Caste versus Class: Social Mobility in India, 1860-2012

Gregory Clark, University of California, Davis
Zach Landes, UCLA, Anderson School[1]
September, 2012

Using surname distributions, we compare long run social mobility rates for elite and underclass groups in India 1860-2012, with those of other societies such as Sweden and the USA. It is not clear whether recent social mobility rates in India should be higher or lower than in the West. The caste system notoriously embedded privilege in elite castes. But since Independence a quota of places in higher education, and in government jobs, have been reserved for the former lower castes. These quotas are now as great as 50% of such positions. Social mobility rates in India, however, prove to be extremely low, and not any higher now than under the Raj. Despite extensive social engineering India seems to be an unusually immobile society. We hypothesize that this immobility stems from continued strong marital endogamy in India.

Introduction

India is an interesting society in which to study rates of social mobility. On the one hand it entered the modern era with the legacy of the Hindu caste system, which found echoes also in Muslim society, which limited intermarriage, and even social intercourse, between those of different castes. This system of exclusion was so powerful that different castes and sub-castes, even within small geographic areas, have distinct DNA profiles.[2] There is the underfunded and poorly functioning primary and secondary public education system, which those of means have largely abandoned in favor of private schooling. Further many of the poor are located in rural areas for which educational provision is particularly poor, and private

---

[1] With thanks to Lincoln Atkinson for his great help in digitizing the 2.2 million names of the Kolkata voters roll of 2010.

[2] Reich et al., 2009.

Figure 1: College Graduation Rates by Social Group, India, 2000
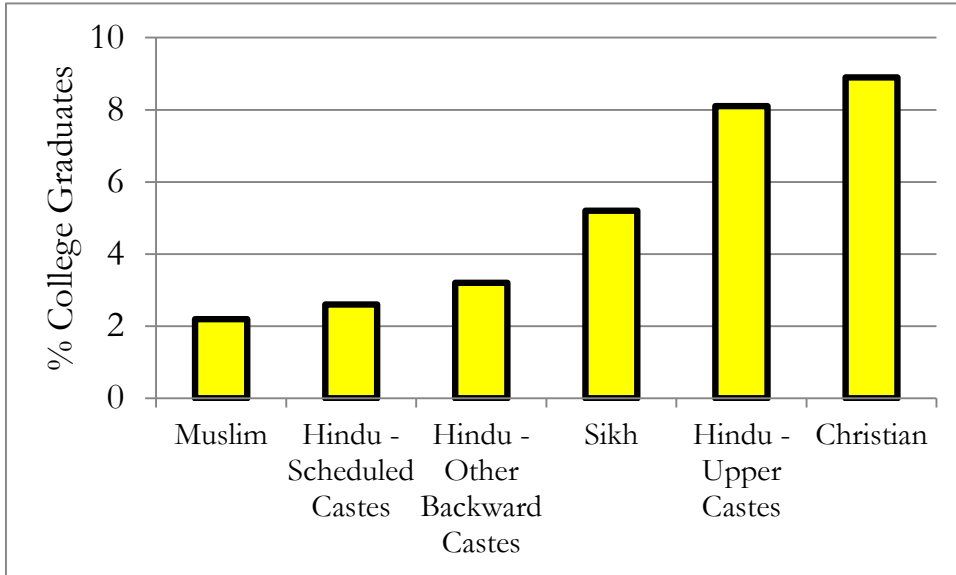
alternatives limited. Thus when we look at College graduation rates by social group, we still see great differences, as portrayed in figure 1, which shows the percent of 23 year olds who had graduated from colleges in India in 2000 by caste and religion. Caste affiliations determined centuries ago still strongly predict current outcomes.

On the other hand, since Independence there has been an extensive system of reservation of positions in universities and government employment, which sets aside up to 50% of positions for traditionally disadvantaged groups. Table 1, for example, shows the candidates admitted to the All India Institute of Medical Sciences for the MBBS degree in 2012, as well as their rank on the entrance exam. Of 72 admissions, only 50% are in the unreserved category. The lowest ranked admit in the unreserved category is 36, compared to 2,007 for the reserved. Suppose the caste system trapped many potentially talented people at the lower levels of the society in the pre-modern era. Then the modern reservation policy could lead to a period of rapid social mobility.

Table 1: Admissions to AIIMS, Delhi, 2012

Category	Number	Rank on Admission Test
Un-Reserved	36	1-36
Scheduled Caste	11	288-1164
Scheduled Tribe	5	177-2007
Orthopedic Physically Handicapped	1	1201
Other Backward Classes (Non Creamy Layer)	19	41-116
All	72	1-2007

Source: Posted Admission List, AIMMS.

Thus while the the situation in figure 1 may speak of continuing dramatic social inequalities, it is not clear whether we should expect high or low rates of social mobility currently.

Formal studies of social mobility in India are, however, modest in number.[3] Thus two recent international surveys of social mobility, one for earnings and the other for education, do not feature India (Corak, 2012, Hertz et al, 2011). However a recent study suggests the Indian intergeneration income elasticity is 0.58 (Hnatkovska et al., 2012).[4] This would indeed classify India on an international scale as one of the world's more immobile societies, as is shown in figure 2. However, since the estimated intergenerational income elasticity for the UK is 0.5, and the US 0.47, this also implies that social mobility rates in India are not too much lower than

---

[3] The large numbers of people still engaged in agriculture make occupational status classifications difficult. Studies of mobility based on occupational classification are thus difficult to interpret, and to compare with those from more developed economies. See for example, Nijhawan, 1969, Kumar et al., 2004, Hnatkovska et al., 2012.

[4] This value corrects for measurement error in income through IV techniques.

Figure 2: Intergenerational Earnings Elasticities and Inequality
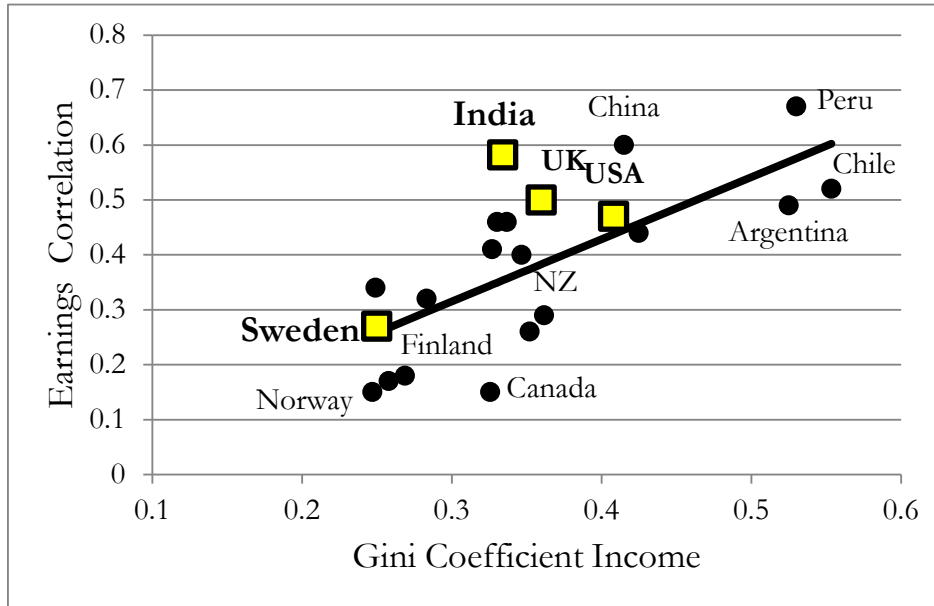
Source: Corak, 2012, Figure 2 (coefficient for Canada, personal communication from Miles Corak). Income elasticity for India from Hnatkovska et al., 2012, table S12. Gini for India from the World Bank.

in the UK or USA (Corak, 2012). Since $(.58)^2 = 0.34$ measures the share of income variance in the next generation explained through inheritance from parents it also implies that even in India the majority if people's position in the income ranks is not derived from inheritance.

This paper uses surname distributions to measure intergenerational social mobility in Bengal from 1860 to 2010, and to ask two questions. The first is, what are long run social mobility rates are in India compared to modern western economies? The second is what are modern social mobility rates are compared to those of India under the Raj? We shall see that measured this way social mobility rates are dramatically lower than suggested by the Hnatkovska et al., 2012 study. The true b seems to be close to 0.9. There is also little sign for many groups of any increase in mobility since the times of the British Raj. However, these long run social mobility rates turn out to be only again modestly higher than the equivalent rates for the UK, USA, and Sweden.

Using Surnames to Measure Social Mobility

The measures we have of status at any time for various surname groups in Bengal is their share in an elite occupation or educational status compared to their population share. We thus measure over time the relative frequency of elite names like Banarji among attorneys, doctors, university attendees, and also of lower class surnames like Shaw.

We define the *relative representation* of each surname or surname type, z, in an elite group as

$$relative\ representation\ of\ z = \frac{Share\ of\ z\ in\ elite\ group}{Share\ of\ z\ in\ general\ population}$$

With social mobility any surname which in an initial period has a relative representation differing from 1 should tend towards 1, and the rate at which it tends to 1 is determined by the rate of social mobility.

The measure we will derive of social mobility is the b in the equation

$$y_{t+1} = by_t + e_t$$

where y is some measure of socio-economic status such as occupation or education.

To extract implied bs from information on the distribution of surnames among elites and underclasses over time we proceed as follows. Assume that social status, y, follows a normal distribution, with mean 0 and variance σ^2. Suppose that a surname, z, has a relative representation greater than 1 among elite groups. The situation looks as in figure 3, which shows the general distribution of status (assumed normally distributed) as well as the distribution for an elite group.

The overrepresentation of the surname in this elite could be produced by a range of values for the mean status, $\bar{y}_{z0}$, and the variance of status, σ^2_{z0}, for this surname. But for any assumption about $(\bar{y}_{z0}, \sigma^2_{z0})$ there will be an implied path of relative representation of the surname over generations for each possible *b*. This is because

$$\bar{y}_{zt} = \bar{y}_{z0}b^t$$

Also since $var(y_{zt}) = b^2 var(y_{zt-1}) + (1 - b^2)\sigma^2,$

Figure 3: Initial Position of an Elite

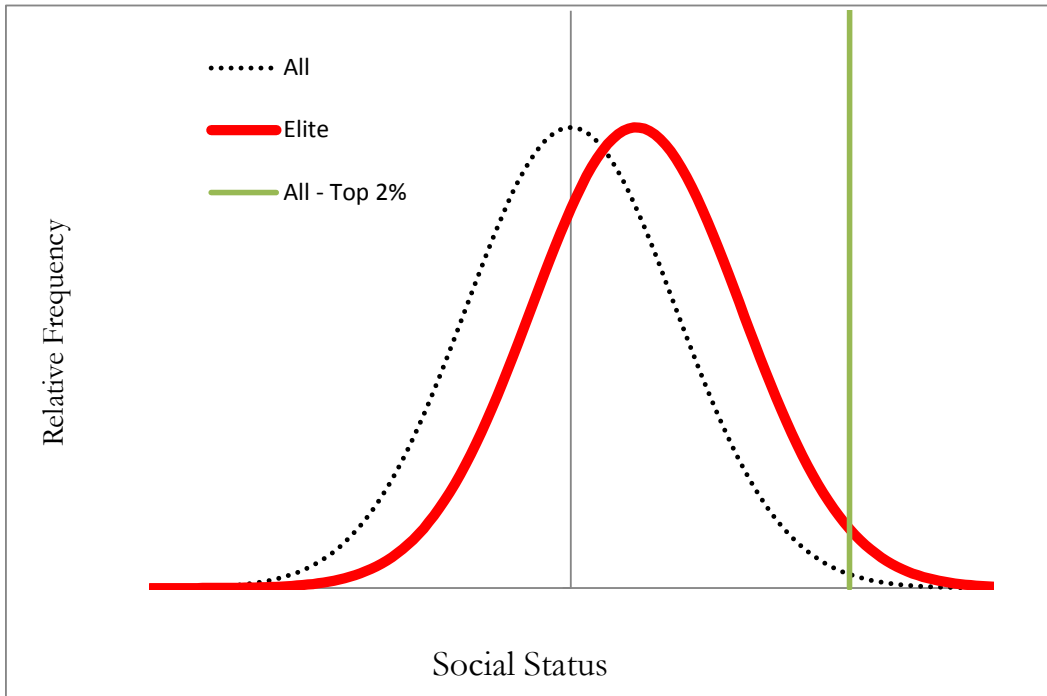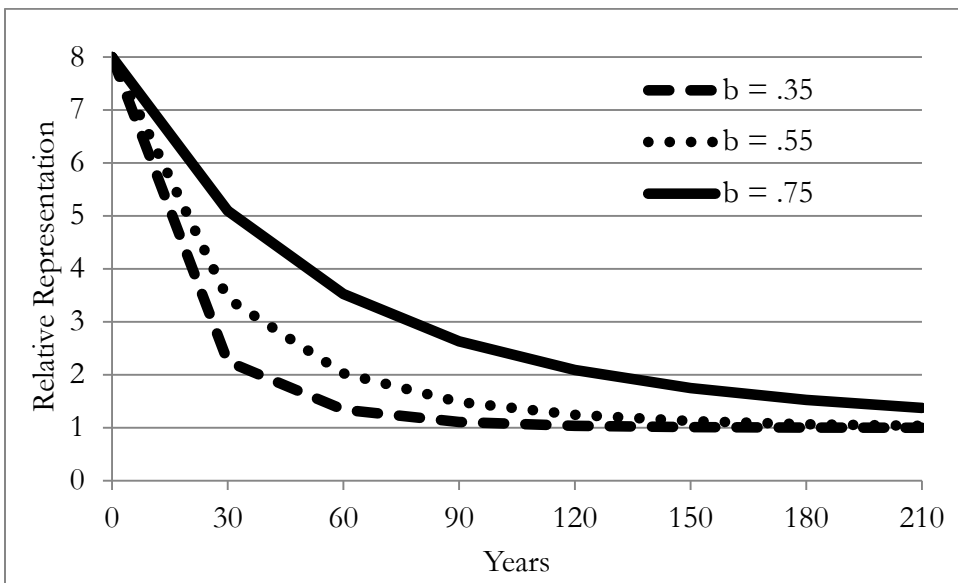**Figure 4: Relative Representation by Generation with Different bs**

$$var(y_{zt}) = b^{2t}\sigma_{z0}^2 + (1 - b^{2t})\sigma^2$$

With each generation, depending on *b,* the mean status of the elite surname will regress towards the population mean, and its variance increase to the population variance (assuming that $\sigma_{z0}^2 < \sigma^2$). Its relative representation in the elite will decline in a particular pattern.

Thus even though we cannot initially fix $\bar{y}_{z0}$ and σ_{z0}^2 for the elite surname just by observing its overrepresentation among an elite in the first period, we can fix these by choosing them along with *b* to best fit the relative representation of the elite surname z in the social elite in each subsequent generation. In the case of India

where elite surnames were established mainly before 1800, we can safely assume that the variance of status among the elite is by the modern period as great as that for the general population (it turns out to matter little to the estimated size of b what specific initial variance is assumed). We shall see below that we can confirm this assumption.

Figure 4 shows what we would expect the relative representation of a surname, which had a relative representation of 8 times its share in the population in the first period, to have in each subsequent 30 year interval with different assumptions about b. Figure 5 shows how the process works in practice in the case of the last two generations in the USA. This shows the relative representation of six groups of surnames among US doctors. The three elite groups are the surnames of the Jewish population, the descendants of the rich of the 1920s, and the descendants of those attending Ivy League universities before 1850. The two underclass groups are the surnames of the Black population, and of those of French Canadian origin. The surname Olson is included as a representation of a group always close to the social average.

This US pattern where elites systematically regress to the mean is echoed in England and Sweden. The rate of social mobility is slow, but there is a consistent regression to the mean by advantaged and disadvantaged surname groups, as in figure 5.

Figure 5: Convergence to the Mean among US Surname Types, 1920-2009
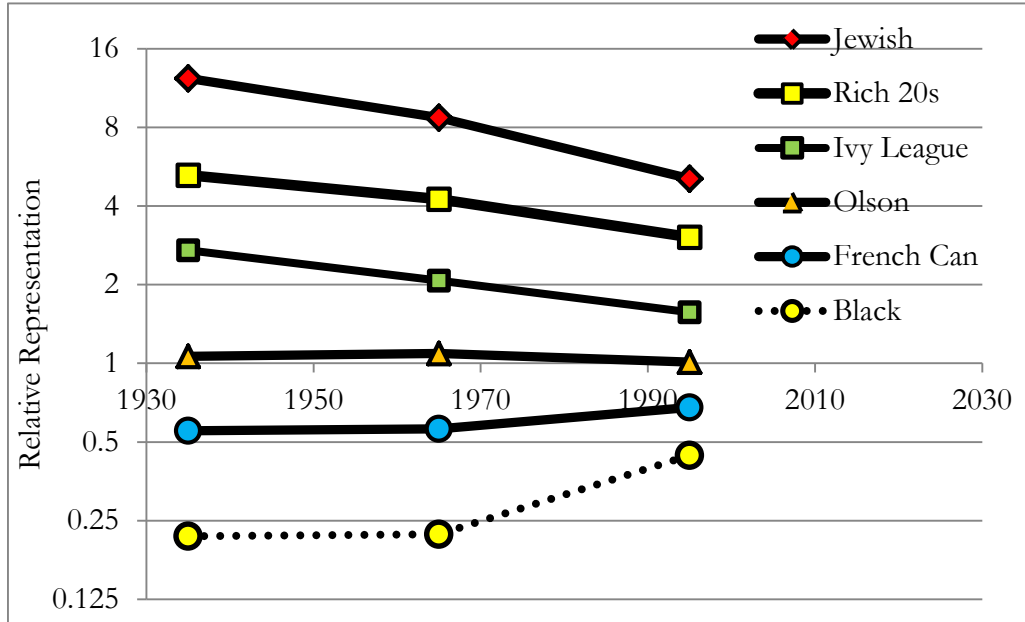
Social Mobility in Bengal, 1757-2010

For the upper classes in Bengal family surnames date from the eighteenth century or earlier. Thus petitioners to the East India Company courts in Bengal in the 1770s typically have surnames, and these are the same surnames still common in Bengal: Banarji, Basu, Chattarji, Datta, Ghosh, Haldar, Khan, Mandal, Mitra, Sen.[5] If there had been substantial social mobility in Bengal, even a b as high as 0.6, then over the last 200-250 years, 7-8 generations, common surnames would all have regressed towards having an average representation at the top and the bottom of society. However, as figure 6 illustrates, common surnames vary enormously in their relative representation among elites in modern Bengal such as doctors or attorneys.

The Hindu community in India was traditionally divided into four castes in descending order of status, Brahmins (priests), Kshatriya (rulers, administrators, soldiers), Vaishya (farmers, bankers, traders) and Shudra (laborers and servants). Even within each castes there were sub-castes of different levels of prestige. The highest status group within the Brahmins of Bengal were the Kulin Brahmins. There are a set of seven surnames that are associated with this group: Mukhopadhyaya (Mukherjee), Bandopadhyaya (Banerjee), Chakraborty (Chakravarty), Chattopadhyaya (Chatterjee), Bhattacharya (Bhattacharjee), Gangopadhyaya (Ganguly), and Goswami (Gosain).[6] This sub-caste of Brahmins supposedly migrated to Bengal from north India in the 10th or 11th centuries AD. If they maintained this status by descent into the modern era then this implies a society of astonishing social rigidity. The surnames of the Kulin Brahmins, however, are generally the most over-represented of all surnames among modern elites in Bengal. They are more than four times as common among doctors first registering in 1980-2009 than their share in the population.

Similarly other surnames associated with the high status Brahmin and Kayastha castes in Bengal are both still overrepresented among doctors and judges in figure 6, even though not as prominent as the Kulin Brahmin surnames. In contrast the surnames of the Muslim population are dramatically underrepresented in both cases, are as some surnames of poor Hindu groups, such as Shaw, who have not benefited from Scheduled Caste status.

---

[5] Government of Bengal, Political Department, 1930.
[6] The association of these surnames with the Kulin Brahmin sub-caste can be confirmed by looking at the surnames of those listing themselves as Kulin Brahmin on matrimonial web sites. All these surnames are found also, however, under other sub-castes of Brahmins.

Figure 6: Relative Representation of Surnames among Doctors and Judges
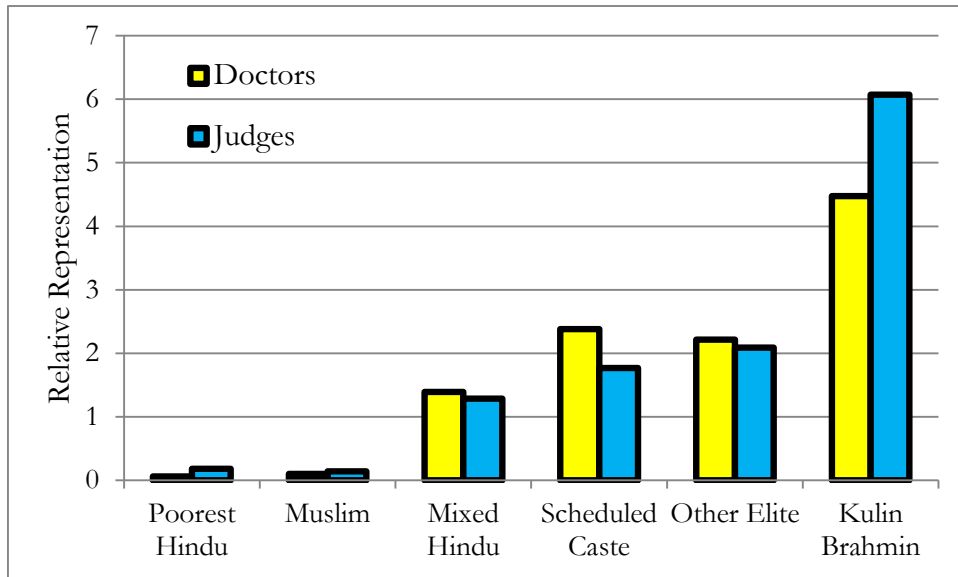
Sources: Surname frequency among doctors in West Bengal from Indian Medical Register. Surname frequency among the population 20-29 in West Bengal estimated from the Kolkata Voter Roll, 2010. Surname frequency among West Bengal judges, 2011, from the High Court Roll of High Court and District Judges.

The extraordinary small share of Muslim surnames among elites such as doctors and attorneys in West Bengal, and their large share of the population, means that Hindu surnames all tend to be overrepresented among doctors. In considering social mobility rates we shall see that they have to be low overall in West Bengal 1947-2012 because there is essentially no upwards social mobility among a large sector of society, the Muslim population.

To measure social mobility over generations in Bengal we look at the relative representation of surname types among doctors in Bengal and West Bengal, 1860-2017. Doctors are just one of a number of high status occupations in West Bengal, but figure 6 suggests that what is true for doctors will be true for other occupations such as attorneys and engineers.

The information for the years 1910-2009 comes from the Indian Medical Register, which includes doctors registering in Bengal from 1915 onwards. Before

1910 we estimate the surname frequencies among doctors from a list of register doctors in the Province of Bengal, 1903, and from lists of doctors registered in Bihar and Orissa, and in Burma (but trained in Bengal) in 1930.   Muslim and Hindu first names are also quite distinct, so we can easily track the fraction of Muslim doctors in Bengal 1860-2009.

Social Mobility of the Muslim Population

For the Muslim population we have census reports on the size of the Muslim population by decade. Figure 7 shows by 30 year generations, starting in 1860-89, the relative representation of the Muslim population among doctors in Bengal, 1860-1947, and West Bengal 1950-2009. This is just the fraction of Muslim doctors in each decade divided by the estimated fraction of the population which was Muslim in the first year of that decade from the census.

The striking feature is the very low representation of this group among doctors in all periods. Under British rule the group was becoming represented more in proportion to its estimated population share over the generations. But, as the figure shows, the implied persistence of status even then was high, with a b of 0.76. However, following Independence the Muslim community in West Bengal has, since the 1970s seen a decline in representation among doctors until (possibly) very recently, with no implied regression to the mean. Indeed going from the generation entering practice 1950-79 to 1980-2009 the implied b is 1.2. This implies the Muslim community has been diverging further from the mean, even though it was already a community concentrated in lower status occupations.

The system introduced in Bengal after Partition that reserves 22-28% of places in all higher education institutions and government employments for Scheduled Castes and Scheduled Tribes, explicitly excludes anyone of the Muslim or Christian religion from the reservation. Only Hindus, Sikhs, and Buddhists qualify for a Caste or Tribe certificate. Bengal has not yet introduced in education any reservation for "Other Backward Classes" which would include Muslims, though this is under discussion.[7] Thus Muslims would be disadvantaged in admission to medical practice compared to the Hindu, Sikh and Buddhist population from 1947 onward. They can
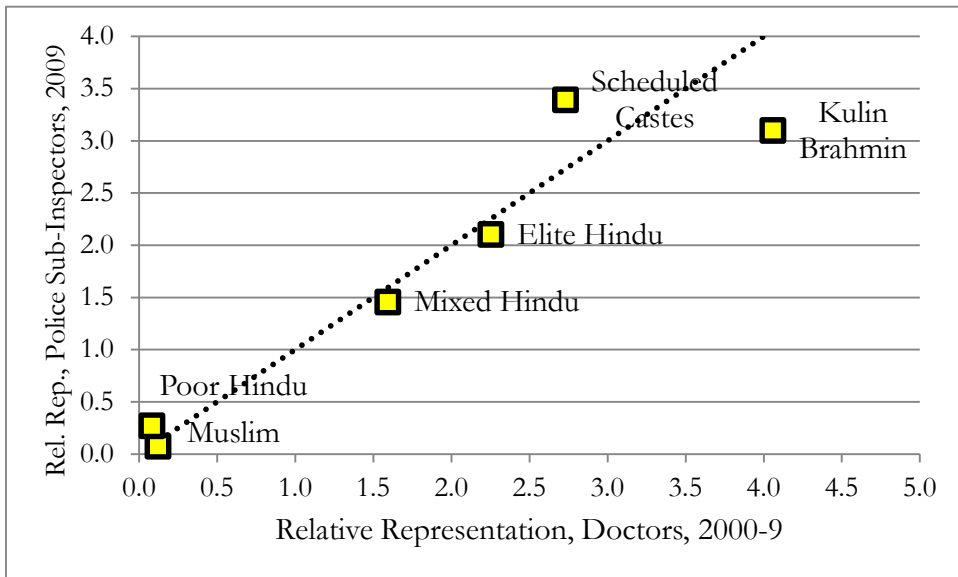
---

[7] In 2012 a Law was passed reserving 17% of government jobs for these groups.

Figure 7: Relative Representation among Doctors, Muslims, Bengal, 1860-2009
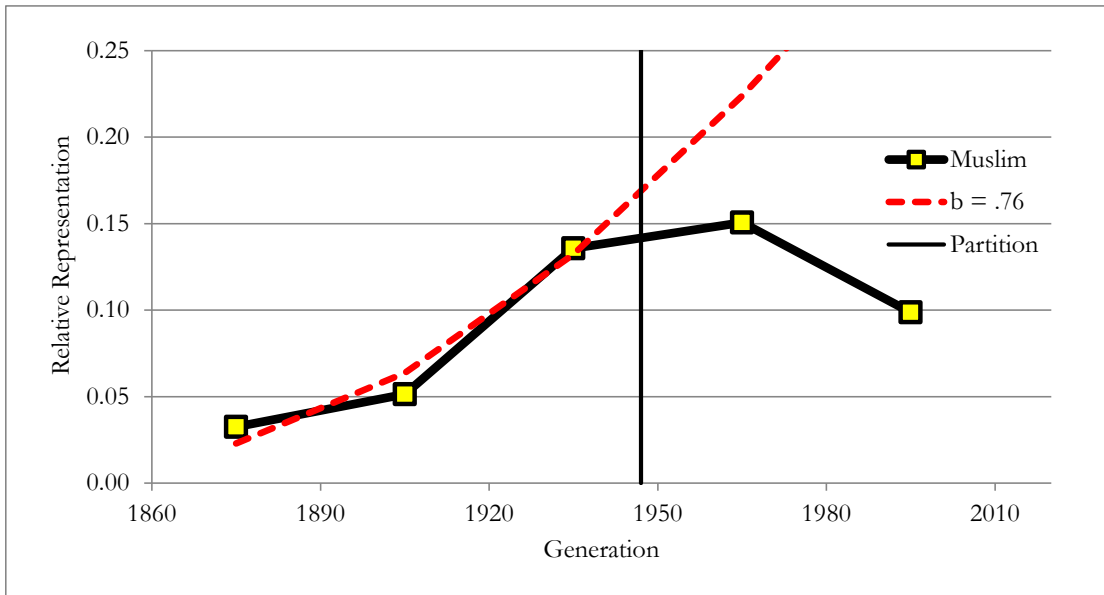
Figure 8:   Relative Representation of Surnames, Doctors versus Police Sergeants

compete on equal terms for the 72% of non-reservation positions. But the existence of the reservation will hit particularly hard such an excluded group whose members would be concentrated on the lower rungs of the admissions if all 100% were open. This may partly explain the surprising negative social mobility implied by these statistics for the West Bengal Muslim community.

Since Muslim representation among doctors, a high status group, is limited it might be thought that the relative representation among doctors does not capture well overall movements in the status of the Muslim minority. However, even in much lower status occupations Muslims seem to be similarly underrepresented. Figure 8, for example, shows the relative representation of Muslims among those admitted as Sergeants and Sub-Inspectors in the Kolkata Police Force, 2009, compared to their relative representation among doctors, 2000-9. These lower level police posts are still coveted positions, but with an education requirement only of High School Graduation. As can be seen Muslims are equivalently underrepresented in promoted positions in the police.

Social Mobility of the Brahmin Population

The seven Brahmin surnames listed above have always constituted a substantial share of doctors in Bengal. Figure 9 shows the percentage of doctors licensed in Bengal and then West Bengal in each decade from 1860-9 to 2000-9 with one of these seven surnames. Until recent decades these 7 surnames account for 15-25% of all doctors in Bengal. Also shown from the AMA registry in the USA are the share of doctors in the US who graduated from West Bengal medical schools in each decade. These proportions are very similar to those found for doctors domiciled in West Bengal.

Figure 9 also shows the estimated population share of these surnames. For 1950 and later this is estimated from the voting roll of Kolkata which lists 2.2 million names, as well as ages. This assumes that the population of Kolkata was representative of the population of West Bengal as a whole. In practice Kolkata likely contains a higher proportion of these Brahmin surnames than in West Bengal as a whole. The city had a more educated population in the 2001 census than West Bengal as a whole, and also a lower proportion was Muslim (21.4% compared to 25.2%). Another feature that appears with the Kulin Brahmin Surnames is that they

**Figure 9: Share of Doctors with Kulin Brahmin Surnames, 1860-2017, Bengal**
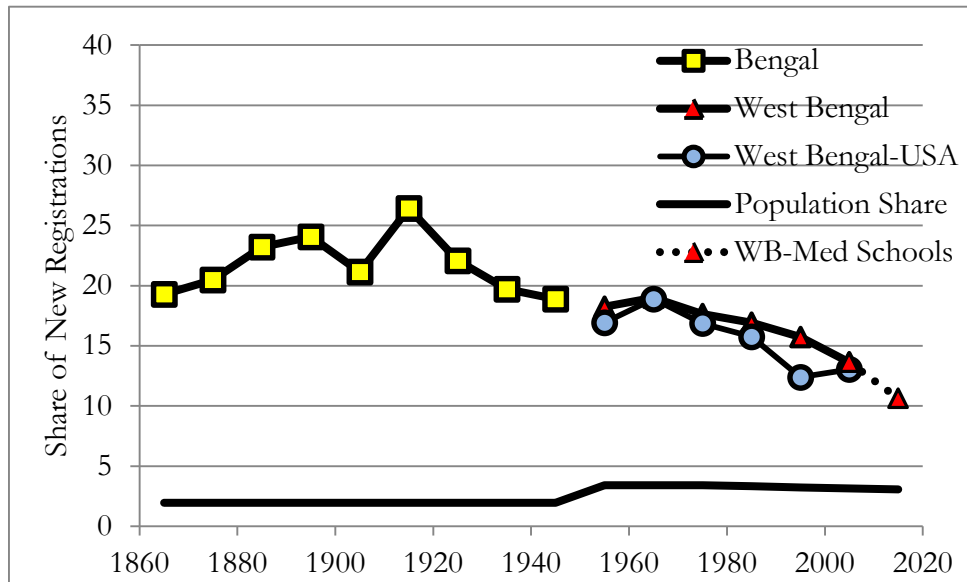
Notes: This proportion was calculated 1910-2009 from a list of 57,407 doctors registered in Bengal and West Bengal between 1915 and 2009.[8] For 1860-1909 it was calculated from four sources: a schedule listing 1,507 doctors in Bengal licensed 1903 or earlier, the 1930 registers just for the component state of Orissa and Bihar, and for adjacent Burma (taking only doctors trained in Bengal), and doctors registered in Bengal 1915 who graduated medical school 1900-1909. For doctors qualifying 2010-2017 the posted admit lists of the Calcutta Medical College, Kar Medical College, KPC Medical College, and Bankura Medical College were employed.

become a steadily larger share of the voters at higher ages, as shown in figure 10. In contrast the poor Muslim community declines sharply as a share of voters at older ages. This presumably mainly reflects differences in average life span between rich and poor, since the differences accelerate with age. But in part it relects differential rates of population growth. In figure 6 the population share of Kulin Brahman surnames is thus estimated for doctors graduating 2000-9 as the share of males with these surnames aged 20-29 in the Kolkata voter roll, multiplied by .95 to account for the overrepresentation of these surnames in Kolkata.[9] For 1990-9 the share is taken

---

[8] This list was obtained from the Indian Medical Register. The earliest registration year listed is 1915, but there seems to have been a listing in that year of some large section of doctors in Bengal, since medical school graduation dates from as early as the 1880s are listed in this cohort.

[9] The share of males was taken since for lower class Hindus, and for Muslims, many women are denoted by non-inherited surnames such as Khatoon, Begum, or Devi, or do not have surnames.

**Figure 10: Surname Shares by Age, Kolkata Voters Rolls, 2010**
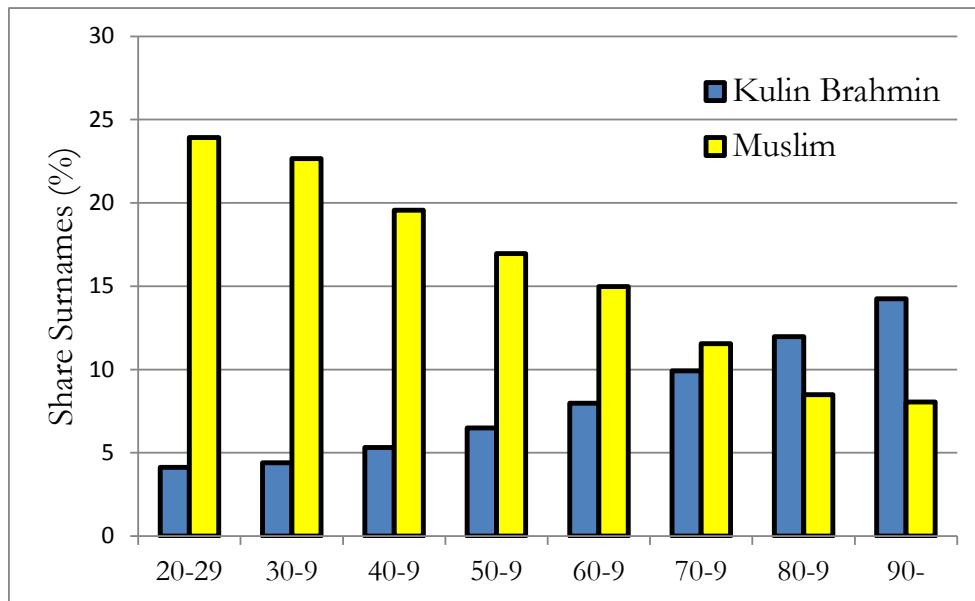
from the Brahmin share among male 30-9 year olds in 2010, similarly adjusted. For 1970-9, 1980-9, and 2010-9 the shares of 1990 and 2000 are projected forward and back based on the % gain 1990-9 to 2000-9. This sets the 1970-9 share at 3.4%, and the 2010-9 share at only 3.15%.

Before 1950-1979 the Kulin Brahmin population share of the Hindu population is assumed constant, based on the fact that the poor Muslim community was not growing as a share of the population 1951-1971.[10] Based on the decline of the Muslim population share from 54.29% of Bengal in 1941 to 19.85% of West Bengal in 1951, the Kulin Brahmin share is estimated at 1.94% in 1940-7. Since the Muslim share again seems to have been constant in Bengal before 1941 (it was 54.43% in 1931), we again take the Kulin Brahmin share as a constant 1.94% 1860-1947. However, this assumes that the Kulin Brahmin surnames grew at the same rate as the rest of the Hindu population 1860-1960. Kingsley Davis, for example, shows that in 1931 the Brahmins in India (a broader group than the surname population we look at here) had a lower ratio of children 0-6 to women 14-43 than any other Hindu group. Indeed the ratio for Brahmins was only 88% of that for other Hindu groups on average. This was mainly a consequence of the social taboo on widow remarriage

---

[10] That share was 19.85% in 1951, 20.46% in 1971.

Figure 11: Kulin Brahmin Relative Representation Among Doctors, 1860-2016
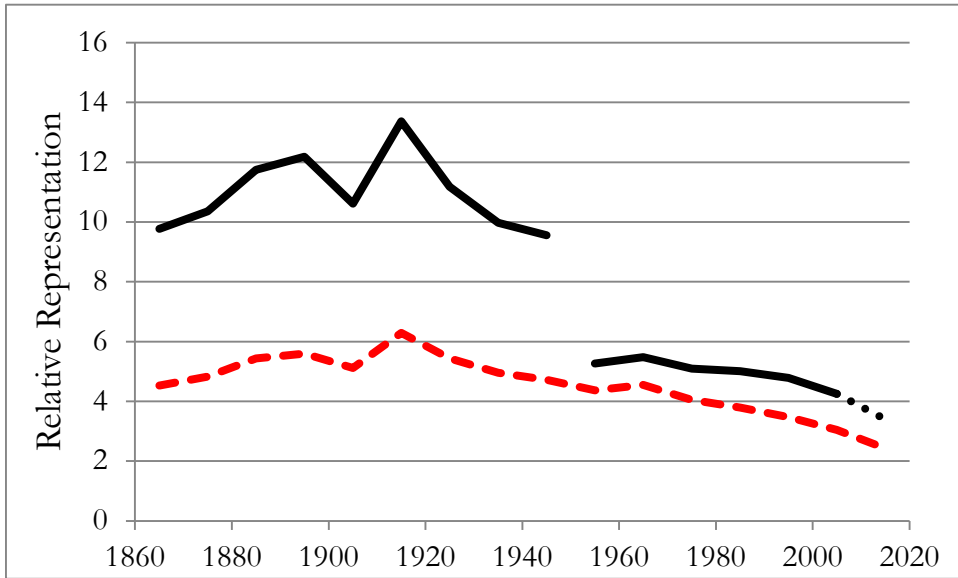
Figure 12: Estimated b for Kulin Brahmins, 1860-2009
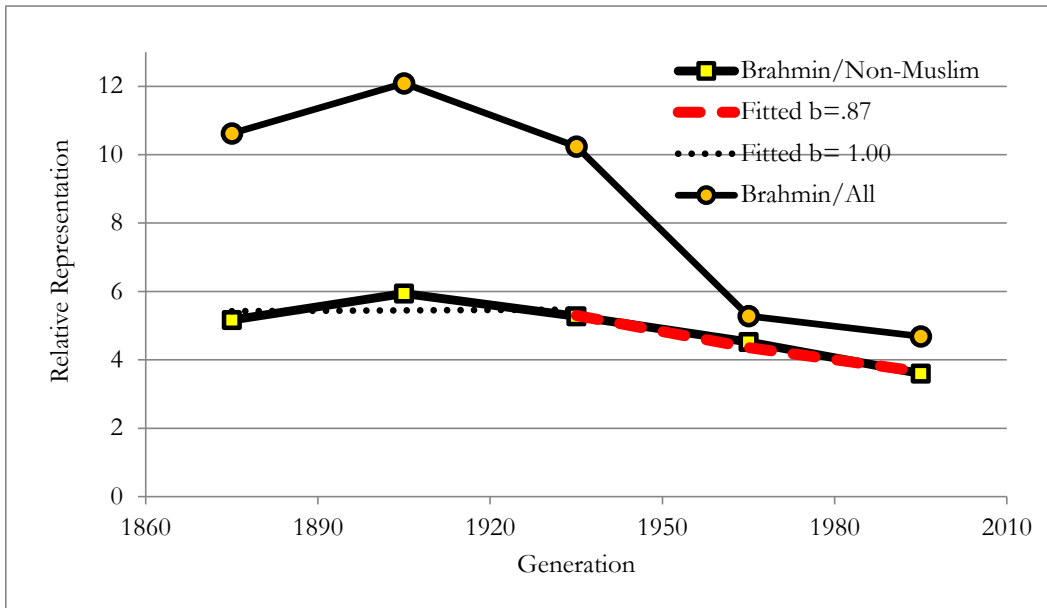
among Brahmins (Davis, 1946, table 3, 248). Since the Brahmins as a group with higher incomes on average presumably had better child survival rates in years subsequent to age 6, we cannot be sure they on net they had lower fertility than the bulk of the population before 1950. That is why we assume Kulin Brahmins were a constant share of the population before 1950.

Figure 11 shows the implied relative representation of these Kulin Brahmin surnames among doctors 1860-1947 in Bengal, and 1950-2016 in West Bengal. Overall there is a decline in relative representation from 7.6 times the average in 1860-89 to 3.5 times the average in 1980-2009, which would imply some modest social mobility. However, this decline is almost all associated with the partition of Bengal in 1947. Since Independence where we actually observe licensed doctors, 1947-2009, these surnames show little sign of declining towards average representation except in the last period which is based on a much smaller sample of medical school admissions. However, within the Colonial period relative representation also shows little sign of decline over time.

The reason for the sharp decline in the overrepresentation of these surnames among doctors around 1947 was the Partition, which removed a large part of the Muslim community. The Partition meant that the share of Kulin Brahmins in the West Bengal population was now significantly higher. But their share among doctors from outside the Muslim population changed little.

To give a better representation of the true rate of social mobility among Kulin Brahmins we thus look at their relative representation in the non-Muslim population among doctors. Figure 11 also shows this. Figure 12 shows these relative representation rates by generation. In the Colonial period the Brahmin surname group shows no signs of downward mobility. The b then is essentially 1. After independence the relative representation of the Brahmin surname group declined, but at a very slow rate. The implied b across the generations 1920-47, 1950-1979, and 1980-2009 is 0.87. Surprisingly the Reservation system introduced after Independence in 1948, which reserved up to 28% of all medical School places in Bengal to Scheduled Castes and Scheduled Tribes, seemingly produced little downward mobility of the Brahmin surname group.

We shall see below that the Reservations did increase sharply the representation of a group of surnames which were heavily represented among Scheduled Castes.

What would have been the rate of downward mobility of the Brahmin surnames had the reservations not been implemented? Assuming that the Brahmin community did lose access to 28% of medical training opportunities, then the relative representation in the final period 1980-2009 would have risen to 5 among non-Muslims, which is very little below its level in 1920-47 under the British.[11] Thus the suggestion here is that absent reservations, there would have been little, if any, downward mobility by the Brahmin community in Bengal all the way from 1860 to 2012.

Other Elite Surnames

There are a number of other surnames – Basu/Bose, Datta/Dutta, Ghosh, Kundu, Mitra and Sen/Sengupta – which we all high status in the nineteenth century. Basu/Bose, Ghosh, and Mitra, for example, are associated the Kulin Kayastha subclans, which were regarded as next in status in Bengal to the Brahmins in pre-modern times. *Kayastha* means "scribe", reflecting the caste's traditional role as administrators. The Kulin pre-fix denoted as with the Brahmins a superior subgroup.

These surnames started in 1860 with a fourfold overrepresentation among doctors in Bengal, as figure 13 shows. The implied b for the Imperial period is slightly different depending on whether it is calculated for the population as a whole, or for the non-Muslim population. But as Table 2 shows it is 0.91-0.93. Again this represents very slow regression to the mean. For the modern period, however, the b for this group of surnames is significantly lower. Now it is 0.73-0.74, depending on the reference group. This group centered somewhat lower on the social scale than the Kulin Brahmin seemingly faced more competition from compression of open places at universities through the Reservation system. But again the implied rates of downward social mobility for this group of surnames remains very slow.

As with the Kulin Brahmin surnames, absent reservations, the rate of relative representation of these surnames in 1980-2009 would have risen to 2.5, very modestly less than the rate of 3.2 in the generation 1920-47 under British rule. This

---

[11] Scheduled caste and tribe candidates who score high enough on the general list are admitted under that list. So the reservation takes away from competition of the higher castes that number of seats at the educational institution.

Figure 13: Relative Representation of Kulin Kayastha Surnames.
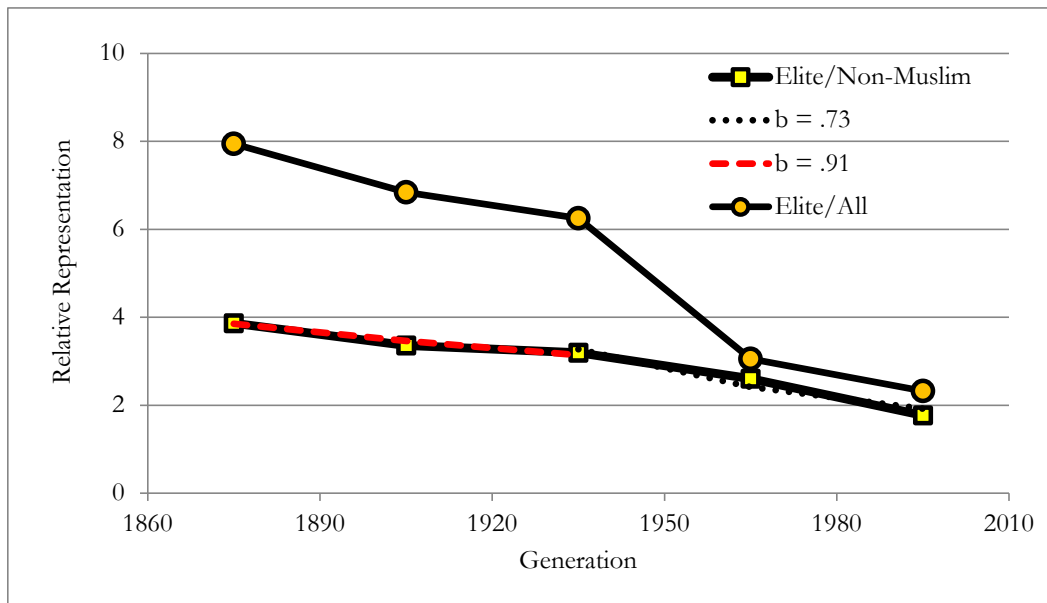
Table 2: b Calculated for Various Groups and Periods

Surname group	Share of Population 2010 (%)	b 1860-1947 All	b 1860-1947 Non-Muslims	b 1950-2009 All	b 1920-2009 Non-Muslims
Muslim	28.8	0.76	-	1.20	-
Kulin Brahmin	3.4	0.99	1.00	0.92	0.87
Other Elite	5.2	0.93	0.91	0.74	0.73
Poorest, pre 1947	6.9	1.02	1.01	0.85	0.82
Scheduled Castes	3.8	1.79	0.62	4.28	-
Mixed Hindu	13.6	1.00	-	2.44	-

implies the underlying rate of social mobility for this group, even absent reservations, implies a b of 0.88. Again the inherent mobility rates are very low.

The Least Represented Surnames, 1860-1947

There are some surnames we can identify which had little or no representation among doctors before 1947. The major one is Shaw or Show, which represents 3.7% of men in Kolkata in 2010 from the voting roll. Others were Rauth/Routh, Paswan, Dhanuk, Balmiki, Mahata/Mahato. Together these surnames constitute 4.5% of the modern population of West Bengal. Despite the Reservation system of post 1947 these surnames show a very small presence among elite groups such as doctors even now, as is shown in Figure 14 and Figure 8 (for Police Sergeants and Sub-Inspectors in Kolkata). As table 2 shows the implied b for this group is, depending on the reference group, 0.82-0.85 under the British and 0.85-0.86 after Independence.

Some of these surnames, such as Dhanuk, belong to groups which while poor, did not qualify as Scheduled Castes because they were not included on the 1931 Census list of Untouchable Castes.[12] While at least some *Shaw/Shows* were included within the Scheduled Castes, many clearly were not. Thus in the list of 499 recruits to the Kolkata Police with the rank of Sergeant or Sub-Inspector, the 4 Shaws were all found only in the General Category. In a sample of medical school admissions 2010-11, 3 of 4 *Shaws* were in the General category.

Scheduled Castes

From lists of those admitted to colleges in West Bengal in recent years, and from lists of successful candidates for Police posts in Kolkata, we can identify some surnames where the majority of holders appear in the Scheduled Caste reservation. These names are Barman/Burman, Biswas, Haldar/Halder, Mandal/Mondal, and Naskar. These names account for 3.8% of the population of Kolkata in 2010.

---

[12] This list was initially promulgated by the British in 1936 based on Untouchable Castes identified in the 1931 Census. The British classification was largely adopted by the Government of India in 1953 in establishing its Reservation Policy. Jadhav, 2008.

Figure 14: Social Mobility of the Poorest Groups, 1860-1947.
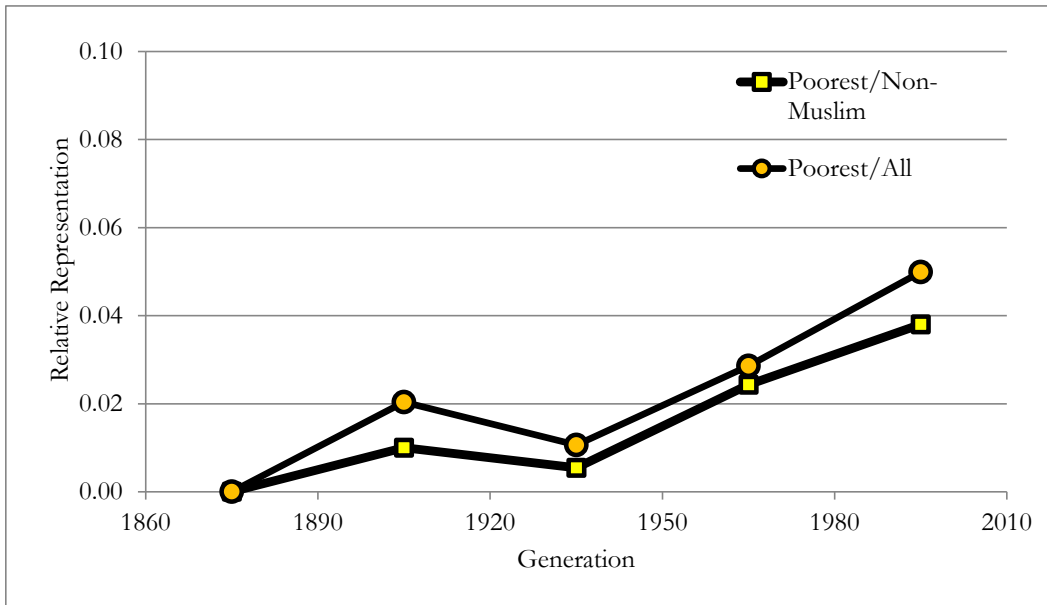
Figure 15: The Curious History of the Scheduled Caste Surname Group
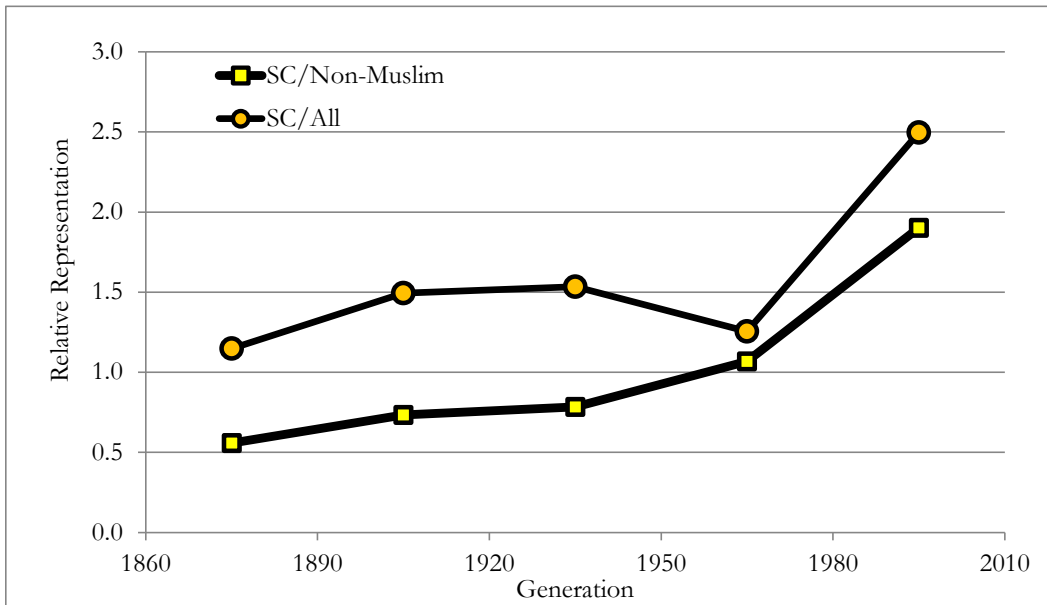
The peculiarity of these surnames is that while all of them figure prominently in the Scheduled cast list, they are all surnames that did figure significantly in the list of Bengal doctors before 1947. Indeed, as figure 15 shows, these surnames were already overrepresented relative to their share in the population among doctors 1860-1889, and were becoming more overrepresented in subsequent generations (b = 1.79). If we look just at the share of these surnames among the Hindu population, then they were at half their expected representation in 1860-1889, but were converging relatively fast towards proportional representation (b = 0.62).

The success of bearers of these surnames under the Reservation Policy has led to these surnames becoming as overrepresented as many surnames of the Hindu higher castes among doctors (and in Police recruiting). See figure 8. For the population as a whole going from 1950-79 to 1980-2009 this implies that there was dramatic divergence of this group away from the mean (b = 4.22). If we look just at the Hindu population, then there is no implied b for the period 1920-1947 to 1980-2009, since the group went from being underrepresented among doctors to being substantially overrepresented.

This recent overrepresentation of these surnames among doctors, even comparing them just to the non-Muslim population, seems likely to be strongly driven by Reservation policy. In a list of recent admits to medical schools in West Bengal which identifies some students by Reservation Category, this group, with 141 admissions, was at double the average representation for the non-Muslim community. However, in the cases where the Reservation status of the surnames was listed, only 30% of these surnames were admitted in the General classification, 70% through the Scheduled Caste category.[13] Had all candidates been admitted through the General classification, then only an expected 58 surnames from this group would have appeared, and this group would have had a relative representation of only 0.84, instead of 2.04.[14]

These results seem to be driven by the arbitrariness of inclusion in the Scheduled Caste and Scheduled Tribe lists. The broad sweep of these lists, starting with the British in the 1931 census, ended up classifying even moderately prosperous

---

[13] Bankura Medical College, class of 2012, and Kar Medical College, classes of 2010 and 2011, had admission lists with Reservation Status.

[14] This is assuming that these surnames would have shown the same relative frequency among additional names admitted under a general category as they showed for the existing names.

groups as "untouchables." These misclassified communities were then best positioned to gain from the Reservation policy.[15]

Mixed Hindu Surnames

The last surname group we consider are Hindu surnames found mainly in the general admission lists for colleges and the police, but also found in significant numbers in the Scheduled Caste lists. These surnames are Das, Dasgupta, Majumdar, Ray/Roy, Saha and Sarkar. Figure 16 shows their pattern of relative representation.

In terms of the population as a whole, these were elite surnames in the era of the Raj, and showed no tendency then to regress to the mean. From 1950-2009 these names have tended again to regress away from the mean in terms of the population as a whole, becoming more elite relative to the general population. But in terms of the non-Muslim community these surnames have an average representation among doctors throughout the period of the Raj and the Independence era. It is thus not possible to estimate a rate of regression to the mean for them since they already are at the mean.

This surname group is both benefitted by, and suffers from, the Reservation Policy. Those without a Scheduled Caste certificate have a lower chance of admission to college, but that is roughly balanced by the improved chances of those within these surname groups with the Scheduled Caste certificate. Looking at the two medical schools with recent information on the Reservation Status of their admits, 58% of this surname group was admitted on the General admission list. Had the Reservation categories been abolished then the relative representation of this group of surnames, compared to other non-Muslim surnames, would have dropped from slightly above 1, to 0.8.

[15] "caste designations appearing on the Scheduled Tribe and Caste lists include many which refer to the broadest and most generalised regional status groupings, and certainly not to homogeneous 'communities' with an inheritance of uniform 'disability' (Bayly, 1999, 277).

Figure 16: Social Mobility of Mixed Hindu Surnames, 1860-1947.
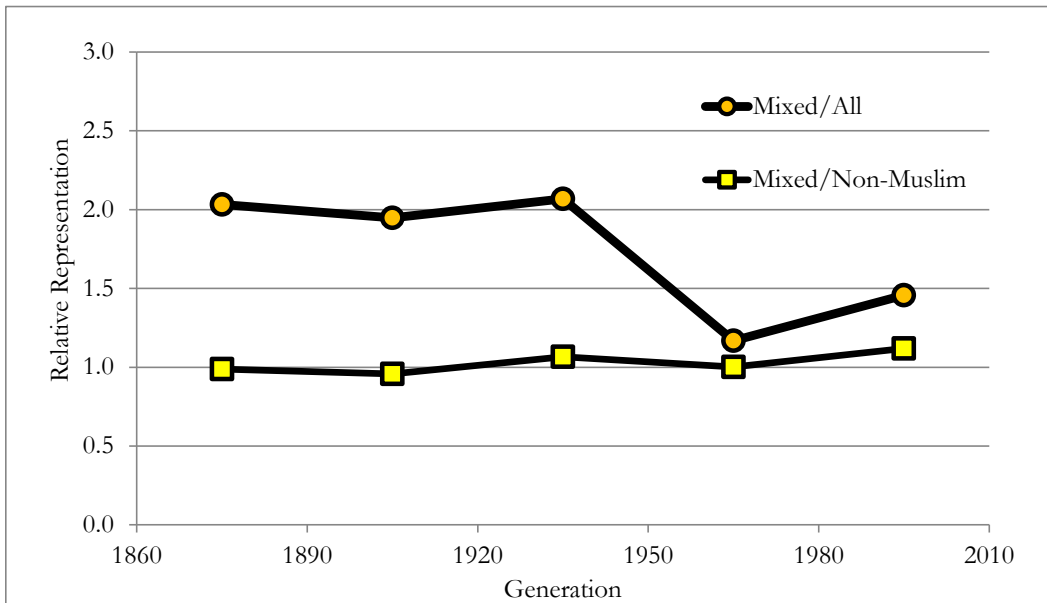
Figure 17: A Summary of Social Mobility by Surname Type, 1860-2009
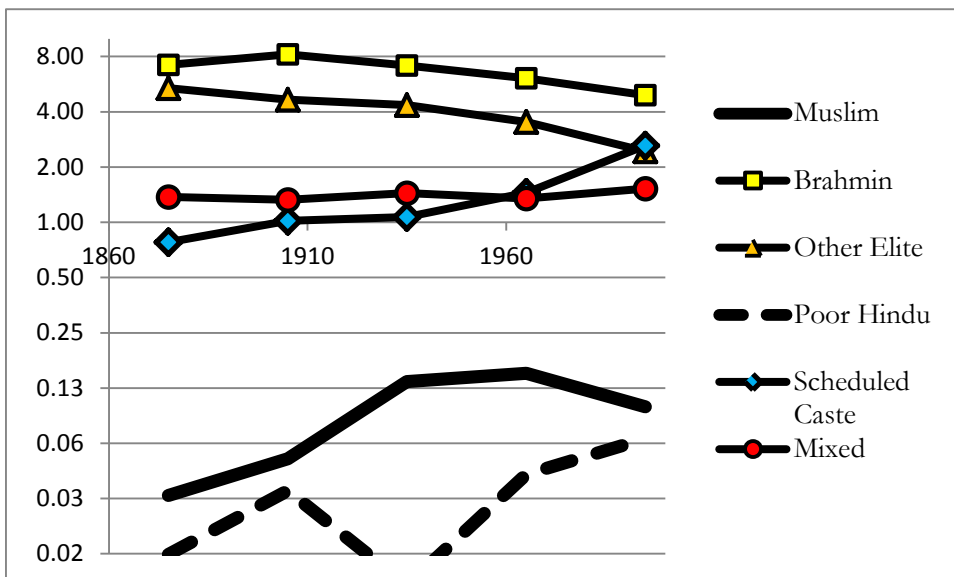
Social Mobility Rates Overall in Bengal

Figure 17 summarizes the relative representation of these six surname groups among doctors over the last five generations. To control for the changing proportions of the Muslim population, the relative representation of the Hindu groups was estimated adjusting the Muslim population to it share in 1980-2009. Unlike in the USA shown in figure 5 we do not get a consistent pattern of movement of groups initially under and over represented towards the mean. Over the last generation the Scheduled Caste Surname group has moved further above the mean, and Muslim representation has declined further below the mean. Surnames with some representation in Scheduled Caste groups, the "Mixed" surnames, have persisted at a higher than average representation. This shows in the estimated values of b for 1920-2009 shown in table 2, which range from 0.74 to 4.28.

However this strange pattern of convergence and divergence seems to be an artifact of the Reservation policy for college admissions in place since the 1950s. Table 3 shows the relative representation of each of our six surname groups in doctors first registering in Bengal in 2000-9. Also shown, from the records of admission to the BMMS for 2010-12 for two medical schools in Bengal is an estimate of the share of admissions to medical school for each group that came through reservations. This is based on only 395 admissions so for one group, the poorest Hindu surnames, there are only four observations. With this information of who entered through reservations, we can estimate what the doctor stock would have been 2000-9 had all admission been by open competition. Column 4 shows the implied relative representation in this case. Figure 18 shows the implied movement of relative representation for each group between 1920-47 and 2000-9 without reservations. Now there is a pattern much more reminiscent of the USA in figure 5 of consistent regression to the mean.

From the counterfactual movement of relative representation excluding the effects of reservations we can calculate an implied persistence coefficient, b, between the generation of doctors 1920-47 and that of 2000-9, 70 years or 2.33 generations later. These estimates are shown in the last column of table 3. While there would then be regression to the mean for the groups that deviate from the mean, the implied rates of social mobility for all of these groups are very low. The average implied value of b per generation is 0.88. Note that this is the implied rate of persistence excluding just some of the effects of reservation policy. It is showing how well the grandchildren of the original generation of 1920-47 in Bengal are now doing in competition for medical school places (absent Reservation). But there will be some effect from the previous round of admissions by reservation. The Brahmin group for example, with an estimated persistence of 0.95, in the previous generation had fewer people in high status occupations as a result of Reservation policy in the

Table 3: Implied b without Reservation Policy, Doctors, Bengal

Surname Group	Relative Representation 2000-9	Share Admitted Through Reservation 2010-12	Relative Representation 2000-9 no reservation	Implied b 1935-2005, no reservation
Kulin Brahmin	4.25	0	5.90	0.95
Other Elite	2.36	3	3.17	0.89
Mixed	1.67	45	1.28	0.84
Scheduled Caste	2.86	70	1.18	-
Poorest Hindu	0.10	25	0.11	0.77
Muslim	0.12	0	0.17	0.96

Figure 18: Social Mobility by Surname Type without Reservations, 1920-2009
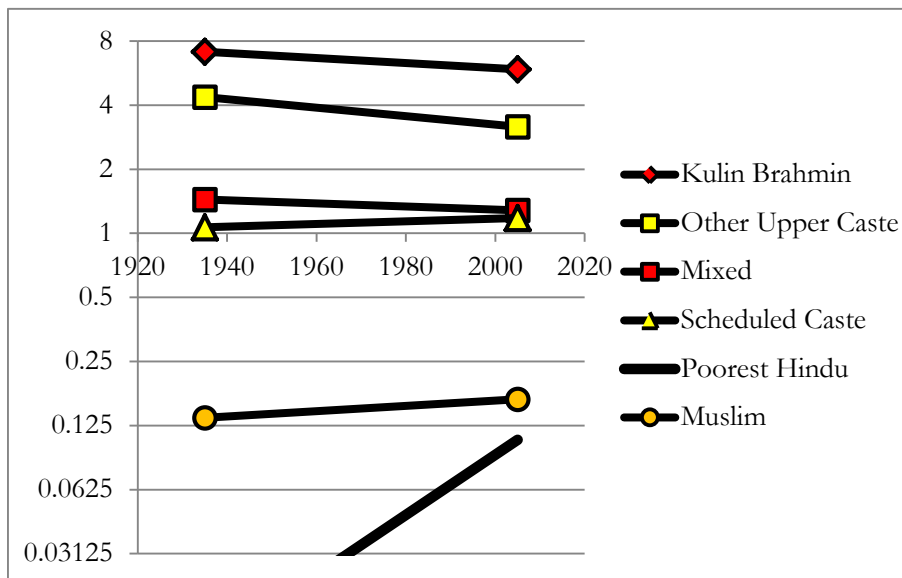
previous generation. Thus their children competing for medical school places in 2000-9 had parents from less advantages backgrounds than would have happened absent Reservations. This implies that the true b for the Brahmin group in Bengal, absent Reservation policy in its entirety, would be 0.95 or even greater.

For the Muslim community the estimated b excluding reservations is 0.96, again very high. In this case, however, reservation policy in the generation previous to the doctors observed in 2000-9 would have reduced the average economic status of parents, and so the true b in this case may be lower than 0.96. Thus while the effects of Reservation Policy between 1950 and 1999 cannot be controlled for, on balance while they may have reduced the b for the initial high status groups, they would have increased it for the low status groups that were excluded by this policy. So absent Reservations, the evidence is that on balance there would have been very strong persistence of social position in modern Bengal.

It is also not clear if Reservations are doing much to change these overall slow social mobility rates in Bengal. As witnessed by surname distributions, the 70% of the population that lies outside the reserved categories has seen little change in its relative social position over the last two generations. Within the groups falling into the Reserved categories, a few seem to have reaped disproportionately the gains, while others seem to have experienced few benefits. Thus despite the intergenerational mobility injected, at least in the short run, by the Reservation system the impression from the surname groups is of an overall long run rate of social mobility that must be significantly less than the 0.59 estimated by Hnatkovska et al., 2012.

However, the objection could be raised that while the bs implied by the surname grouping mobilities is low, could there be considerable social mobility within these surname groupings. Thus while the surname *Banarjee* may be consistently overrepresented among elite groups, could it be that completely different groups of *Banarjee*s are found in the elite occupations each generation?

Even if there was considerable random movement up and down within families within each of these surname groups, however, the surname data shows that this individual mobility data will not predict the social mobility of larger groups of advantaged and disadvantaged castes, religions, or classes within Indian society. The question of the social mobility of classes or of castes would not then be amenable to answer from estimated individual mobility rates.

The second answer is that based on experience in other countries the high persistence of surname groupings in relative status is echoed by the high persistence of individual families within these groupings.

Why is Social Mobility so Low in Bengal?

The social mobility rates for modern West Bengal estimated in table 3 are among the lowest observed in a series of surname studies for England, 1300-2012, USA 1920-2012, Sweden, 1700-2010, China, 1700-2010, Japan, 1870-2010, and Chile, 1920-2010. Table 4 shows the persistence rates estimated in these other cases in recent years, as well as in some earlier periods. The typical rate is 0.7-0.8, still very high, but significantly less than the average of 0.89 observed in West Bengal since Independence. Why are these persistence rates so high in Bengal?

Why are rates of social mobility consistently so low in Bengal? The hypothesis offered here is that this is caused by low rates of intermarriage between different surname groups in Bengal. There has been surprisingly little study of intermarriage rates between different social groups in India in general and in Bengal specifically, despite the importance of the caste system in Indian history and politics. As late as the 1960s caste endogamy still seemed to be the rule for most marriages in Bengal, as seen in a detailed study of a modest sized town in Bengal in the late 1960s (Corwin, 1977). Another study of a high caste group in Hyderabad, Kayasths, looking at marriages 1900-1975, found that rates of marriage within the caste were 98.5%, 1900-25, 97.1% 1926-50, and 94.8% 1951-75 (Leonard and Weller, 1980, tables 1-3). But information on the endogamy rates of marriages in Bengal in the 1970s to 1980s, which produced the most recent crop of doctors, is not readily available.

One source we do have on the likely rate of endogamy is the 2010 Kolkata voter roll, which gives surnames, first names and ages. There are many first names that are highly specific to the Hindu, Muslim, and Christian/Jewish communities. Table 5 shows the most common ten first names for women in each category. If we take, for example, the Kulin Brahmin surnames then women who married into this surname group from the Muslim or Christian communities would almost always have different first names than women born into this group. Also if families with these surnames were identifying as Muslim or Christian, as a result of intermarriage and adoption of at least some elements of the culture of the wives, then the children would again have different first names. However, as table 6 reports, the percentage of women in the Kulin Brahmin surname group having non-Hindu first names is extremely small. Since Muslims are nearly a quarter of the Kolkata population this implies that intermarriage rates between Kulin Brahmin men and women of Muslim origin must be extremely low, in the order of 0.1%. A similar result holds for other high caste Hindu surnames.

More women with Muslim surnames have first names that are Hindu in origin, 0.9%. But given the absence of sign of any intermarriage with high caste Hindu groups, if these reveal marriage alliances it is likely with lower caste Hindus.

Table 4: Estimates of b from Surnames, other Societies

Country	Measure	Period	B
USA	Attorneys	1950-2011	0.67-0.77
USA	Doctors	1950-2011	0.73-0.74
England	Attorneys, Doctors	1950-2012	0.69-1.00
England	Wealth	1950-2012	0.70
England	Education	1950-2012	0.77
England	Education	1300-1500	0.75
Chile	Occupations	1940-2010	0.74
China	Education	1905-2011	0.71
Japan	Education	1940-2012	0.84

<u>Sources</u>: England, Clark and Cummins, 2012**,** China, Hua and Clark, 2012, India, Clark and Landes, 2012, Japan, Clark and Tatsuya, 2012, USA, Clark et al., 2012. Chile communication from Daniel Diaz.

Table 5: Most Common Female First Names by Community

Kulin Brahmin	Other High Caste Hindu	Muslim	Christian/ Jewish
Krishna	Geeta/Gita	Salma	Mary
Soma	Krishna	Yasmin	Elizabeth
Geeta/Gita	Soma	Shabana	Maria
Arati	Arati	Asma	Margaret
Swapna	Meera/Mira	Sultana	Helen/Helena
Meera/Mira	Namita	Anwari	Agnes
Kalpana	Kalpana	Shabnam	Veronica
Ratna	Anjali	Afsana	Rosemary
Sumita	Swapna	Shahnaz	Dorothy
Anjali	Pratima	Farzana	Teresa

Table 6: Female First Name Origins by Surname Group

Female First Names	Kulin Brahmin	Other High Caste Hindu	Muslim	Christian
Hindu	99.6	99.3	0.9	30.2
Muslim	0.1	0.1	98.9	0.4
Christian	0.3	0.6	0.2	57.4
Hindu and Christian	0.0	0.0	0.0	11.9

Source: Kolkata Voter Roll, 2010.

There is sign of potentially much more intermarriage between Christians and high caste Hindus. Christian origin surnames are a very small share of the surname stock in Kolkata, about 0.3%, and are mainly Portuguese in origin, indicating their long history in India. The small share of women with high caste surnames who have Christian surnames is compatible with significant intermarriage, given the small Christian population share. But these female Christian first names may alternately stem from daughters from some high caste Hindus being given Christian surnames at birth, as opposed to intermarriage. The possibility of significant intermarriage between Christians and Hindus is, however, supported by 30% of women with Christian surnames having first names that are Hindu. Also 12% of women with Christian surnames have a mixture of Christian and Hindu first names. But again there is little sign of marriages that cross the Muslim-Christian social divide.

The surname evidence thus suggests almost no intermarriage between the largely poor Muslim community and either Hindus or Christians. Within the Hindu community it is not see easy with the first name evidence to see whether there is still marital endogamy within the surnames that are associated with the high caste groups. This is because there is not such dramatic variation in surname frequencies between high status and low status Hindu groups in first name types. There are only a few female first names that vary dramatically between high caste and low caste Hindus.

One of these is *Munni*, found at the rate of 0.007% among high caste surname women, and at the rate of 0.20% among other Hindu surname groups. If *Munni* was distributed representatively in the rest of the Hindu population, maintaining this incidence disparity would require than less than 4% of elite surname men married women from the general Hindu population. Again there would be a high degree of marital endogamy among elite populations. However there is clear indication that the poorer the Hindu surname the more prevalent is the first name *Munni*. For the poorest Hindu surnames it is found at a rate of 0.9%. Thus the failure of the first name *Munni* to appear among women with elite surnames may not reflect a general marital endogamy among these groups, but only a failure to marry women drawn from low in the social scale.

Another source of evidence of the continuing strength of marital endogamy are web sites advertising for potential wedding partners in Bengal. A survey of 200 women identified as Kulin Brahmin finds that 83% specify that they are seeking a Brahmin husband, 2% specify Brahmin or other high caste, and only 15% state that caste status is no barrier to a potential union.[16] However, among that 15% open to any caste, 8% list this in a form such as "Brahmin - Kulin, Caste no bar". Thus a full 93% of advertisements indicate a preference for a Brahmin spouse.

Why would marital endogamy among surname groupings slow the rate of social mobility? We hypothesize that this stems from the fact that the current status of a person, y_t, on any of the various aspects of social status in generation t – income, wealth, education, occupation - has two components, a systematic one and a random element. Specifically $y_t = \theta x_t + e_t$, where x_t, is some the fundamental social competence or status of families, and e_t is some random component. The random component exists for two reasons. First there is an element of luck in the status attained by individuals given their underlying aptitudes. People happen to choose a successful field to work in, or firm to work for. They just succeed in being admitted to college, as opposed to just failing. But, second, people trade off income and other aspects of status. They choose to be philosophy professors as opposed to finance executives. The systematic component is strongly inherited, it could be by social or genetic inheritance (they would be observationally equivalent), but the random component is not inherited at all.

If people match up in marriage based on the current status of families only, then they look just at how the family ranks on current status y_t. High ranked families will tend to be those with positive luck, and the children of these unions getting on average no such bonus, will tend to regress to the mean. This is the normal process of social mobility. Even if castes and the associated surname groups differ in terms of the average value of the underlying competence, x, as long as matching in

---

16 bengalimatrimony.com

marriage is by attained current status, *y*, the average underlying competence or abilities of the castes and associated surname groups will converge over time.

However if marriage is endogenous to caste or religion, then while there will be social mobility within each caste, there is no mechanism to eliminate the underlying differences in the average level of ability or competence of different castes. At the level of castes and the associated surname groupings there will be little or no social mobility. The differences in socio-economic rankings between these groups diminishes little, or not at all, over time.

Conclusions

Long run social mobility rates in India as measured by the frequency of surname types in high status occupations such as doctors or judges turn out to be even lower than the low rates observed in countries such as England, the USA and Sweden using equivalent methods. The underlying b before the effects of the reservation system is estimated at 0.89 in Bengal, higher even than in medieval England. This is what allows the maintenance even today of the great social disparities illustrated in figure 1, despite 60 years, two generations of extensive Reservations in education for lower castes.

This unusually low rate of social mobility, we argue, is the result of high rates of marital endogamy among social groups in India. We show that in Kolkata there is little or no intermarriage between Muslims and high caste Indians and Christians. There is also sign of significant continued endogamy within caste groups within the Hindu population.

Sources

Kolkata Electoral Roll, 2010.
http://ceowestbengal.nic.in/AcNoMatrix.asp?m_menu_id=40&m_district_id=11

(downloaded and digitized by Lincoln Atkinson).

Kolkata Residential Phone Directory
(http://www.calcutta.bsnl.co.in/newdirectory/dqresidential.php).

Thacker's Bengal Directory. 1870. Calcutta: Thacker, Spink and Company.

Thacker's Indian Directory. 1942. Calcutta: Thacker, Spink and Company. Alphabetical List of Principal Indian Residents.

Indian Medical Register,
http://www.mciindia.org/InformationDesk/IndianMedicalRegister.aspx

Government of Bengal, Political Department. 1930. *Press List of Ancient Documents relating to the Provincial Council of Revenue at Calcutta preserved in the Secretariat Room of the Government of Bengal.* Series II – Intermediate Revenue Authorities, Vol III (part 1) 1773-1775.

Judges Sitting in the High Court and District Courts, West Bengal from
http://calcuttahighcourt.nic.in/

Bibliography

Bayly, Susan. 1999. *Caste, Society and Politics in India from the Eighteenth Century to the Modern Age. The New Cambridge History of India, IV.3.* Cambridge: Cambridge University Press.

Clark, Gregory. 2012. "What is the True Rate of Social Mobility in Sweden? A Surname Analysis, 1700-2012." Working Paper, UC Davis.

Clark, Gregory and Neil Cummins. 2012. "What is the True Rate of Social Mobility? Surnames and Social Mobility in England, 1800-2011." Working Paper, University of California, Davis.

Clark, Gregory, Daniel Marcin, Firas Abu-Sneneh, Wilfred M. Chow, Kuk Mo Jung, Ariel M. Marek, Kevin M. Williams. 2012. "Social Mobility Rates in the USA, 1920-2010: A Surname Analysis." Working Paper, UC Davis.

Corak, Miles. 2006. "Do Poor Children Become Poor Adults? Lessons from a Cross Country Comparison of Generational Earnings Mobility." IZA Discussion Paper No. 1993. Bonn: Institute for the Study of Labor.

Corwin, Lauren A. 1977. "Caste, Class and the Love-Marriage: Social Change in India." *Journal of Marriage and the Family, 39*(4): 823-31.

Davis, Kingsley. 1946. "Human Fertility in India." *American Journal of Sociology, 52(3)*: 243-254.

Hao, Max and Gregory Clark. 2012. "Social Mobility in China, 1700-2012." Working Paper, UC Davis.

Hertz, Thomas, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith and AlinaVerashchagina. 2007. "The inheritance of educational inequality: International comparisons and fifty-year trends." The B.E. Journal of Economic Analysis & Policy 7, Article 10.

Hnatkovska, Viktoria, Amartya Lahiri, and Sourabh B. Paul. 2012. "Breaking the Caste Barrier: Intergenerational Mobility in India." *Journal of Human Resources*, forthcoming.

Jadhav, Praveen. 2008. "Relative Disparity in the Implementation of Reservation Policy in India." In Jagan Karade (ed.), *The Development of Scheduled Castes and Scheduled Tribes in India.* Newcastle, U.K.: Cambridge Scholars Publishing, 1-10.

Khan, Abdul Majed. 1969. *The Transition in Bengal, 1756-1775. A Study of Saiyid Muhammad Reza Khan.* Cambridge: Cambridge University Press.

Kumar, Sanjay, Anthony Heath and Oliver Heath. 2002. "Determinants of Social Mobility in India." *Economics and Political Weekly, 37*(29): 2983-2987.

Leonard, Karen and Susan Weller. 1980. "Declining Subcaste Endogamy in India: The Hyderabad Kayasths, 1900-75." *American Ethnologist,1 (3), 504-17.*

McLane, John R. 1993. *Land and local kingship in eighteenth-century Bengal.* Cambridge: Cambridge University Press.

Nijhawan, N. K. 1969. "Inter-Generational Occupational Mobility." *Economic and Political Weekly, 4, No. 39*: 1553-7.

Reich, David, Kumarasamy Thangaraj, Nick Patterson, Alkes L. Price and Lalji Singh. 2009. "Reconstructing Indian population history**."** *Nature 461*: 489-494.

Sethi, Rajiv and Rahini Somanathan. 2010. "Caste Hierarchies and Social Mobility in India." Working Paper, 2010.

University Grants Commission. 2008. *Higher Education in India.* New Delhi.

Watkins, WS, R Thara, BJ Mowry, Y Zhang, DJ Witherspoon, W Tolpinrud, MJ Bamshad, S Tirupati, R Padmavati, H Smith, D Nancarrow, C Filippich and LB Jorde. 2008. "Genetic variation in South Indian castes: evidence from Y-chromosome, mitochondrial, and autosomal polymorphisms." *BMC Genetics,* **9:** 86.