Social Heterogeneity and Economic Mobility

Nazneen Alam [*]
Sukanta Bhattacharya [†]
Anirban Mukherjee [‡]

Abstract

This paper aims to examine the link between ethnic heterogeneity within a locality and inter-generational occupational mobility of its residents. We see that the more mixed/heterogeneous the locality (in terms of caste), the higher the odds of upward occupational mobility, possibly as a result of positive inter-group contact. Further, we see that in case of more homogeneous localities, if the predominant population belongs to the privileged caste, they are more likely to be concentrated in high skilled, high wage jobs. On the other hand, homogeneous localities with predominantly less privileged caste residents, show a shifting concentration towards lower skilled, lower paid jobs. The observations of this paper may provide another perspective on why the historically and conventionally privileged section of the society might keep flourishing under the existing societal paradigm, while the marginalised section would always need to pay a higher cost of movement to access information and opportunities.

[*]University of Calcutta
[†]University of Calcutta
[‡]University of Calcutta

1 Introduction

The traditional caste system structures the society in a hierarchical formation of caste groups where one's caste identity is tied to their occupation [1]. Therefore, in its purest form, the caste system does not allow any mobility at all. In its contemporary form however, the institution of caste does not stop mobility altogether but restricts mobility. In this backdrop, researchers have looked into the role of state and market in bringing the convergence between the disadvantaged and the advantaged castes (Hnatkovska et al., 2012, 2013; Chin and Prakash, 2011). In this paper, we approach the same issue from a different angle. Here, we examine if residential segregation affects social mobility of disadvantaged castes.

The way residential ethnic composition affects one's economic mobility is not trivial. Existing literature shows that the level of public good in an area depends on its ethnic composition. In their seminal paper on this issue, (Alesina et al., 1999) argue that preference for public goods is aligned with ethnicity and therefore, in an ethnically mixed area, the preference for public goods is diverse. This leads to less public contribution for and less provisioning of public goods in such areas. Similar evidence of negative relationship between ethnic heterogeneity and public goods provisioning was also found in India as well (Banerjee et al., 2005). The mechanism connecting ethnic heterogeneity and private income however, takes a different route. The basic premise of our paper comes from the literature that examines the role of peer-groups on one's educational and occupational achievements. Following this approach, researchers have evaluated the impact of Moving to Opportunity (MTO) program in the United States. Starting in 1994, MTO program provided 4600 randomly chosen low income families living in disadvantaged urban neighbourhoods with an opportunity to move to less distressed neighborhoods (Mogstad and Torsvik, 2021). It was found that children below 13 who moved with their families to less poor neighbourhoods earn 31% more than their control group counterpart (Chetty et al., 2016). In another study, using national data for the United States, Chetty and Hendren (2018) found that For children growing up in low-income families, each year of childhood

[1] The broad classification of the caste system consists of four caste groups – Brahmin, Khatriya, Baishya and Shudra. The main occupation for Brahmins is scholarly work; for Khatriyas, it is warfare; for Baishyas, the main occupation is business and for the Shudras, it is menial jobs (Jaffrelot, 2013). The groups which lie below these four groups in the traditional caste hierarchy are the untouchables who, adopting the terminology coined by B.R. Ambedkar, are now called the *dalits*(Muthukkaruppan, 2014).

exposure to a one standard deviation better county increases adulthood income by 0.5%.

But why does neighborhood matter? In a more recent survey article Chyn and Katz (2021) discusses different mechanisms through which neighbourhood effect can influence economic outcomes. Examples of such mechanisms include peer influences, neighborhood safety and exposure to violence, school quality, the physical environment, and access to employment and criminal opportunities. In his seminal paper, Manski (1993) categorizes the channels through which peer groups affect an individual's behaviour into three groups. These are endogenous effects, contextual (exogenous) effects and correlated effects. [2]

All there of these mechanisms apply to a neighbourhood peer effect. Besides these channels neighbourhood norms may also affect people's behaviour. One major example of such incident is "acting white"' in the United States. Austen-Smith and Fryer (2005) defines "acting white" as socioeconomic signalling which determines one's acceptance in peer groups. Fordham and Ogbu (1986) highlights the effect of "acting white" on black students in America. Black adolescents face rejection by their peer group and a clash in their collective identity or "fictive kinship" if they "act white". Attempts to distance oneself from one's ethnic identity results in academic underachievement and hurdles in future career prospects. In our paper however, the effect comes through a different channel. In this paper, we rather compare the labor market outcomes of the kids growing up in an ethnically mixed neighborhoods vis-a-vis those growing up in (ethnically) homogenous communities. We explain our proposed mechanism using a model below. But in a nutshell, we argue that in a mixed neighbourhood, kids are exposed to a more diverse peer learning and higher aspirational norms which pays-off when they enter the labor market.

Our paper is related to the general economics literature that focuses on inter-generational economic mobility. In their pioneering paper, Galor and Zeira (1993) showed that given credit market imperfection and lumpiness in educational investment, inter-generational economic mobility through

[2]Endogenous effects capture the effects wherein an individual's behaviour varies with the behaviour of the group. Contextual effects are the effects wherein individual's behaviour varies with the exogenous characteristics of the group. In case of correlated effects, the members of a group behave in a similar manner because they face similar institutional environment.

human capital accumulation may not be possible; high wealth families will invest in education and go to high paying skilled occupations, while their low wealth counterparts will fall back in the trap of unskilled occupations. In a related paper that links community formation and human capital investment, Benabou (1996) models how heterogenous families form communities, choose local public expenditures and accumulate human capital. Then he goes on to show how a minor difference in education technologies and wealth lead to high degree of divergence across communities. The issue of residential segregation through skill formation is analyzed in an earlier paper as well (Benabou, 1993).

The labour market advantages transferred by affluent parents to their offspring are not just restricted to the more obvious benefits from schooling and inter-generational wealth. A sizeable contribution actually comes from personality, psychology, and behavioural skills that are passed on by parents and acquaintances who possess a high social and economic status (Bowles et al., 2001). Continuing further along this vein, Collier and Gunning (1999) in their study of African growth state that it is not ethnic fragmentation by itself, but ethnic fragmentation combined with low levels of political rights, that in turn leads to low levels of social capital formation. And this social capital, specifically the informational aspect of it - which is learnt through one's social network and peer groups - ultimately has a bigger impact on the adoption of new techniques and innovations, than factor endowments such as formal education.

Our paper also contributes to the section of the literature that discusses the issue of inter-generational mobility in the context of Indian economy. Besides the papers that looked at the role of market and state in bringing in inter-caste convergence, studies have looked at the patterns of inter-generational mobility in education and occupation. For example, Azam and Bhatt (2015) found that for the birth cohort of 1940-1985, educational persistence between the father and the son declined for the lower tail of fathers' educational distribution, but it increased for the upper tail. In another paper, Reddy (2015) found that after controlling for the change in the occupational structure over the years, inter-generational occupational mobility declined during 1983-2012. However, the decline was sharper for the disadvantaged castes (scheduled castes and tribes). Using a more recent, novel administrative data set, Asher et al. (2018) found that on an average the intergenerational occupational mobility remained constant for the entire population. But occupational mobility changed for different

groups – rise in mobility for the scheduled castes has been almost offset by the decline in mobility for the Muslims. They also find that mobility for children at the lowest tail of the income distribution is higher in southern, urban India and places where average education level is higher.

There is not much research on residential segregation in India. In one such study, Adukia et al. (2019) showed that both Scheduled castes/tribes (SC/ST) and Muslims are concentrated in poor cities and this is more so for Muslims. Cities with more Muslims have worse public goods than cities with more SC/ST. If we look within cities, Muslim and SC/ST neighborhoods are characterized by lower consumption and access to public goods. One key mechanism that leads to segregation is lack of mobility in Indian labor market. In a seminal work on this issue, Munshi and Rosenzweig (2009) showed that mobility rates across caste-based societies in India have been quite stagnant. This is observed for both social mobility (intermarriages) as well as spatial mobility (migration for the purpose of job hunting). The main reason behind this, according to the authors, is that communities founded on identities of castes and sub-castes (*jatis* are bound by mutual insurance, a stable system of sharing risks and rewards. The higher the economic reward within this social arrangement, the lower will be the incentive for mobility. While such immobility might be countered by better economic and risk-sharing opportunities outside of these caste networks, they would probably not be enough to surpass the social insurance provided within a *jati*, for the poorer sections of these societies.

The next key issue is the welfare implication of segregation. A key element in determining whether increasing diversity — in terms of marginalised members moving into a community — will be faced with acceptance or backlash depends on whether native residents see the entrants as an opposing group competing for resources, or positive inter-group interactions as individuals. This is illustrated by Posta (2013), which shows that while competitive threat is stronger at the state/regional level, contact hypothesis is stronger at the commune (town/city) level where individual interactions are possible, and leads to a reduction in prejudice, further leading to a significant fall in anti-immigration attitudes. Alesina and Tabellini (2022) states that immigrant backlash depends heavily on non-economic factors like misperceptions and stereotyping, hyper-nationalistic rhetoric, and media bias. However, if conditions for inter-group contact are encouraged, it would prompt natives' sensitization and acceptance, potentially snowballing liberal outlook into the general native population.

In this background, we in this paper examine whether growing up in a mixed neighbourhood have any impact for economic mobility of Indian workers. The rest of this paper is ordered in the following sequence: Section 2 will outline the Data and Methodology. Section 3 will describe our theoretical Model, followed by Section 4, Data Analysis, which will explain the empirical findings. Next, Section 5 will lay out the Results of our analysis, with detailed interpretation of our findings. Finally, Section 6, Conclusion, will wrap it up with an interpretation of the results, along with possible prospects and implications therein.

2 Conceptual Framework

2.1 Model

We construct a theoretical model that presents an analysis of skill acquisition within the social framework of caste hierarchy, and its link to wage earning capabilities. We consider a community with two castes, one higher on the socioeconomic scale, the other lower. We refer to them as castes h and l respectively. Our model only looks at transitional mobility between two generations and do not explore the implications of long run equilibrium. We simply look at a generation's skill acquisition incentives, given the caste composition of a locality, the proportion of skilled workers in the social groups (castes) and the degree of social privilege enjoyed by the upper castes in the caste hierarchy.

We normalize the village population to 1. Let α be the fraction of population belonging to h, the privileged caste. There are two types of occupations - skilled and unskilled with respective wages w_s and w_n. We normalize w_n to 0. The determination of w_s is explained later. The fraction of caste i who are engaged in skilled occupation are represented by p_i, where $i = h, l$, with $p_h > p_l \geq 0$. The total number of skilled workers in the village is thus $\theta = \alpha p_h + (1 - \alpha) p_l$.

We assume that each member of the community has one descendant who decides whether to acquire skill or not. Skill acquisition is costly. The cost of skill acquisition depends on the existing number of skilled workers in the village and ones position in social hierarchy. Thus, for a person belonging to caste i, the cost of skill acquisition is $c_i(\theta)$, with $c_i'(.) \leq 0$ and $c_h(\theta) \leq c_l(\theta)$ for all θ. To avoid analytical complications, we assume linear cost functions for both types. More specifically, we assume

$$c_l(\theta) = 1 - \theta$$

and

$$c_h(\theta) = \beta - \theta$$

with $\beta < 1$. In our model, β captures the social privilege of the higher caste with lower β implying higher privilege.

In our model, the earning from the high skilled occupation depends on a person's ability. We assume that ability is uniformly distributed across caste. A person's ability $a \in [0,1]$ determines the skilled wage w_s. We assume that $w_s = w + a$ without loss of generality where $w \in (0,1)$ is the minimum earning from skilled occupation.

An individual belonging to caste i acquires skill iff

$$a - c_i(\theta) \geq 0$$

$$\Leftrightarrow a \geq c_i(\theta)$$

Given our specification of the cost function, we encounter different scenarios regarding skill acquisition depending on initial values of θ.

Under our specification, from the lower caste group only people with ability $a \geq 1 - \theta$ would acquire skill. For $\theta < \beta$, only people with ability $a \geq \beta - \theta$ from the higher caste group will acquire skill. On the other hand, for $\theta \geq \beta$, people of all abilities from the higher caste group will acquire skill. Thus,

$$p_l' = \theta$$

and

$$p_h' = \begin{cases} 1 - \beta + \theta & \text{for } \theta < \beta \\ 1 & \text{for } \theta \geq \beta \end{cases}$$

Hence, the fraction of skilled workers in the population would be

$$\begin{aligned} \theta' &= \alpha p_h' + (1-\alpha) p_l' \\ &= \begin{cases} \alpha(1 - \beta + \theta) + (1-\alpha)\theta & \text{for } \theta \in [0, \beta) \\ \alpha + (1-\alpha)\theta & \text{for } \theta \in [\beta, 1] \end{cases} \end{aligned}$$

Notice that in our model $\theta \in [p_l, p_h]$. If $\beta > p_h$, $\theta < \beta$ for all $\alpha \in [0,1]$. In this case, $\theta' = \alpha(1-\beta) + \theta$. On the other hand, if $\beta < p_l$, $\theta' = \alpha(1-\theta) + \theta$. The more interesting case arises when $\beta \in (p_l, p_h)$. Since θ is strictly rising in α, there exists a critical $\bar{\alpha}$, such that $\theta' = \alpha(1-\beta) + \theta$ for $\alpha < \bar{\alpha}$ while $\theta' = \alpha(1-\theta) + \theta$ for $\alpha \geq \bar{\alpha}$.

Our measure for occupational mobility in the village is the change in the number of skilled workers between two periods, i.e. $\theta' - \theta$. We now state our main result in following proposition.

If the social privilege enjoyed by the upper caste is relatively low ($\beta > p_h$), occupational mobility rises as the fraction of upper caste in the population rises. Otherwise, the relationship between occupational mobility and the caste composition depends on the proportions of the skilled workers among different castes. If the proportions of skilled workers among higher and lower castes are relatively close ($p_h \leq \frac{1+p_l}{2}$), occupational mobility rises as the fraction of upper caste in the population rises. But if the ratio of skilled workers among the upper caste is much higher than the same ratio among the lower caste ($p_h > \frac{1+p_l}{2}$), occupational mobility first rises and then falls as the fraction of upper caste in the population rises.

First consider the case $\beta > p_h$. In this case, $\theta' - \theta = \alpha(1-\beta)$. Hence, occupational mobility rises with α.

Now suppose $\beta < p_l$. Then our measure for mobility is

$$\theta' - \theta = \alpha(1-\theta)$$

In this case,

$$\begin{aligned} \frac{d}{d\alpha}(\theta' - \theta) &= 1 - \theta - \alpha\frac{d\theta}{d\alpha} \\ &= 1 - p_l - 2\alpha(p_h - p_l) \end{aligned}$$

If $p_h \leq \frac{1+p_l}{2}$, $\frac{d}{d\alpha}(\theta' - \theta) \geq 0$ for $\alpha \in [0,1]$. On the other hand, if $p_h > \frac{1+p_l}{2}$,

$$\frac{d}{d\alpha}(\theta' - \theta) \geq 0$$

if and only if

$$\alpha \leq \tilde{\alpha} = \frac{1-p_l}{2(p_h - p_l)}$$

Hence, $\theta' - \theta$ rises with α if and only if $\alpha < \tilde{\alpha}$.

Finally, consider the case $p_l < \beta < p_h$. For $\alpha < \bar{\alpha}$, $\theta' - \theta$ rises with α. On the other hand, for $\alpha \geq \bar{\alpha}$, $\theta' - \theta$ continues to rise with α for all $\alpha \in [\bar{\alpha}, 1]$ if $p_h \leq \frac{1+p_l}{2}$. However, for $p_h > \frac{1+p_l}{2}$, $\theta' - \theta$ rises with α if and only if $\alpha < \tilde{\alpha}$. Hence, $\theta' - \theta$ rises with α if and only if $\alpha < \max\{\bar{\alpha}, \tilde{\alpha}\}$. This completes the proof of the proposition.

The relationship between $\theta' - \theta$ (mobility, caused by acquiring better skills) and α (the fraction of privileged caste population within a locality) resemble an inverted-U curve when the community is characterized by high level of privilege for the upper caste as well as a marked difference between the proportions of skilled workers between the two castes. At the two extreme

points of α, skill learning is at its lowest. With α close to 0, we have a locality with only underprivileged, skill-deprived residents, with little access to the know-how required for occupational betterment. At the other end, when α is close to 1, we are left with a locality consisting of privileged but similarly high-skilled people who have no more to offer each other in terms of upward mobility. Our theory predicts that with caste privilege for upper caste and historically given significantly large proportion of skilled workers coming from the upper caste, the more socially heterogeneous the locality, the higher the prospects for skill acquisition and upward wage mobility.

2.2 Implication

For underprivileged castes, a possible way out of this vicious cycle of impoverishment might be through a gradual and deliberate osmosis of skills and knowledge from the more privileged. Each budding generation finds its teachers and guides in their preceding generation. While this tutelage could potentially come from any educator-figure in their vicinity, the easiest skills to learn for a child would be to replicate that of a parent's. That is, for those choosing to stay in the same occupation as their parent, acquiring those skills is the easiest path to take, with a reasonably sure outcome. However, if an individual wants to branch out of such familial constraints and upwards on the occupational ladder, they will need to learn new skills for higher paying jobs, which are more likely to be held by those from higher castes. So we need to consider the probability of meeting a higher caste individual — and hence gaining access to occupational knowledge and skill. The more heterogeneous the *mohalla* (neighbourhood), the higher the odds of getting to learn about different occupations and opportunities that might be available.

The knowledge and skill development required to climb up the occupational ladder consists not only of academic education and technical know-how, but also soft skills: a knack for interpersonal communication, building bridges and weaving networks. So while better schooling is definitely a start, a key component of inter-generational socioeconomic betterment is getting to learn first-hand the ins and outs of a profession from someone who already works in the same field. Teaching the minutiae and subtleties of a profession goes beyond formal schooling. On that account, we can safely infer that the presence of highly skilled individuals is an essential requirement for the current generation of learners in a neighbourhood

or locality. And the higher the number of high skilled neighbours, the greater the transfer, dissemination, and assimilation of these skills would be. However, moving to a higher skilled job — beyond the skills that are much easier to teach or pass on from parent to child — through the mixed *mohalla* mechanism comes with its costs and risks.

3 Data and Empirical Model

In this section we examine the empirical validity of our theoretical predictions. For our empirical section we use the Indian Human Development Survey data (survey year 2011-12). IHDS database was initially formed through a survey of 41554 households in 1503 villages and 971 urban areas spanning across 35 Indian states and union territories conducted by Indian Council of Applied Economic Research (NCAER), New Delhi and University of Maryland in 2004-05. The survey consists of two parts, household questionnaire with household characteristics on demography, health, education, income, work, occupation, production, consumption, assets, social capital, fertility, children schooling, etc. and individual questionnaire with work, income, gender relation, fertility decision, marriage practices, exposure to mass media, reading, writing skill etc. The respondent households of 2005 survey were re-interviewed in 2011-12 to form a longitudinal database. The number of households increased slightly in the second round and it interviewed 42152 households. We however, do not use both the rounds. We only use the more recent one which was conducted in 2011-12. Using this data we estimate the following equation:

$$Y_i = \beta_0 + \beta_1 N_i + \beta_2 X_i' + \epsilon_i \tag{1}$$

where Y_i is a measure of inter-generational economic mobility, N_i represents ethnic heterogeneity in the neighbourhood individual i is living and X_i' represents other individual specific controls. In the following subsections we discuss these variables.

3.1 Mobility Measure

A vital cornerstone in our empirical analysis is the concept of economic mobility. In the literature, the standard way of measuring inter-generational mobility is to rank occupations according to wage and skill level and then examine whether child's occupation's is higher/lower ranking than that of the parents. In one such study on India, Motiram and Singh (2012) uses National Classification of Occupations (NCO 2004) to classify occupations

in 10 single digit codes viz. (1) legislators, senior officials and managers; (2) professionals; (3) technicians and associate professionals; (4) clerks; (5) service workers and shop and market sales workers; (6) skilled agricultural and fishery workers; (7) craft and related trades workers; (8) plant and machine operators and assemblers; (9) elementary occupations (e g, sweepers, street vendors, loaders, etc); and (10) workers not classified by occupations. It is clear from the classification that as one move from category 1 to 10, wage goes down. Then the authors create a transition matrix to examine if, compared to the parents, the children moved up or down in the ladder.

We, in this paper, have taken a different approach. We know the occupations of both the respondents (children) and their fathers. Rather than classifying occupations in high paying and low paying jobs in an ad-hoc manner, we calculate the state level average wage of the occupation the father was in. But the wage data is from the current time. Hence, this is not the wage the father actually earned. Rather, this is the average wage the child would have earned now, had he (she) stuck to the occupation of his (her) father. Hence, this is a counterfactual wage. Let's call this w_i^f. We also have current wage of the child – w_i^c. For the i^{th} family, our mobility measure is defined as

$$M_i = w_i^c - w_i^f$$

If $M_i > 0$, the son is earning more than what he would have earned if he had stuck to his father's occupation. This is a measure of upward mobility. On the other hand, there will be downward mobility if $M_i < 0$. Overall, the degree of mobility rises with rise in M_i.

3.2 Measure of Neighbourhood Heterogeneity

The IHDS documents whether the respondents reside in mixed *mohallas* or not (from the Village Questionnaire: "In your village do different *jatis*/groups reside in separate hamlet/*mohalla*/locality or do they live together?"). The dummy variable associated with this question takes the value 1 if they live together, 0 otherwise.

3.2.1 Social Capital

The respondents are asked whether they know any doctor, teacher, school worker, other government employee, elected politician, political party official, police and military personnel within or outside their castes/communities. Using responses to thess questions, we construct two social

capital variables – within community social capital (SCW) and outside community social capital (SCO). If the respondent knows any of the above mentioned people within community, SCW takes 1, and 0 otherwise. Similarly, if the respondent knows any of the above mentioned people within community, SCO takes 1, and 0 otherwise. We use both of these variables in our regression. In this definition of social capital, connection with people holding important position is established and therefore, we can call it vertical social capital.

There is another type of social capital captured by membership in different groups. Because of its nature, we can call it horizontal social capital. Some villages provide options for participation in groups of various kinds, categorised by hobbies, political interests, or other social organisations. In order to examine the effect this sort of fraternisation might have on job prospects, we have also considered the presence of such groups in the surveyed villages, and the respondents' membership in them, if any.

3.3 Control Variables

Next we introduce control variables, focusing on indicators of social capital and local Institution. We have generated variables to check for the respondents' sex, age, educational qualification, and caste. We have taken in to account the difference in educational qualification between the household head and their parent. We have also included a variable measuring distance to the nearest town, in order to limit the effect of rural-urban divides.

4 Results

4.1 Descriptive Statistics

Table 1 presents the descriptive statistics for our key variables.

Table 1 here

A few values that stand out here include extreme negatively skewed with high kurtosis results of our mobility measure. The severely leptokurtic result is due to the presence of outliers, while the left-skewed result tells us that the bulk of these outliers lie on the left tail, which is also reflected in the overall range of values for mobility. The mean of mobility is negative. The presence of groups in the respondent's village is also similarly negatively

skewed and leptokurtic, possibly because of the heavy presence of outliers on the lower end. In other words, a considerable number of villages have no such group presence. The variable for confidence in courts shows similar results as well.

Next we plot the mean of mobility for each caste group.

Figure 1 here

While mean of mobility is negative for all caste groups, with no other parameters being considered, it worsens as we climb down the caste hierarchy. However, if we separate these caste groups further by whether or not they reside in mixed localities, we get a clearer picture.

Figure 2 here

For Brahmin, Forward, and ST caste groups, mean mobility worsens in mixed neighbourhoods. This result is much starker for Brahmins. On the other hand, OBC and SC groups are better off in mixed neighbourhoods.

4.2 Baseline Regression

In our baseline regression, we regress our measure of mobility on mixed neighbourhood dummy along with other control variables. The control variables include a dummy for female respondent, whether the respondent has social capital outside his/her own community (SCW and SCO), whether the respondent has social capital within his/her own community, age, education, inter-generational education difference, distance from the nearest town, whether groups such as self help groups exist in the village, whether the respondent has a group membership, confidence in police and confidence in court.

Table 2 here

In the column (1) of table (2) we regress our measure of wage mobility on mixed neighbourhood and other control variables for the full sample. We find that the coefficient for the mixed neighborhood is negative but not significant. We also find that female respondents are associated with lower degree of mobility than their male counterpart. Age is associated with lower mobility. As expected, inter-generational educational difference is positively associated with upward wage mobility – the children earn more than their parents if they are more educated than their parents.

More interestingly, vertical social capital within one's own community is negatively associated with wage mobility while the coefficient for outside community social capital (SCO) is positive but not significant. Existence of groups in village is positively associated with upward mobility but membership in these groups have no significant impact.

Next, we split our full sample in separate cast groupings to check their mobility results separately. We run the same regression for upper castes and lower caste groups and report the results in columns (2) and (3). In columns (4)-(6), we report the regression results for difference lower caste groups such as Other Backward Class (OBC), scheduled caste (SC) and scheduled tribes (ST). We find that coefficient for mixed neighbourhood is negative and significant for upper castes. For the aggregated lower caste groups the coefficient is positive but not significant. The same is true for SC and OBC groups. For ST, the coefficient is negative but not significant. For the female dummy, the coefficient is negative for all caste groups. The coefficient for the age variable is negative and significant for all the groups. Education for the respondent is not significant for any groups. But inter-generational educational difference is positive significant for both the aggregate caste categories of upper and lower castes. Among the lower caste groups, the coefficient is only positive and significant for the scheduled castes (SC).

For the results concerning social capital outside community (SCO), we find that the coefficients are not significant for any of the groups. For within community social capital (SCW) the coefficient is negative for the full sample and lower castes. We find that the results is being driven by the SC group as they are the only group with negative, significant coefficient.

Our hypothesis is that the driving force behind generational occupation/earning mobility is the heterogeneity of one's locality, and the social capital accumulated through interaction within a diverse neighbourhood. Our baseline regression strategy however, cannot deal with the possibility of reverse causality – people from different ethnicities are likely to move to areas with greater economic opportunities which in turn leads to greater degree of economic mobility. Such possibility can be seen as a variant of the tiebout sorting (Tiebout, 1956). To rule out this problem, at least partially, we considers only those residents who have been living here for at least 20 years (the approximate length of a generation) or more. The results are reported in table 3.

Table 3 here

We run the same regression as before, but this time it is conditional upon years of residence of the respondent or their family having lived there being at least 20 years if not more. Table 1B demonstrates that our empirical results remain very similar to that of our previous regression. As before, the coefficient for mixed neighbourhood is negative and significant for Upper Castes, for the female dummy for all groups, and also for increasing age for all groups (though not significant for the Upper Caste group). Social capital within one's community still shows a negative and significant association with mobility for the entire sample, as well as for the group of Lower Caste, and the SC group. Inter-generational educational difference continues to play an important and positive role in encouraging mobility (though not statistically significant for the OBC and ST groups). The presence of groups in the village also displays a positive and significant association for the full sample, the Upper Castes, and the OBC. This rules out any endogeneity error, and our initial hypothesis continues to hold. To investigate the issue further, in the next subsection, we examine the effect of the interaction between mixed neighbourhood and social capital. We carry out this analysis only for the long term residents.

4.3 The Mechanism

In our theory, we argue that neighbourhood level ethnic heterogeneity helps economic mobility through skill learning from other communities. Given that the level of skill is historically high among the upper caste, the lower castes are more likely to benefit from social mixing. Hence, to unbox the mechanism empirically, we divide our respondents by caste and see the interaction of neighbourhood heterogeneity with one's social network *within* and *outside* one's community. In our first analysis of the underlying mechanism, we regress our mobility measure on the interaction between social capital within community and the mixed neighbourhood dummy. The results are reported in table 4. The control variables added to regression include a dummy for female sex of respondent, the respondent's age, education level achieved, inter-generational difference in education level between the respondent and their parent, distance from the nearest town, the presence of groups like self help-groups in their village, the respondent's membership in such groups, their confidence in the police and in the courts.

Table 4 here

In Table 4, we find that similar to the baseline specification, significant negative associations exist for female respondents throughout the groups, and the coefficient for age is negative throughout and significant in all but the privileged castes. Inter-generational educational difference has positive and significant association for the privileged and underprivileged caste groups. But once we split the underprivileged population into separate sections by caste, the SC is the only underprivileged caste with such positive and significant association. The presence of groups in villages is positive and significant for the privileged castes, and for the OBC. The OBC groups also shows a negatively significant association between mobility and confidence in the police. The SC group has negative and significant association with the respondent's education level. The ST group on the other hand, show positively significant association for the same.

The main variable of interest in this table is the interaction between within caste social capital and mixed neighbourhood. We find that the coefficient is negative and significant for the upper caste groups. For all other groups including the full sample the coefficients are not significant. This result means that if upper caste people know important people within their community, they are less likely to achieve economic mobility in a mixed neighbourhood.

Next, we focus on one's status regarding social capital from outside the community. The results are reported in table 5 Like the last analysis, we only take long term residents to avoid endogenous sorting. In this analysis also, we take the same control variables such as sex dummy, age, education, intergenerational education difference, group membership status and trust on the formal institutions. The signs of the coefficients are almost the same as the last table – negative and significant coefficient for female dummy and age and positive, significant effect of inter-generational education difference. We also find positive effect of existence of group in the village for upper caste and OBC. But no effect of group membership.

Table 5 here

In this table the main variable of interest is the interaction between social capital outside community and mixed neighbourhood. We find that mixed neighbourhood has a positive impact on economic mobility of the lower castes if they have social capital outside community. From the result reported in column (5) we also find that the result is driven by scheduled castes. This result is consistent with our expectation. The result essentially

means that the underprivileged castes benefit from mixed neighbourhood if they know people outside their community. Similar to the result generated by the model, this means that in mixed neighbourhood the lower caste people get in touch with upper castes who leave a positive impact on them either by training or by role model. This mechanism leads to economic mobility for the lower caste people.

5 Conclusion

In this paper, we examine the effect of living in a ethnically mixed neighbourhood on one's economic mobility. We find that on an average there is not much effect of the mixed neighbourhood on one's economic mobility – the coefficient is negative but not significant. However, when we divide the sample according to the caste of the respondents an interesting pattern emerges. The effect of mixed neighbourhood on economic mobility is negative and significant for upper caste. The coefficients for mixed neighborhood is however, positive when we consider respondents from SC, ST and OBC caste groups. But they are not significant. To deal with endogeneity issue arising from Tiebout sorting, we then restrict the sample for long time residents – people who lived in the same place for twenty years or more. The results remain the same qualitatively.

More interesting patterns emerge when we do a heterogeneity analysis using two types of social capital – social capital within community and social capital outside community. We find that for upper caste members who have social capital within community, living in mixed neighbourhood leads to less economic mobility. No such significant effect for other caste groups. A different pattern emerge when we look at the social capital outside the community. We find that low caste members who live in a mixed neighbourhood, experience higher economic mobility if they have social capital outside their caste. This result is particularly strong for scheduled caste.

The empirical results are consistent with our theoretical predictions. A key policy prescription that comes from our result is the one that discourages *ghettoization* and encourages living in the mixed neighbourhood.

6 Tables and Figures

6.1 Tables

Table 1: Descriptive Statistics

	Observations	Minimum	Maximum	Mean	SD	Variance	Skewness	Kurtosis
Mobility	22466	-131.7827	0.9587	-0.3123	1.4708	2.1633	-36.9173	2936.2440
Mixed Neighbourhood	26230	0	1	0.3868	0.4870	0.2372	0.4650	1.2162
SCO	40654	0	1	0.6927	0.4614	0.2129	-0.8352	1.6976
SCW	40654	0	1	0.5647	0.4958	0.2458	-0.2611	1.0682
Female Respondent	40654	0	1	0.1436	0.3506	0.1229	2.0331	5.1337
Age	40654	1	91	41.7627	13.5983	184.9148	0.2813	2.5844
Education	40625	1	17	6.6786	4.9238	24.2440	0.3097	1.9690
Intergen Edu Diff	40248	-15	17	3.9706	5.5281	30.5596	-0.0571	3.0078
Distance from Nearest Town	26201	0	110	13.5498	10.9598	120.1170	2.1860	12.1115
Groups in Village	40654	0	1	0.9955	0.0669	0.0045	-14.8040	220.1575
Group Membership	40654	0	1	0.4057	0.4910	0.2411	0.3842	1.1476
Confidence in Police	40654	0	1	0.7736	0.4185	0.1752	-1.3074	2.7092
Confidence in Courts	40654	0	1	0.9306	0.2542	0.0646	-3.3883	12.4807

Table 2: Baseline Regression

VARIABLES	(1) Full Sample	(2) Upper Caste	(3) Lower Caste	(4) OBC	(5) SC	(6) ST
Mixed Neighbourhood	-0.0241	-0.144**	0.00272	0.0245	0.0204	-0.105
	(0.0349)	(0.0550)	(0.0356)	(0.0339)	(0.0728)	(0.0662)
Female Respondent	-1.062***	-1.321***	-1.027***	-0.937***	-1.317***	-0.639***
	(0.136)	(0.212)	(0.158)	(0.0990)	(0.404)	(0.144)
Social Capital outside Community (SCO)	0.0251	-0.0259	0.0324	-0.00875	0.0475	0.107
	(0.0303)	(0.0489)	(0.0327)	(0.0318)	(0.0696)	(0.0862)
Social capital within Community (SCW)	-0.0860**	-0.0143	-0.101**	-0.0518	-0.138**	-0.142
	(0.0331)	(0.0519)	(0.0413)	(0.0338)	(0.0688)	(0.129)
Age	-0.00447***	-0.00185	-0.00519***	-0.00593***	-0.00533**	-0.00354**
	(0.00100)	(0.00225)	(0.00128)	(0.00164)	(0.00262)	(0.00168)
Education	-0.00729	-0.00956	-0.00733	-0.00366	-0.0245**	0.0108
	(0.00537)	(0.0119)	(0.00555)	(0.00645)	(0.0119)	(0.00651)
Intergen Edu Diff	0.0130**	0.0277**	0.00954**	0.00602	0.0214***	-0.00249
	(0.00516)	(0.0121)	(0.00447)	(0.00510)	(0.00754)	(0.00789)
Distance from Nearest Town	-0.00317	-0.00274	-0.00326	-0.00330	-3.50e-05	-0.00492
	(0.00214)	(0.00294)	(0.00242)	(0.00255)	(0.00200)	(0.00627)
Groups in Village	0.249*	0.570***	0.205	0.425**	0.0852	-0.0847
	(0.126)	(0.157)	(0.128)	(0.196)	(0.131)	(0.0737)
Group Membership	0.0281	0.0117	0.0297	0.0605	0.0212	-0.00991
	(0.0292)	(0.0669)	(0.0333)	(0.0379)	(0.0642)	(0.0611)
Confidence in Police	0.0543	0.0727	0.0515	-0.0644**	0.175	0.155
	(0.0443)	(0.0917)	(0.0483)	(0.0321)	(0.153)	(0.102)
Confidence in Court	0.0298	0.179	0.00423	0.0549	-0.0584	0.0112
	(0.0458)	(0.122)	(0.0520)	(0.0585)	(0.101)	(0.144)
Constant	-0.358**	-0.920***	-0.261*	-0.403*	-0.146	-0.236
	(0.137)	(0.254)	(0.139)	(0.206)	(0.183)	(0.190)
Observations	14,575	2,429	12,146	5,646	4,506	1,988
R-squared	0.046	0.121	0.040	0.088	0.032	0.057

Robust standard errors in parentheses
$***p < 0.01, **p < 0.05, *p < 0.1$

Table 3: Baseline Regression for Long Term Residents

VARIABLES	(1) Full Sample	(2) Upper Caste	(3) Lower Caste	(4) OBC	(5) SC	(6) ST
Mixed Neighbourhood	-0.0223	-0.150**	0.00572	0.0266	0.0192	-0.0893
	(0.0348)	(0.0571)	(0.0355)	(0.0339)	(0.0735)	(0.0685)
Female Respondent	-1.074***	-1.365***	-1.036***	-0.937***	-1.344***	-0.636***
	(0.139)	(0.221)	(0.162)	(0.100)	(0.416)	(0.141)
Social Capital outside Community	0.0234	-0.0332	0.0320	-0.0129	0.0471	0.115
	(0.0303)	(0.0515)	(0.0328)	(0.0330)	(0.0691)	(0.0829)
Social capital within Community	-0.0858**	-0.0155	-0.100**	-0.0497	-0.138*	-0.142
	(0.0334)	(0.0522)	(0.0418)	(0.0356)	(0.0692)	(0.130)
Age	-0.00460***	-0.00183	-0.00533***	-0.00599***	-0.00560**	-0.00361**
	(0.00104)	(0.00237)	(0.00130)	(0.00167)	(0.00263)	(0.00173)
Education	-0.00894	-0.00958	-0.00943	-0.00464	-0.0268**	0.00609
	(0.00560)	(0.0124)	(0.00570)	(0.00655)	(0.0125)	(0.00667)
Intergen Edu Diff	0.0141***	0.0274**	0.0110**	0.00671	0.0225***	0.00170
	(0.00524)	(0.0126)	(0.00447)	(0.00515)	(0.00781)	(0.00821)
Distance from Nearest Town	-0.00334	-0.00275	-0.00344	-0.00328	-0.000187	-0.00532
	(0.00225)	(0.00300)	(0.00253)	(0.00258)	(0.00224)	(0.00654)
Groups in Village	0.244*	0.553***	0.201	0.422**	0.0795	-0.0905
	(0.127)	(0.155)	(0.128)	(0.196)	(0.131)	(0.0725)
Group Membership	0.0239	0.00815	0.0254	0.0604	0.0138	-0.0108
	(0.0302)	(0.0705)	(0.0343)	(0.0392)	(0.0666)	(0.0628)
Confidence in Police	0.0594	0.0800	0.0569	-0.0631*	0.178	0.164
	(0.0453)	(0.0958)	(0.0496)	(0.0323)	(0.156)	(0.105)
Confidence in Court	0.0329	0.181	0.00740	0.0588	-0.0590	0.0210
	(0.0464)	(0.135)	(0.0517)	(0.0594)	(0.0996)	(0.145)
Constant	-0.351**	-0.906***	-0.255*	-0.403*	-0.123	-0.245
	(0.139)	(0.264)	(0.140)	(0.206)	(0.185)	(0.194)
Observations	14,255	2,340	11,915	5,533	4,433	1,943
R-squared	0.046	0.123	0.040	0.087	0.033	0.058

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Heterogeneity Analysis 1: SCW for Long Term Residents

VARIABLES	(1) Full Sample	(2) Upper Caste	(3) Lower Caste	(4) OBC	(5) SC	(6) ST
Social Capital within Community	-0.0569	0.0756	-0.0870	-0.0192	-0.161	-0.0995
	(0.0446)	(0.0580)	(0.0562)	(0.0414)	(0.125)	(0.150)
Mixed Neighbourhood	0.00612	-0.00941	0.00961	0.0727**	-0.0371	-0.0960
	(0.0310)	(0.0818)	(0.0317)	(0.0308)	(0.0457)	(0.106)
SCW*Mixed Neighbourhood	-0.0549	-0.241**	-0.00774	-0.0943	0.112	0.00122
	(0.0484)	(0.0954)	(0.0555)	(0.0661)	(0.160)	(0.116)
Female Respondent	-1.074***	-1.371***	-1.035***	-0.937***	-1.342***	-0.639***
	(0.139)	(0.219)	(0.162)	(0.101)	(0.417)	(0.142)
Age	-0.00460***	-0.00179	-0.00536***	-0.00595***	-0.00569**	-0.00361**
	(0.00104)	(0.00230)	(0.00130)	(0.00165)	(0.00264)	(0.00173)
Education	-0.00866	-0.00889	-0.00913	-0.00472	-0.0265**	0.00700
	(0.00559)	(0.0122)	(0.00570)	(0.00670)	(0.0126)	(0.00648)
Intergen Edu Diff	0.0138***	0.0265**	0.0107**	0.00663	0.0220***	0.000565
	(0.00512)	(0.0123)	(0.00430)	(0.00523)	(0.00756)	(0.00792)
Distance from Nearest Town	-0.00327	-0.00243	-0.00338	-0.00325	-0.000134	-0.00515
	(0.00228)	(0.00299)	(0.00257)	(0.00261)	(0.00230)	(0.00671)
Groups in Village	0.245*	0.532***	0.203	0.412**	0.0677	-0.0573
	(0.122)	(0.150)	(0.126)	(0.193)	(0.121)	(0.0801)
Group Membership	0.0260	0.0181	0.0274	0.0587	0.0129	-0.0170
	(0.0308)	(0.0715)	(0.0348)	(0.0384)	(0.0722)	(0.0615)
Confidence in Police	0.0598	0.0832	0.0562	-0.0612*	0.176	0.170
	(0.0454)	(0.0948)	(0.0499)	(0.0330)	(0.152)	(0.104)
Confidence in Court	0.0330	0.163	0.00556	0.0596	-0.0726	0.00235
	(0.0464)	(0.122)	(0.0528)	(0.0597)	(0.106)	(0.157)
Constant	-0.353**	-0.952***	-0.241*	-0.418**	-0.0545	-0.207
	(0.132)	(0.254)	(0.139)	(0.202)	(0.172)	(0.202)
Observations	14,255	2,340	11,915	5,533	4,433	1,943
R-squared	0.046	0.125	0.040	0.087	0.033	0.056

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Heterogeneity Analysis 2: SCO for Long Term Residents

VARIABLES	(1) Full Sample	(2) Upper Caste	(3) Lower Caste	(4) OBC	(5) SC	(6) ST
Social Capital outside Community	-0.0365	0.0140	-0.0477	-0.0266	-0.118	0.0707
	(0.0417)	(0.0503)	(0.0477)	(0.0404)	(0.111)	(0.0954)
Mixed Neighbourhood	-0.0762	-0.0656	-0.0772	0.0287	-0.191*	-0.0769
	(0.0471)	(0.0598)	(0.0529)	(0.0397)	(0.103)	(0.131)
SCO*Mixed Neighbourhood	0.0772	-0.125	0.121*	-0.00512	0.321**	-0.0306
	(0.0605)	(0.116)	(0.0656)	(0.0680)	(0.152)	(0.132)
Female Respondent	-1.075***	-1.366***	-1.036***	-0.938***	-1.344***	-0.634***
	(0.139)	(0.221)	(0.162)	(0.101)	(0.416)	(0.141)
Age	-0.00470***	-0.00189	-0.00542***	-0.00600***	-0.00585**	-0.00394**
	(0.00104)	(0.00234)	(0.00130)	(0.00168)	(0.00256)	(0.00181)
Education	-0.0117*	-0.00943	-0.0126*	-0.00611	-0.0305**	0.00175
	(0.00602)	(0.0120)	(0.00632)	(0.00689)	(0.0141)	(0.00758)
Intergen Edu Diff	0.0160***	0.0273**	0.0133***	0.00778	0.0246***	0.00455
	(0.00540)	(0.0123)	(0.00449)	(0.00518)	(0.00797)	(0.00761)
Distance from Nearest Town	-0.00327	-0.00266	-0.00331	-0.00322	-0.000115	-0.00515
	(0.00227)	(0.00296)	(0.00257)	(0.00257)	(0.00218)	(0.00657)
Groups in Village	0.236*	0.560***	0.189	0.436**	0.0297	-0.102
	(0.127)	(0.150)	(0.131)	(0.196)	(0.130)	(0.0796)
Group Membership	0.0193	0.00732	0.0218	0.0598	0.00577	-0.0249
	(0.0299)	(0.0711)	(0.0335)	(0.0389)	(0.0662)	(0.0590)
Confidence in Police	0.0638	0.0803	0.0619	-0.0628*	0.188	0.171
	(0.0464)	(0.0942)	(0.0515)	(0.0326)	(0.159)	(0.105)
Confidence in Court	0.0406	0.184	0.0183	0.0610	-0.0445	0.0226
	(0.0469)	(0.138)	(0.0505)	(0.0600)	(0.0945)	(0.150)
Constant	-0.344**	-0.955***	-0.241*	-0.430**	-0.0329	-0.253
	(0.140)	(0.250)	(0.143)	(0.202)	(0.197)	(0.212)
Observations	14,255	2,340	11,915	5,533	4,433	1,943
R-squared	0.046	0.124	0.040	0.086	0.033	0.054

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6.2 Figures

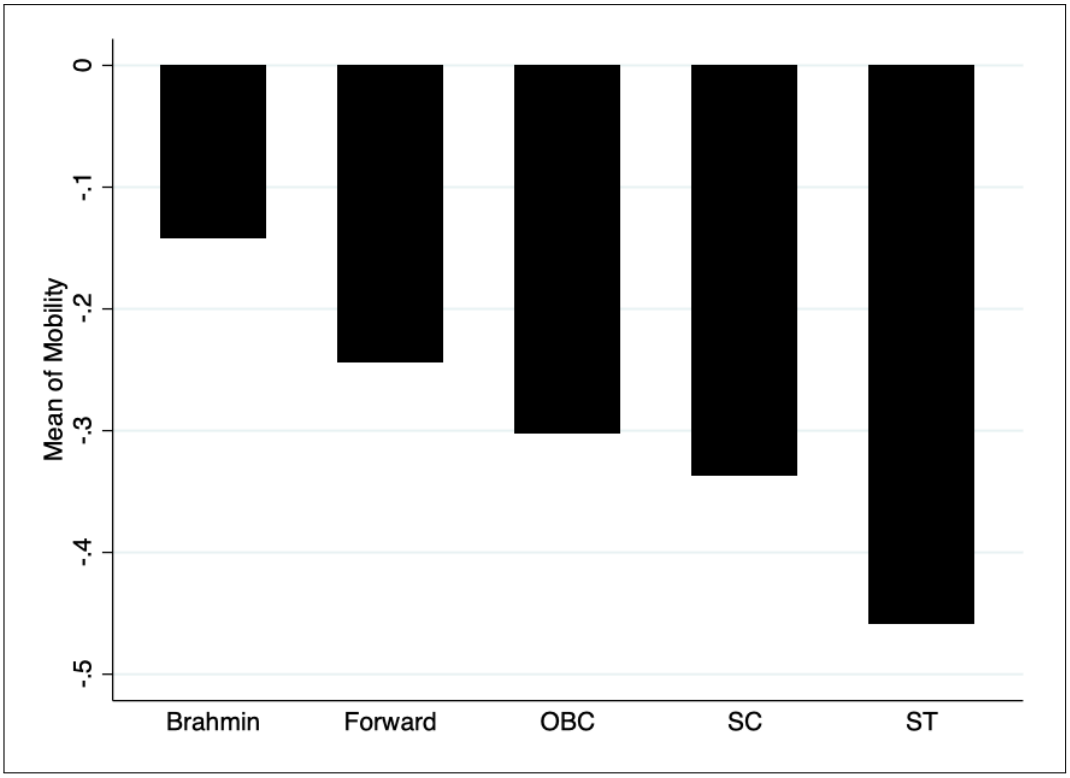

Figure 1: Mean Mobility across Caste

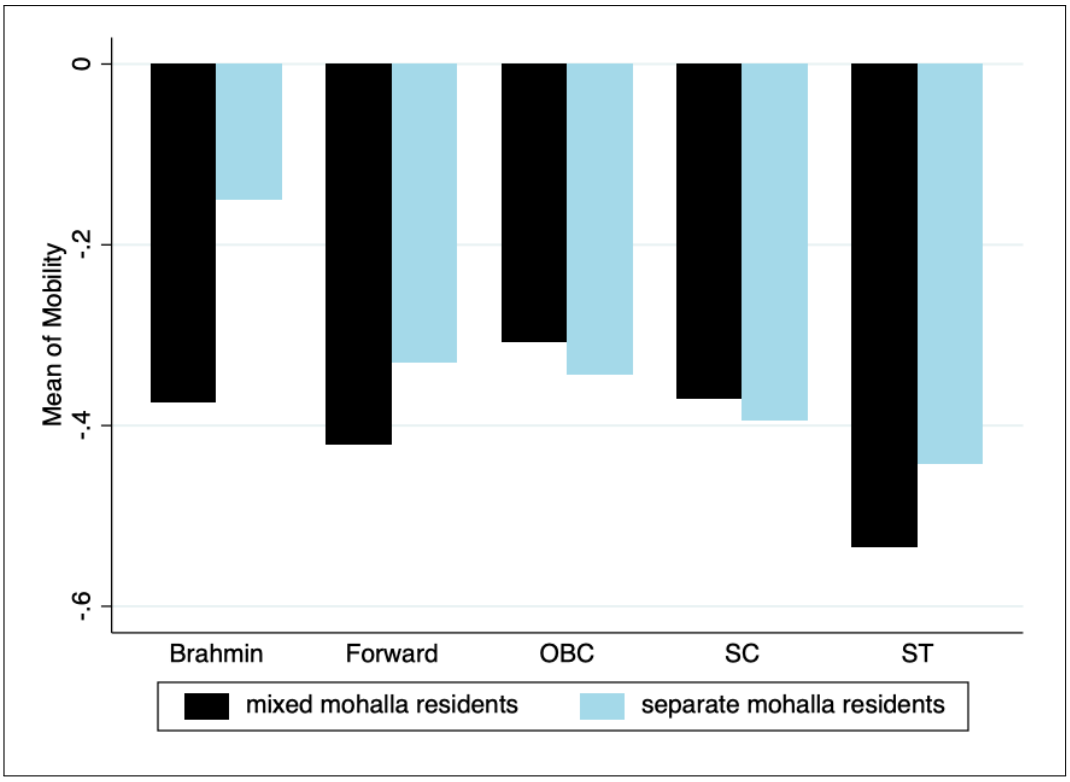

Figure 2: Mean Mobility for Mixed and Separate Neighbourhood Residents

References

Adukia, A., Asher, S., Novosad, P., and Tan, B. (2019). Residential segregation in urban india. *Unpublished manuscript*.

Alesina, A., Baqir, R., and Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, 114(4):1243–1284.

Alesina, A. and Tabellini, M. (2022). The political effects of immigration: Culture or economics? Working Paper 30079, National Bureau of Economic Research.

Asher, S., Novosad, P., and Rafkin, C. (2018). Intergenerational mobility in india: Estimates from new methods and administrative data. *World Bank Working Paper*.

Austen-Smith, D. and Fryer, R. G. (2005). An economic analysis of "acting white". *The Quarterly Journal of Economics*, 120(2):551–583.

Azam, M. and Bhatt, V. (2015). Like father, like son? intergenerational educational mobility in india. *Demography*, 52(6):1929–1959.

Banerjee, A., Iyer, L., and Somanathan, R. (2005). History, social divisions, and public goods in rural india. *Journal of the European Economic Association*, 3(2-3):639–647.

Benabou, R. (1993). Workings of a city: location, education, and production. *The Quarterly Journal of Economics*, 108(3):619–652.

Benabou, R. (1996). Equity and efficiency in human capital investment: the local connection. *The Review of Economic Studies*, 63(2):237–264.

Bowles, S., Gintis, H., and Osborne, M. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature*, 39(4):1137–1176.

Chetty, R. and Hendren, N. (2018). The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates. *The Quarterly Journal of Economics*, 133(3):1163–1228.

Chetty, R., Hendren, N., and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902.

Chin, A. and Prakash, N. (2011). The redistributive effects of political reservation for minorities: Evidence from india. *Journal of development Economics*, 96(2):265–277.

Chyn, E. and Katz, L. F. (2021). Neighborhoods matter: Assessing the evidence for place effects. *Journal of Economic Perspectives*, 35(4):197–222.

Collier, P. and Gunning, J. W. (1999). Explaining african economic performance. *Journal of Economic Literature*, 37(1):64–111.

Fordham, S. and Ogbu, J. U. (1986). Black students' school success: Coping with the "burden of 'acting white'". *The Urban Review*, 18(3):176–206.

Galor, O. and Zeira, J. (1993). Income distribution and macroeconomics. *The review of economic studies*, 60(1):35–52.

Hnatkovska, V., Lahiri, A., and Paul, S. (2012). Castes and labor mobility. *American Economic Journal: Applied Economics*, 4(2):274–307.

Hnatkovska, V., Lahiri, A., and Paul, S. B. (2013). Breaking the caste barrier intergenerational mobility in india. *Journal of Human Resources*, 48(2):435–473.

Jaffrelot, C. (2013). Caste and political parties in india: Do indians vote their caste—while casting their vote? In Kohli, A. and Singh, P., editors, *Routledge Handbook of Indian Politics*, pages 107–118. Routledge, first edition.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542.

Mogstad, M. and Torsvik, G. (2021). Family background, neighborhoods and intergenerational mobility. Working Paper 28874, National Bureau of Economic Research.

Motiram, S. and Singh, A. (2012). How close does the apple fall to the tree? some evidence from india on intergenerational occupational mobility. *Economic and Political Weekly*, pages 56–65.

Munshi, K. and Rosenzweig, M. (2009). Why is mobility in india so low? social insurance, inequality, and growth. Working Paper 14850, National Bureau of Economic Research.

Muthukkaruppan, P. (2014). Dalit: the making of a political subject. *Critical Quarterly*, 56(3):34–45.

Posta, D. J. D. (2013). Competitive threat intergroup contact or both? immigration and the dynamics of front national voting in france. *Social Forces*, 92(1):249–273.

Reddy, A. B. (2015). Changes in intergenerational occupational mobility in india: Evidence from national sample surveys, 1983–2012. *World Development*, 76:329–343.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of political economy*, 64(5):416–424.