for the women.

market networks are effectively available for the men only.

IV. Empirical Analysis

A. *Specification and Identification*

The first implication of the model is that the occupational distribution in the *jati* should persist across generations when networks are active. Because schooling choice maps into occupational choice, equation (1) in Section II expressed this implication in terms of schooling choice in the current generation and the occupational distribution in the previous generation:

$$\Pr(E_{ij} = 1) = 1 - P_j.$$

Recall that $E_{ij} = 1$ if individual i belonging to *jati* j is schooled in English; $E_{ij} = 0$ if he is schooled in Marathi; and P_j is the proportion of men in the *jati* in the previous generation who are employed in working-class jobs and so in a position to provide referrals.

The particular relationship between E_{ij} and P_j in the equation above is, of course, a consequence of the modelling assumption that schooling choice maps perfectly into future occupational outcomes. More generally, we would expect to see a negative coefficient, but not necessarily with magnitude one, on P_j. The model laid out in Section II also does not allow for intergenerational state dependence in schooling choice at the level of the *household.* Moreover, we noted above that pre-schooling human capital and family incomes appeared to vary systematically across castes with different occupational backgrounds. The schooling regression that we estimate is consequently specified as

$$\Pr(E_{ij} = 1) = \alpha P_j + X_{ij}\beta + \omega_j, \tag{2}$$

where X_{ij} includes the parents' language of schooling to reflect household-level state dependence in schooling choice, as well as a cohort variable to capture the increase in the returns to English over successive cohorts in the current generation; ω_j measures unobserved or imperfectly observed pre-school human capital and family income in the *jati,* which could independently determine schooling choice.

P_j measures the proportion of men in the previous generation employed in working-class jobs. Although the model assumes that only two types of jobs—working-class and white-collar—are available, as many as 90 occupations are listed in the data. A relatively strong relationship between the level of referrals and the type of occupation was observed earlier in Table 3, and so one convenient statistic that accurately and parsimoniously describes the occupational distribution in the *jati* would be the proportion of fathers (the previous generation) who received a job referral. Working-class jobs were also associated with lower levels of English schooling (Table 3). An alternative measure of the occupational distribution in the previous generation would compute the proportion of fathers who attended English secondary schools. Most of the regressions reported in this paper will use the referrals statistic to measure P_j; English schooling levels were generally low in the previous generation and so there is substantially more variation in the referrals statistic across *jatis.* We will, however, verify that the results hold up with the English-schooling statistic as well.

Following the discussion above, we expect to find $\alpha < 0$ when networks are active and the occupational distribution persists across generations. Recall that α must also remain stable across cohorts in the current generation to explain the absence of convergence in Figure 3. Although much of the analysis treats α as constant, we will later verify that α does indeed remain stable across cohorts.

An identification problem arises when P_j and ω_j are correlated in equation (2). Although *jatis* might have been the same to begin with, we noted in the previous section that their members now have very different characteristics (income and education), depending on the type of occupation that the *jati* has historically been engaged in. A traditionally working-class *jati* could thus be associated with high P_j and low ω_j, in which case a family effect would be erroneously interpreted as a network effect because individuals with lower family resources independently select into Marathi schools.

Our solution to this identification problem exploits the fact, documented in Table 2, that networks are concentrated in working-class jobs

dominated by men. We mentioned earlier that the levels of referrals for women were relatively low, consistent with the significant change in the occupational distribution across generations for women indicated in Table 2. Thus, although the networks might affect schooling choice for the boys, they should have had little or no impact on the girls. The model in Section II then applies to boys only. Instead of using variation in the level of referrals across *jatis* to identify the presence of networks, as in equation (2), we proceed instead to exploit this gender difference in the access to job networks by pooling both sexes in the schooling regression to identify the presence of the network *within* the *jati:*

$$(3)\quad \Pr(E_{ij}=1) = (\alpha - \tilde{\alpha})P_j \cdot B_{ij} + X_{ij}\tilde{\beta} + X_{ij} \cdot B_{ij}(\beta - \tilde{\beta}) + \gamma B_{ij} + f_j,$$

where $\tilde{\alpha}$, $\tilde{\beta}$ represent the effect of the network and parents' language of schooling on the girls. B_{ij} is a dummy variable that takes a value of one for boys and zero for girls. The advantage of pooling the boys and girls is that the schooling regression can be estimated with *jati* fixed effects, $f_j \equiv \tilde{\alpha}P_j + \omega_j$. Although we can no longer identify α directly, we can obtain a consistent estimate of $\alpha - \tilde{\alpha}$, the coefficient on the $P_j \cdot B_{ij}$ interaction term. For the special case with exclusively male networks, $\tilde{\alpha} = 0$ and the coefficient on the interaction term identifies network-based occupational persistence for the boys directly. More generally, the coefficient on the interaction term provides a conservative estimate of the effect of caste-based networks on schooling choices for the boys.

The identifying assumption in this estimation strategy is that no variable $\phi_j \cdot B_{ij}$ appears in the residual of equation (3), where ϕ_j is correlated with P_j. A sufficient condition for this identifying assumption to be satisfied is that no unobserved determinant of schooling choice should vary by gender or have a differential effect on schooling choice by gender, *within* the *jati.* Later in Section V we will discuss alternative explanations for the negative and significant $\alpha - \tilde{\alpha}$ coefficient we obtain in the schooling regression. These explanations either relax the assumptions of the model, made earlier in Section II, or build on the failure of the identifying assumption. We will argue that none of these explanations fits the data quite as well as the male labor market network explanation we put forward in this paper.

B. *Caste-Based Networks and Schooling Choice*

Table 4, column 1, reports the estimates of the schooling choice regression, equation (2), for the boys. As noted, the sample covers 20 cohorts of students age 6 to 25, who entered school between 1982 (cohort = 1) and 2001 (cohort = 20). The student's cohort (1 to 20), the proportion of fathers in his *jati* who received a referral, and the father's and the mother's language of instruction in secondary school are included as regressors.

The cohort term is included in this regression to account for the increase in the returns to English over time. While the linear cohort effect we specify in Table 4 is clearly restrictive, we verify below that the estimated referral coefficient is unchanged when we allow for more flexible cohort effects. The referral coefficient is also specified to be constant over time in Table 4, and we will subsequently relax this restriction as well. For now, we see that the referral coefficient is negative and significant; children belonging to (historically) working-class and more heavily networked *jatis* are less likely to be schooled in English, consistent with the first implication of the model. The cohort effect is positive and significant, implying a shift into English over time, which is consistent with the increase in the returns to English we saw in Figure 1. Finally, the results imply that a boy is much more likely to be schooled in English if his parents were educated in that language, indicating significant state dependence in schooling choice at the level of the household.

Table 4, column 2, reports estimates from a specification that includes variables that determine the student's pre-schooling human capital as well as the household budget constraint, which could independently determine schooling choices. The parents' years of education, conditional on their language of instruction in secondary school and the level of referrals in the *jati,* are likely significant determinants of children's pre-schooling human capital. The family's access to own resources is measured by the

TABLE 4—CASTE-BASED NETWORKS AND SCHOOLING CHOICE

Dependent variable	English schooling					
Sample	Boys only		Girls only		Boys and girls	
	(1)	(2)	(3)	(4)	(5)	(6)
Referrals	−1.060	−0.377	−0.646	0.124	—	—
	(0.164)	(0.148)	(0.160)	(0.167)		
Referral - boy	—	—	—	—	−0.398	−0.464
					(0.091)	(0.105)
Cohort	0.013	0.009	0.013	0.009	0.017	0.010
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Father studied in English	0.320	0.236	0.388	0.309	—	0.301
	(0.037)	(0.033)	(0.037)	(0.026)		(0.026)
Mother studied in English	0.351	0.220	0.441	0.269	—	0.259
	(0.041)	(0.028)	(0.071)	(0.045)		(0.043)
Father's years of education	—	0.023	—	0.020	—	0.021
		(0.004)		(0.003)		(0.003)
Mother's years of education	—	0.023	—	0.026	—	0.024
		(0.003)		(0.003)		(0.003)
Family income	—	0.005	—	0.009	—	0.007
		(0.005)		(0.003)		(0.003)
Boy	—	—	—	—	0.270	0.297
					(0.049)	(0.077)
Cohort - boy	—	—	—	—	−0.002	−0.001
					(0.002)	(0.002)
Father studied in English - boy	—	—	—	—	—	−0.091
						(0.044)
Mother studied in English - boy	—	—	—	—	—	−0.044
						(0.042)
Father's years of education - boy	—	—	—	—	—	0.002
						(0.005)
Mother's years of education - boy	—	—	—	—	—	−0.001
						(0.004)
Family income - boy	—	—	—	—	—	−0.003
						(0.005)
R^2	0.173	0.274	0.146	0.272	0.163	0.299
Number of observations	2,405	2,286	2,228	2,093	4,635	4,379

Notes: Standard errors in parentheses are robust to heteroskedasticity and clustered residuals within each *jati.* English schooling = 1 if the child is sent to an English school, 0 if the child is sent to a Marathi school. Referrals measures the proportion of fathers in the *jati* who received a referral. Boy = 1 if the student is a boy, 0 if girl. Family income is measured in thousands of 1980 Rupees in the year that is closest to the year in which the child entered school. Columns 1–2: schooling choice for boys. Columns 3–4: schooling choice for girls. Columns 5–6: schooling choice for both boys and girls, including a full set of *jati* dummies.

total income of the father and the mother at the time when the child entered school.[26] Inclusion of these variables results in a substantial decline in the referral coefficient, suggesting that the level of referrals was previously proxying to some extent for unobserved, family-specific determinants of schooling choice, but it remains negative and significant. The coefficient on the cohort variable is quite stable. And the coefficients on the additional regressors all have sensible signs; the boy is more likely to be schooled in a more-expensive English-medium school if his father or mother are more educated, or if the family is wealthier.

The estimates for girls are reported in columns 3 and 4 in Table 4. Column 3 reports the estimates based on equation (2); column 4 reports the estimates from the augmented specification that adds the parents' years of schooling and family income as additional regressors. The estimated cohort effects, and the coefficients on

[26] We use the income in 2000 for students currently age 6 to 10, the income in 1995 for students 11 to 15, the income in 1990 for students 16 to 20, and the income in 1985 for students 21 to 25. All incomes are computed in 1980 Rupees.

parents' language of instruction, parents' education, and family income are similar to those for boys in columns 1 and 2. The referral coefficient, however, becomes negligible for the girls in column 4 once the observed determinants of pre-schooling human capital and access to own family resources are included. One explanation for this result is that girls receive help from the women, not the men, in their *jati.* But we noted that the level of referrals for the women is very low, across all castes. Although not reported, we also found no correlation between referrals and schooling choice, for both boys and girls, when we replaced the level of referrals for the fathers with the level of referrals for the mothers.

The results we have just described are consistent with the view that caste-based networks, net of individual and family characteristics, affect schooling decisions for the boys, but not for the girls. But up to this point, we have controlled only for unobserved ability with a limited number of family characteristics. A more robust identification strategy estimates the schooling regression with *jati* fixed effects, as in equation (3). These estimates are reported in column 5 of Table 4. As noted, only the referral-boy interaction coefficient, and not the linear referral coefficient, can now be identified. The coefficient on this term is negative and significant, and very similar to the referral coefficient for the boys in column 2. Recall from equation (3) that the coefficient on the referral-boy interaction term provides us with a direct estimate of the referral coefficient for the boys if the referral coefficient for the girls is zero. The result we obtained earlier for the girls, in column 4, suggests that this might well be the case.

The regression specification with *jati* fixed effects in Table 4, column 5, did not include family characteristics. Column 6 includes parents' language of schooling, parents' years of schooling and family income, both interacted and uninteracted with the boy dummy, as additional regressors. Including uninteracted family characteristics in the schooling regression has no effect on the estimated referral-boy coefficient by construction, once *jati* fixed effects are included. But we see that the inclusion of the family characteristics, interacted with the boy dummy, has no effect on the estimated referral-boy coefficient as well. Indeed, this coefficient is no smaller than the corresponding coefficient estimated earlier in column 5 without *any* household characteristics. This stability contrasts with the decline in the referral coefficient in Table 4, columns 1 to 4, when family characteristics were included, providing some support for the view that the *jati* fixed effects absorb much of the unobserved heterogeneity in this environment.[27]

Notice also that parents' years of schooling and family income, which had a strong influence on schooling choice for both boys and girls in columns 1 to 4, do not differentially affect schooling choice by gender (column 6). It is only the *jati*-level referral variable that has such a differential effect on schooling choice, as measured by the negative and significant referral-boy coefficient. This observation will be useful later in Section V when we consider alternative explanations for the results presented in this paper.

C. *Schooling Choice over Time*

The second implication of the model laid out in Section II is that the relationship between schooling choice and the occupational distribution in the previous generation will weaken across successive cohorts in the current generation as the returns to English grow, unless restrictions on occupational mobility are in place. To assess empirically the stability of the referral coefficient, we create 4 cohort categories that evenly divide the 20 cohorts, and then estimate the referral coefficient separately for each category.

We begin with a benchmark *jati*-fixed-effects regression, which maintains a constant referral coefficient but relaxes the restriction imposed

[27] A previous version of the paper (Munshi and Rosenzweig, 2003) reported a number of additional robustness tests. First, we accounted for occupational persistence at the level of the family by including a full set of (90) dummies for the student's father's occupation. Second, we allowed for the possibility that the scope of the network was determined by caste and the region of origin (within Maharashtra) by replacing the *jati* by the *jati*-region as the boundary of the network. Third, we dropped very large *jati*-regions (more than 250 observations) and very small *jati*-regions (fewer than 10 observations). The estimated referral-boy coefficient with these alternative specifications was shown to be very similar to what we report in Table 4.

TABLE 5—SCHOOLING CHOICE OVER TIME

Dependent variable	English schooling			
Additional regressors	Without family characteristics	With family characteristics	Without family characteristics	With family characteristics
	(1)	(2)	(3)	(4)
Referral - boy	−0.426	−0.478	—	—
	(0.090)	(0.106)		
Referral - boy - cohort1	—	—	−0.269	−0.416
			(0.168)	(0.167)
Referral - boy - cohort2	—	—	−0.352	−0.333
			(0.100)	(0.112)
Referral - boy - cohort3	—	—	−0.523	−0.540
			(0.145)	(0.143)
Referral - boy - cohort4	—	—	−0.607	−0.665
			(0.256)	(0.238)
Cohort 1	−0.261	−0.161	−0.261	−0.161
	(0.031)	(0.032)	(0.030)	(0.032)
Cohort 2	−0.231	−0.146	−0.231	−0.146
	(0.031)	(0.028)	(0.031)	(0.028)
Cohort 3	−0.161	−0.121	−0.161	−0.121
	(0.030)	(0.023)	(0.030)	(0.023)
Boy	0.236	0.261	0.338	0.364
	(0.065)	(0.091)	(0.156)	(0.149)
Cohort 1 - boy	0.033	0.031	−0.152	−0.106
	(0.038)	(0.037)	(0.209)	(0.169)
Cohort 2 - boy	0.052	0.031	−0.090	−0.153
	(0.042)	(0.035)	(0.174)	(0.151)
Cohort 3 - boy	0.041	0.041	−0.007	−0.030
	(0.032)	(0.024)	(0.117)	(0.114)
R^2	0.164	0.301	0.164	0.301
Number of observations	4,635	4,379	4,635	4,379

Notes: Standard errors in parentheses are robust to heteroskedasticity and clustered residuals within each *jati.* English schooling = 1 if the child is sent to an English school, 0 if the child is sent to a Marathi school. Referrals measures the proportion of fathers in the *jati* who received a referral. Boy = 1 if the student is a boy, 0 if girl. Cohort 1: age 21–25; Cohort 2: age 16–20; Cohort 3: age 11–15; Cohort 4: age 6–10. Column 2 and column 4 include family characteristics, separately and interacted with the boy dummy. Family characteristics include parents' language of schooling and years of education, and total family income. A full set of *jati* dummies is included in all regressions. Sample includes boys and girls.

thus far that cohort effects are linear, by including the cohort categories in Table 5, column 1. The estimated negative referral-boy coefficient is unaffected by the inclusion of the flexible cohort effect and remains very similar to the results shown in Table 4. Inclusion of the family background variables, uninteracted and interacted with the boy dummy, as additional regressors again has no effect on the estimated referral coefficient (column 2).

Table 5, column 3, allows for changes in the referral coefficient across cohort categories. All the referral-boy-cohort coefficients are negative and significant except for the coefficient on the first cohort category, which is slightly less precisely estimated. The referral coefficient is actually increasing for the later cohorts, and we can easily reject the convergence hypothesis which implies a decline in the referral effect over time. Once more, the estimated referral coefficients are robust to the inclusion of the family background variables as regressors (column 4). Although not reported here, the referral coefficient remained stable when the schooling regression was estimated with boys only, including parents' years of education and family income as additional regressors. It is this *jati*-level effect that presumably sustains the gap in schooling choice between broad caste categories observed for the boys in Figure 3.

TABLE 6—ALTERNATIVE MEASURES OF THE OCCUPATIONAL DISTRIBUTION AND SCHOOLING

Dependent variable	English schooling				Test scores		
Occupational distribution measure	Proportion of fathers schooled in English				Proportion of fathers that received a referral		
Sample	Boys only	Girls only	Boys and girls		Boys only	Girls only	Boys and girls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Occupational distribution	0.847	0.083	—	—	−23.151	−23.650	—
	(0.262)	(0.427)			(5.045)	(4.080)	
Occupational distribution - boy	—	—	0.701	0.869	—	—	−0.734
			(0.224)	(0.221)			(5.761)
Cohort	0.008	0.010	0.017	0.010	−0.505	−0.180	−0.190
	(0.002)	(0.002)	(0.002)	(0.002)	(0.134)	(0.204)	(0.223)
Boy	—	—	0.026	−0.006	—	—	3.794
			(0.028)	(0.031)			(4.357)
Father studied in English	0.217	0.301	—	0.313	4.901	2.323	1.847
	(0.032)	(0.027)		(0.026)	(1.397)	(3.711)	(4.028)
Mother studied in English	0.220	0.266	—	0.259	3.312	−2.596	−2.772
	(0.030)	(0.046)		(0.042)	(2.200)	(1.905)	(1.632)
Father's years of education	0.024	0.019	—	0.020	0.929	0.765	0.812
	(0.004)	(0.002)		(0.003)	(0.195)	(0.225)	(0.238)
Mother's years of education	0.023	0.025	—	0.023	0.617	1.074	0.984
	(0.003)	(0.003)		(0.003)	(0.221)	(0.199)	(0.200)
Family income	0.005	0.008	—	0.007	0.260	0.122	0.107
	(0.004)	(0.003)		(0.003)	(0.118)	(0.076)	(0.076)
R^2	0.275	0.272	0.162	0.298	0.322	0.334	0.354
Number of observations	2,286	2,093	4,635	4,379	849	775	1,624

Notes: Standard errors in parentheses are robust to heteroskedasticity and clustered residuals within each *jati.* The sample in columns 5–7 is restricted to children in cohorts 1–10, past the school-leaving age, who passed the SSC exam. Test scores range from 35 to 100. Boy = 1 if the student is a boy, 0 if girl. Family income is measured in thousands of 1980 Rupees in the year closest to the year in which the child entered school. Column 3 also includes cohort interacted with boy. Column 4 and column 7 also include cohort, father/mother studied in English, father's/mother's years of education, and family income, interacted with boy. Regressions pooling boys and girls (columns 3–4 and 7) include a full set of *jati* dummies.

D. *Robustness and Validation: Alternative Measures of the Occupational Distribution and Schooling*

The regression results reported thus far used the proportion of fathers who received a referral for their first job to measure the occupational distribution. We now proceed to verify the robustness of the results by repeating the schooling regressions for boys, girls, and the pooled sample with the proportion of fathers schooled in English as the measure of the occupational distribution.

The coefficients on the cohort variable and the household characteristics in Table 6, columns 1 to 2, are very similar to the estimates reported in Table 4. The coefficient on the English proportion, which can be interpreted as state dependence in schooling choice at the *jati* level, is positive and significant as expected. The coefficient on the English-boy interaction term in Table 6, columns 3 to 4, which include *jati* fixed effects, is very similar to the English coefficient for the boys in column 1, matching the results reported earlier with the referrals variable. Although not reported, once again parents' education and family income do not have a differential effect on schooling choice by gender.

The language of instruction measures the child's future occupation, and the referral statistic measures the occupational distribution in the previous generation, in most of the regressions that we report in this paper. The negative and significant referral-boy coefficient in the fixed effects regressions then reflects the persistence in the occupational distribution across generations, differentially for boys and girls within the *jati.* To validate this interpretation of the results, we proceed to replace the language

of instruction with an alternative schooling outcome as the dependent variable.

We assume that test scores depend on school quality and the pre-schooling human capital of the student. The comparison of English and Marathi schools in Table 1, and the parents' perception of these schools, indicates that school quality does not vary by the language of instruction. Under the maintained assumption that pre-schooling human capital does not vary by gender within the *jati,* this implies that the referral-boy coefficient should be close to zero in the fixed effects regression with test scores as the dependent variable. The fact that referrals have a differential effect by gender on schooling choice should be irrelevant for test scores, if school quality does not vary by the language of instruction.

Columns 5 to 7 of Table 6 replace the language of instruction with performance on the school-leaving SSC examination as the dependent variable. Referrals are once more used to measure the occupational distribution in the *jati,* to be consistent with the specifications used elsewhere in the paper. We restrict attention to the first ten cohorts (age 16–25), which have already attained school-leaving age, in these regressions. Only 17 percent of the students age 16–25 in the sample never passed the SSC examination, so we focus on the test score conditional on having passed the exam in these regressions.[28]

Table 6, column 5, restricts attention to boys, and includes the cohort, family characteristics, and the level of referrals in the *jati* as regressors. The cohort effect is negative and significant, suggesting a decline in the quality of students over time. The referral coefficient is also negative and precisely estimated, which would be the case if students from high-referral *jatis* have lower levels of pre-schooling human capital. Consistent with this interpretation, family characteristics, particularly parents' years of education, have a very large positive effect on test performance. Subsequently we repeat the exercise just described for the girls (Table 6, column 6). The cohort effect is now absent, but the coefficient on referrals remains negative and statistically significant—both boys and girls from high-referral *jatis* do less well on exams. The fixed-effects estimates, reported in Table 6, column 7, of the cohort effect, the cohort-boy interaction, and the boy dummy are not statistically significantly different from zero. More importantly, the coefficient on the referral-boy interaction term is small and statistically insignificant, in contrast to the specifications with language of instruction as the dependent variable. Caste networks affect the language but not the quality of instruction of their members.

E. *Selection into Marathi Schools over Time*

The framework laid out in Section II also has implications for the compositional change in the students who attend Marathi schools over time by *jati:* first, the pre-schooling human capital of boys entering Marathi schools should decline on average as the returns to English grow. Second, when there are no restrictions on mobility put in place to exploit network externalities, the distribution of pre-schooling ability among the boys entering Marathi schools will converge across all *jatis* over time. It is possible that such convergence across *jatis* will be absent when restrictions are in place. Note that the model has no prediction for selection by ability into English schools.

We do not have a direct measure of pre-schooling human capital. The results in Table 6 suggest, however, that, net of income, parental schooling has a positive and significant effect on school performance. In particular, father's schooling has a significant positive effect on test scores for boys and girls, and the effects do not differ significantly by the gender of the child. We thus use the father's schooling level as a proxy for pre-schooling ability.[29] The question we address is whether boys with more educated fathers increasingly exit Marathi schools and whether, and how, the rate of decline in the pre-schooling ability of boys entering Marathi schools varies by *jati.*

[28] Munshi and Rosenzweig (2003) studied the effect of referrals on the probability of success in the SSC exam and obtained results that are qualitatively the same as what we report below with the test score, conditional on success, as the dependent variable.

[29] The results reported below are essentially the same if we replace father's schooling by mother's schooling.

TABLE 7—SELECTION INTO MARATHI SCHOOLS OVER TIME

Dependent variable	Father's years of education									
Cohort	11–20					1–10				
Sample	Boys		Girls		Boys and girls	Boys		Girls		Boys and girls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Cohort	−0.697	−0.577	0.423	0.219	—	0.153	0.240	−0.543	−0.087	—
	(0.230)	(0.155)	(0.230)	(0.189)		(0.191)	(0.146)	(0.326)	(0.156)	
Cohort - boy	—	—	—	—	−0.792	—	—	—	—	0.117
					(0.252)					(0.232)
Referral - cohort	1.430	1.204	−0.596	−0.280	—	−0.147	−0.258	1.054	0.310	—
	(0.400)	(0.260)	(0.376)	(0.311)		(0.323)	(0.250)	(0.554)	(0.270)	
Referral - cohort - boy	—	—	—	—	1.469	—	—	—	—	−0.231
					(0.414)					(0.415)
Referrals	−30.991	—	−3.651	—	—	−10.256	—	−18.624	—	—
	(6.894)		(6.866)			(2.035)		(2.793)		
R^2	0.106	0.205	0.138	0.205	0.215	0.136	0.254	0.184	0.278	0.285
Number of observations	839	839	815	815	1,654	866	866	851	851	1,717

Notes: Standard errors in parentheses are robust to heteroskedasticity and clustered residuals within each *jati.* Referrals measures the proportion of fathers in the *jati* who received a referral. Column 2, column 4, column 7, and column 9 include a full set of *jati* dummies. Column 5 and column 10 include a full set of *jati* dummies, *jati*-boy dummies, and *jati*-cohort dummies. All regressions are restricted to students in Marathi schools.

To test the implications for school selectivity described above, we estimate regressions on the subsample of boys entering Marathi schools of the form

$$(4)\quad E(S_{ij}|E_{ij}=0) = \kappa + \lambda R_j + \mu C_{ij} + \nu R_j \cdot C_{ij} + \psi\omega_j,$$

where S_{ij} is boy i in *jati* j's father's years of schooling; C_{ij} is the boy's cohort; R_j measures the level of referrals in the *jati;* and ω_j measures pre-schooling ability in the entire *jati.* These terms reflect the fact that pre-schooling ability conditional on selection into Marathi is, in general, a function of ability in the *jati* and the level of referrals. The cohort terms reflect the change in this selection process over time. The model, which ignores the variation in ability ω_j across *jatis,* predicts that $\lambda > 0$, $\mu < 0$, $\nu < 0$ without restrictions, and (possibly) $\nu > 0$ with restrictions.

Because the major shift into English schooling occurred in the 1990s, we first estimate equation (4) for the boys in cohorts 11–20. The estimates are reported in Table 7, column 1. The cohort coefficient is negative and significant as predicted, which implies that the pre-schooling human capital of the boys who entered Marathi schools was declining substantially in the 1990s. The coefficient on the referral-cohort interaction term is positive, consistent with the results in Table 5 showing that restrictions on mobility in the high-referral networks were still in place during this period—the more-able boys in high-referral *jatis* were shifting to English-medium schools at lower rates.

The referral coefficient is negative and significant, consistent with the results in Table 6, which indicate that *jatis* with higher referrals R_j have lower ability ω_j; the negative $R_j - \omega_j$ correlation appears to dominate the positive selection $\lambda > 0$ effect in this case. But this tells us that the positive referral-cohort coefficient that we reported above might also be spurious. To assess the robustness of the results in column 1, we add *jati* fixed effects, which subsume $\kappa + \lambda R_j + \psi\omega_j$ (Table 7, column 2). The referral coefficient λ is no longer identified, but the estimated cohort and referral-cohort coefficients are very similar to the results in column 1.

The within-*jati* estimates allow ability to vary across *jatis* but assume that ability is constant over time (both within and across generations). The level of parental schooling could, however, also depend on the access to education, which might have changed over time. If there was convergence in the access to education across *jatis* in the parent generation, that could explain the positive referral-cohort coefficient in columns 1 to 2 without requiring networks to be active. One test to rule out this alternative interpretation of our result would be to estimate the school selectivity regression for girls rather than boys; we have already seen that the network has no effect on schooling choice for the girls, and so both the cohort and the referral-cohort effect should be absent. In contrast, if the referral-cohort term is picking up convergence in (fathers') schooling levels across *jatis,* then this coefficient should be positive and significant for the girls as well.

Table 7, column 3, reports the basic selectivity regression for the girls attending Marathi schools with cohort, referral-cohort, and referrals included as determinants of father's schooling, while Table 7, column 4, repeats this regression with *jati* fixed effects. The referral coefficient in column 3 is again negative (but insignificant), consistent with the lower levels of ability in high-referral *jatis.* The cohort coefficient is positive but insignificant. More importantly, the referral-cohort coefficient is small in magnitude and statistically insignificant—the point estimate is actually negative—and consistent with the results obtained earlier that girls in families belonging to high-referral *jatis* are not restricted in their mobility.

An alternative strategy to control for the confounding effect of changes in access to schooling among the fathers across *jatis* and over time pools boys and girls in the selectivity regression, which can then be estimated with a full set of *jati* dummies interacted with the cohort variable

$$(5)\quad E(S_{ij}|E_{ij}=0) = (\mu-\tilde{\mu})C_{ij}\cdot B_{ij} + (\nu-\tilde{\nu})R_j\cdot C_{ij}\cdot B_{ij} + f_j + g_j\cdot B_{ij} + h_j\cdot C_{ij},$$

where $\tilde{\mu}$, $\tilde{\nu}$ are the coefficients on the cohort variable and the referral-cohort interaction for the girls, and B_{ij} is a boy dummy as before. The fixed effects, f_j, which subsume $\tilde{\kappa}+\tilde{\lambda}R_j+\psi\omega_j$, allow for the possibility that ability varies across *jatis.* The fixed effects interacted with the boy dummy $g_j\cdot B_{ij}$, which subsume $(\kappa-\tilde{\kappa})B_{ij}+(\lambda-\tilde{\lambda})R_j\cdot B_{ij}$, also allow ability to vary by gender across *jatis.* Finally, the fixed effects interacted with the cohort variable $h_j\cdot C_{ij}$, subsume $\tilde{\mu}C_{ij}+\tilde{\nu}R_j\cdot C_{ij}$ and control for changes in access to schooling for the fathers both across *jatis* and over time.

For the special case with $\tilde{\mu}=0$, $\tilde{\nu}=0$, as is consistent with the model, the estimated coefficients in equation (5) should match the cohort coefficient and the referral-cohort coefficient when equation (4) is estimated with *jati* fixed effects for boys only. Table 7, column 5, suggests that this is indeed the case: the cohort-boy coefficient is negative and significant, the referral-cohort-boy coefficient is positive and significant, and the point estimates are very similar to the corresponding coefficients in columns 1 and 2. These results confirm that in the most heavily networked *jatis,* high-ability girls were exiting to English-medium schools at significantly faster rates than were boys.[30] The 0.1 quantile–0.9 quantile of the referrals distribution ranges from 0.2 to 0.7. The point estimates in column 5 thus suggest that over the period of the 1990s the gap in father's schooling between boys and girls schooled in Marathi grew by 2.3 years in the highest-referral *jatis* (at the 0.9-quantile level). In contrast, the ability-differential measured by the difference in the father's schooling between boys and girls, declined by as much as five years in the low-referral *jatis* (at the 0.1-quantile level) over the same period. This increasing mismatch in ability levels between the sexes within *jatis* and school types could have important implications for the future stability of the caste system, which relies on endogamous marriage, as discussed below.

Columns 6 to 10 of Table 7 report the estimates of the selectivity equations for the first ten cohorts of students, who entered school in

[30] The negative cohort-boy coefficient implies that the boy-girl pre-schooling human capital differential is declining over time, independent of the influence of the male job network. This may be due to differences in labor force participation or changes in the returns to English by gender (as in Figures 1 and 2).

the 1980s. Schooling choices were stable over this period and thus we do not expect to find changing selectivity effects for the boys or the girls. As before, the referral coefficient, in column 6 and column 8, is negative and significant, reflecting the persistent differences in ability across *jatis.* As expected, however, and in contrast to the cohorts making schooling choices in the post-1990s new economy, the cohort effect and the referral-cohort effect, both uninteracted and interacted with the boy dummy, are insignificant in the pre-reform period.

F. *Alternative Interpretations of the Empirical Results*

The discussion that follows considers alternative explanations for the results we have presented. The identifying assumption in the fixed effects schooling choice regression is that unobserved determinants of schooling choice that are correlated with the occupational distribution should not vary by gender, or have a differential effect by gender on schooling choice, within the *jati.* Some of the alternative explanations we pursue are associated with the failure of this identifying assumption. Other explanations are generated by relaxing the assumptions of the model. We will argue that none of these alternative explanations matches all the results as well as our preferred explanation, based on underlying male labor market networks.

Liquidity Constraints.—The model assumes that schooling choices are based entirely on the individual's ability and the historical occupational distribution in his *jati,* which determines the labor market network that he inherits. When credit markets function imperfectly, liquidity constraints could, in addition, prevent individuals belonging to working class *jatis* from choosing more expensive English schooling. Liquidity does not vary by the gender of the child *within* the *jati,* and so the *jati* fixed-effects regression would appear to rule out this alternative explanation. Boys and girls have different labor market opportunities, however, and it is thus conceivable that liquidity could have a differential effect on schooling choice by gender. The schooling regression accounts for liquidity constraints by including family income at the time the child entered school. Schooling expenses for the children in our sample are relatively low (6.3 percent of family income in English schools and 6.0 percent in Marathi schools), and, not surprisingly, family income has a relatively weak effect on schooling choice for both boys and girls. Moreover, the effect of family income on school choice does not differ by gender at conventional levels of significance (Table 4).

Differences in Ability.—The *jati*-level fixed effects absorb all variation in the *jati* that is not gender specific. But in an economy where men and women historically performed very different roles, the parental and societal inputs that boys and girls received in childhood might have been very different. The results reported earlier, however, provide no evidence of gender distinctions in pre-schooling human capital within households or *jatis.* The estimates reported in Table 4 do not reject the hypothesis that the effects of parental human capital characteristics on school choice are equal for boys and girls. The results reported in Table 6 with test performance as the dependent variable are also consistent with the assumption in the fixed effects schooling regression that pre-schooling human capital does not vary by gender within the *jati.*

Discrimination.—Unless there is a gender-based component to caste discrimination, it will be subsumed entirely by the *jati* fixed effects. But it is possible that firms or schools discriminate against boys from working-class backgrounds, perhaps because they are difficult to discipline, while treating girls from different backgrounds more equally. The referral-boy coefficient would proxy for underlying discrimination in that case.

Recall that household characteristics, such as parental education and family income, had the *same* effect on schooling choice for boys and girls within the *jati.* It was only the *jati*-level referrals statistic that had a gender-specific effect on schooling. If these results are attributed to discrimination, then it implies that firms or schools do not discriminate by family background *within* the *jati,* but by *jati* affiliation alone. It is not obvious why we would expect to see gender discrimination purely along caste lines. Family characteristics, such as parental education and income, were seen to be correlated with pre-schooling human capital and are

at least as easy to observe as caste identity. Historically there does not appear to have been a policy of caste discrimination by employers in any industry in Bombay in any case (Morris, 1965).

Restrictions on High-Caste Women.—We focused in Figure 3 and Figure 4 on the absence of convergence for the boys, which was attributed to restrictions on occupational mobility among the lower castes. An alternative interpretation of these figures is that the ability distribution varies across the population such that it remains optimal for individuals to sort by caste into different careers, even as the returns to English grow. The convergence among the girls with this alternative interpretation is attributed to restrictions on the *high-caste* girls.

There is no evidence that such restrictions are in place, or have been in place historically. High-caste women in Bombay have always had higher labor-force participation rates and more English schooling than lower-caste women, as observed in Table 2. Within the high castes, boys are substantially more likely to be schooled in English than girls, and so the girls could easily switch into English schools without creating a mismatch on the marriage market.

Moreover, although the pooled schooling choice regression with fixed effects cannot distinguish between the alternative explanation, based on female restrictions, and our view that male networks shape schooling choice for the boys alone, recall that we also ran regressions separately for boys and girls. The *jati*-level statistic, measured either by the proportion of referrals or the proportion of fathers with English schooling, affects schooling choice for the boys but not for the girls in these regressions. We thus appear to be picking up restrictions on mobility that are specific to the boys.

Sampling Bias.—We noted three potential sources of sampling bias in Section IIIA. First, particular households might school their children outside the Dadar area. Second, particular households might have moved from Dadar over the past 20 years. Third, children from particular households might have dropped out of school.

The first two sources of sampling bias are easily accommodated in the fixed-effects regression framework. Although school locations and out-migration might vary by *jati,* there is no reason to expect these decisions to vary by the gender of the child *within* the *jati.* The third source of sampling bias is potentially more problematic, because drop-outs could vary by gender within the *jati.* The decision to drop out would depend on the child's pre-schooling human capital and future employment opportunities, both of which determine schooling choice. Selective dropouts, by gender across *jatis,* could consequently violate the identifying assumption underlying the fixed-effects estimation procedure.

However, the sex-ratio of students in the most recent eight cohorts (grades one through eight) in which there would be relatively few dropouts is statistically indistinguishable from the sex-ratio in the older 12 cohorts. Regressions not reported here also reveal that the sex-ratio is uncorrelated with the level of referrals in the *jati,* both in the first 12 cohorts and in the 8 most recent cohorts.[31]

V. Conclusion

As modernization proceeds around the world, there is a perception that indigenous existing institutions may importantly shape the course of the development process across different countries. Yet little is known about how such institutions actually affect the transformation of economies undergoing change, or their impact on the economic mobility of particular groups of individuals. This paper examines the role of one long-standing traditional institution—the Indian caste system—in shaping career choices by gender in a rapidly globalizing economy.

We have found that male working-class networks, organized at the level of the subcaste or *jati,* continue to channel boys into traditional occupations, despite the fact that returns to nontraditional (white-collar) occupations have risen substantially during the post-1990s reform period. In contrast, girls, who have had historically low labor-market participation rates and few network ties to

[31] As an additional test, we also verified that the sex-ratio in the most recent cohort that entered school in 2001 is uncorrelated with the level of referrals in the *jati.* Thus, there does not appear to be selective enrollment by gender and *jati* either.

constrain them, appear to be taking full advantage of the opportunities that have become available in the new economy. It is generally believed that the benefits of globalization have accrued disproportionately to the elites in developing countries. In this setting we find, instead, that a previously disadvantaged group (girls) might surpass boys in educational attainment and employment outcomes in the future in the most heavily networked *jatis.*

Although we have focused on how traditional institutions shape the responses of particular groups of individuals to the new opportunities that accompany globalization, our findings suggest that these institutions are likely to be affected in turn by the forces of change. In our framework, an individual schooled in English no longer needs the traditional caste network; indeed, it has been remarked that "the English educated form a caste by themselves" (M. P. Desai, quoted in Julian Dakin et al., 1968 p. 24). Simple statistics on marriage and migration that we computed for the elder siblings of the students in our sample would appear to support the view that English education will ultimately undermine the caste network. Among the 825 married siblings in our sample, 11.9 percent married outside their *jati.* This contrasts with the parent generation, in which only 3.7 percent of the partners were not members of the same *jati.* Schooling in English appears to be contributing to this increase in inter-caste marriage, as 31.6 percent of the English-educated siblings married outside their *jati,* versus only 9.7 percent of the Marathi-educated siblings. And among the 1,073 siblings who are currently employed, 13.9 percent of the English-educated work outside Maharashtra, versus only 2.1 percent of the Marathi-educated (these differences between the Marathi-educated and the English-educated are statistically significant at the 5-percent level). Both marriage outside the *jati* and out-migration weaken caste ties and the caste network. Increasing exposure to the modern economy through English education, and the mismatch in educational choices and future occupational outcomes between boys and girls in the same *jati* that we have documented, suggest that the forces of modernization could ultimately lead to the disintegration of a system that has remained firmly in place for thousands of years.

REFERENCES

Acemoglu, Daron; Johnson, Simon and Robinson, James A. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, 2001, *91*(5), pp. 1369–1401.

Banerjee, Abhijit and Iyer, Lakshmi. "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." *American Economic Review*, 2005, *95*(4), pp. 1190–213.

Burnett-Hurst, Alexander R. *Labour and housing in Bombay: A study in the economic conditions of the wage-earning classes of Bombay*. London: P. S. King and Son, 1925.

Chandavarkar, Rajnarayan. *The origins of industrial capitalism in India: Business strategies and the working classes in Bombay, 1900–1940*. Cambridge: Cambridge University Press, 1994.

Cholia, R. P. *Dock labourers in Bombay*. Bombay: Longmans, Green and Co. Ltd., 1941.

Dakin, Julian; Tiffen, Brian and Widdowson, Henry G. *Language in education: The problem in Commonwealth Africa and the Indo-Pakistan sub-continent*. Oxford: Oxford University Press, 1968.

Dandekar, Hemalata C. *Men to Bombay, women at home: Urban influence on Sugao Village, Decan Maharashtra, India, 1942–1982*. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan Press, 1986.

D'Monte, Darryl. *Ripping the fabric: The decline of Mumbai and its mills*. New Delhi: Oxford University Press, 2002.

Gokhale, R. G. *The Bombay cotton mill worker*. Bombay: Millowners' Association, 1957.

Gore, M. S. *Immigrants and neighborhoods: Two aspects of life in a metropolitan city*. Bombay: Tata Institute of Social Sciences, 1970.

Morris, Morris David. *The emergence of an industrial labor force in India: A study of the Bombay cotton mills, 1854–1947*. Berkeley: University of California Press, 1965.

Munshi, Kaivan. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." *Quarterly Journal of Economics*, 2003, *118*(2), pp. 549–99.

Munshi, Kaivan and Rosenzweig, Mark. "Traditional Institutions Meet the Modern World: Caste, Gender and Schooling Choice in a Globalizing Economy." Bureau for Research and Economic Analysis of Development, Working Paper: No. 038, 2003.

Patel, Kunj M. *Rural labour in industrial Bombay*. Bombay: Popular Prakashan, 1963.

Rees, Albert. "Information Networks in Labor Markets." *American Economic Review*, 1966 (*Papers and Proceedings*), *56*(2), pp. 559–66.