Caste and Infant Mortality in India

Jasleen Kaur

In this paper, I shed light on a puzzle in India. Schedule Caste children are less likely than Schedule Tribe children to survive their first birthday, even though Schedule Castes have higher wealth, educational attainment, and access to state services than the Schedule Tribes. This highlights a very critical inequality puzzle with far-reaching policy implications.

I. Introduction

In this paper, I put forward an inequality puzzle in the health and the development literature. Schedule Tribe children have lower wealth, educational attainment, and face worse access to state services such as piped water and health infrastructure, compared to the majority Schedule Caste children. However, 6 out of 1000 more schedule caste children as compared to schedule tribe children are likely to die before their first birthday. 4 out of the 6 of 1000 schedule caste children are likely to die before their 1st birthday month. I describe this puzzle and then try to understand the causal mechanism leading this difference in infant and neonatal mortality rates between the Schedule castes and Schedule tribes in India.

India is a diverse country with different castes and cultures. There are 4 main categories into which the Indian community is divided – General castes, Other backward class, Schedule caste and Schedule tribes. General castes are the social and economically advantaged as compared to the other caste groups. Schedule Castes and Schedule tribes are the officially designated groups of historically disadvantaged people in India. Schedule castes (“dalits”) rank lowest in the Hindu society and are often called the untouchables. They often face discrimination from other castes. Schedule Tribes are the economically most disadvantaged population of the country. They are the indigenous tribe (called the “adivasis”). The schedule tribes have a social identity that is outside the caste system. They often live in isolated villages that are not well served by state services. Schedule caste and schedule tribes form about 16.6% and 8.6% respectively, of India’s population (according to 2011 census).

II. Literature

India has a high Infant mortality rate and neonatal mortality rate. With 22.7 out of 1000 deaths before the first month birthday (UNICEF, 2018), India is one of worst ranking countries in terms of child survival rates. Bassani et al. (2010) using a nationally representative survey of 1.1 million households find the main causes of neonatal and child mortality in India. The three causes that account for 78% (0.79 M/1.01 M) of all neonatal deaths in India are prematurity & low birthweight (0.33 M; 99%CI 0.31-0.35 M), neonatal infections (0.27 M; 99%CI 0.25-0.29 M) and birth asphyxia & birth trauma (0.19 M; 99%CI 0.18-0.21 M). Two causes that account for 50% (0.67 M/1.34 M) of all deaths at ages 1-59 months: pneumonia (0.37 M; 99%CI 0.35-0.39 M) and diarrhoeal diseases (0.30 M; 99%CI 0.28-0.32 M). Kishor (1993), Murthi et al. (1995), Claeson et al. (2000), Pande (2003), Bassani et al. (2010) study the difference in child mortality rate by gender. Bhalotra et al. (2008) study the difference in infant mortality rates by religion. Geruso and Spears (2018) in their seminal paper find that Hindu children are more likely than Muslim children to die before their first birthday because of the practice of open defecation among Hindus. Very few studies look at the differences in infant mortality by caste status in India. Dommaraju et al.

(2008) use NFHS-2 to indicate that net of individual-level and community-level controls, children belonging to low castes have higher risks of death. Bora et al. (2019) show that 78% of the caste-based gap in Under 5 Mortality is due to the effect of women's level of educational attainment and household wealth between the SC/ST and non-SC/ST population. Ram et al. (2016), using the NFHS-2 and NFHS-3 address that mortality rates among STs are lower than SCs, but stop at that and do not find why that could be.

This paper is the first to use the recent NFHS-4 data to address the puzzle that despite higher wealth and education among SCs, they are more likely to die before their first birthday as compared to ST children. This paper is also the first to try and find our causal mechanism behind this difference. I follow Geruso and Spears (2018) methodologically for this paper.

III. The puzzle of SC mortality disadvantage

A. Data

For my main analysis, I use the data from the recent round of the National Family Health Survey (NFHS) of India: 2015-2016. The NFHS (India's version of the Demographic and Health Survey) is a large, nationally representative survey that collects data from women aged 15 to 49. NFHS-4 sample size is expected to be approximately 568,200 households. This is expected to yield a total sample of 625,014 women and 93,065 men eligible for the interview. Respondents report birth histories, including deaths and stillbirths, from which I calculate infant and neonatal mortality rates. The NFHS also includes information on household assets, household physical infrastructure, and health behaviors. The NFHS-4 sample is a stratified two-stage sample. The 2011 census served as the sampling frame for the selection of Primary Sampling Units (PSUs). PSUs were villages in rural areas and Census Enumeration Blocks (CEBs) in urban areas. PSUs with fewer than 40 households were linked to the nearest PSU. Within each rural stratum, villages were selected from the sampling frame with probability proportional to size (PPS). In each stratum, six approximately equal substrata were created by crossing three substrata, each created based on the estimated number of households in each village, with two substrata, each created based on the percentage of the population belonging to scheduled castes and scheduled tribes (SCs/STs). Within each explicit sampling stratum, PSUs were sorted according to the literacy rate of women age 6+ years. The final sample PSUs were selected with PPS sampling.

Table 1 tabulates the summary statistics for our main analysis sample, which consists of Schedule Castes and Schedule Tribes in the 4th wave of the NFHS. My primary outcome of interest is the infant mortality rate (IMR) and the neonatal mortality rate (NMR), defined respectively as the number of deaths among children less than one year old and less than one month old, scaled per 1,000 live births. Throughout the paper, I organize our analysis at the level of the child, constructing mortality rates from birth history information on around 54000 Schedule Caste and Schedule Tribe children in India in the most recent round.

Table1 - Summary Statistics: SCs and STs in NFHS -4

	SC Subsample		ST Subsample	
	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)
Infant Mortality (IMR), year 1	57.094	0.77	51.357	0.929
Neonatal Mortality (NMR), month 1	39.011	0.614	35.018	0.74
Household Open Defecation	0.535	0.005	0.546	0.007
Local (PSU) open defecation	0.534	0.005	0.547	0.007
Household has electricity	0.977	0.005	0.95	0.007
Household has piped water	0.135	0.003	0.102	0.004
Household is urban	0.255	0.007	0.18	0.007
Household has radio	0.225	0.005	0.197	0.006
Household has TV	0.748	0.006	0.6	0.008
Household has refrigerator	0.345	0.006	0.262	0.007
Household has bicycle	0.755	0.006	0.649	0.007
Household has motorcycle	0.436	0.006	0.381	0.007
Household has car	0.179	0.005	0.16	0.006
Mother's height	157.715	10.258	157.97	6.312
Mother no education	0.532	0.004	0.55	0.005
Mother completed primary	0.389	0.003	0.353	0.004
Child's birth order	2.316	0.007	2.289	0.009
Child's female	0.476	0.001	0.479	0.002
Observations (live births)	246513		298428	

Notes: The table displays summary statistics for my main analysis. Neonatal and Infant mortality are defined, respectively, as the number of deaths among children less than one month old and less than one year old, scaled per 1,000 live births. Observations are children (live births).

Table 1 shows that Infant and neonatal mortality is 11% higher for Schedule Caste as compared to Schedule Tribes. Neonatal Mortality shows a similar pattern, with a 11% survival deficit. This is despite Schedule Castes having higher educational attainment and wealth, measured in survey by assets.

B. The Puzzle

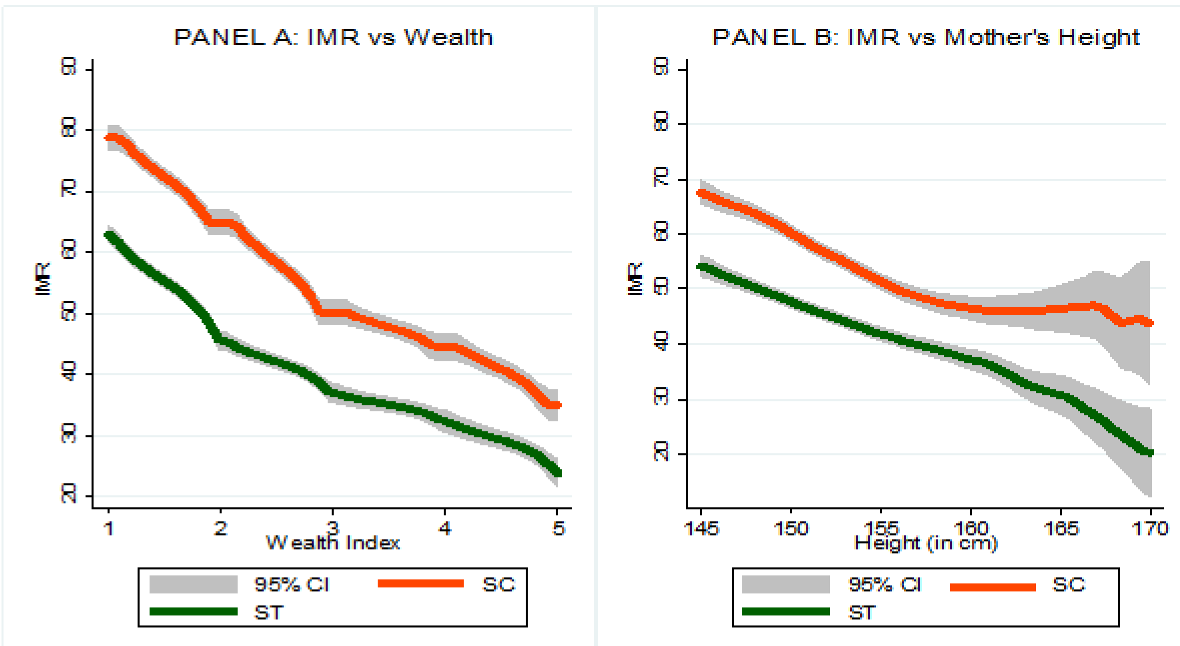

FIGURE 1. PUZZLE: AT ALL LEVELS OF PARENTAL WEALTH AND HEALTH, SC IMR IS HIGHER

Notes: The figure plots local regressions of infant mortality on measures of economic well-being. The dependent variable is an indicator for death in the first year of life × 1,000. Panel A plots mortality against asset wealth rank, constructed as described in the text. Panel B plots mortality against mother's height. Observations are children (live births). The grey area in the figure correspond to 95 percent confidence intervals that are not adjusted for clustering.

Figure 1 demonstrates the mortality puzzle. At all levels of socioeconomic status, mortality is higher among Schedule Castes than Schedule Tribes. The figure plots infant mortality, separately by the two groups, against two alternative summary measures of household economic well-being. The NFHS, like all DHS surveys, does not measure income or consumption. Therefore, in panel A of Figure 1, I follow the literature (see, for example, Filmer and Pritchett 2001) in using asset ownership as a proxy for wealth. I use the wealth index that is measured by the DHS itself as a proxy for assets in the household. This gives the horizontal axis clear rank interpretation. As an alternative measure of parental endowment, I use mother's height along the horizontal axis in panel B of Figure 1. Maternal adult height predicts maternal adult health and reflects maternal economic well-being earlier in the mother's life (Case and Paxson 2008, Steckel 2009). Consistent with asset ownership and mother's height capturing meaningful variation in endowments that is correlated with child survival, Figure 1 shows that infant mortality is steeply decreasing in both measures. The SC disadvantage is apparent in the large and statistically significant mortality differences at any fixed level of either measure of well-being.

C. Empirical Strategy

I run the following linear regression to try and understand patterns in the society.

$$Mortality_{ij} = \alpha + \beta Caste_{ij} + \mu_j + f(X_i) + \varepsilon_{ij} \quad (1)$$

where i indexes live births, j indexes PSUs. $Mortality_{is}$ is an individual-level mortality indicator for infant and neonatal mortality. It is scaled such that coefficients reflect deaths per 1000. $Caste_{ij}$ is the regressor of interest. It takes the value 0 if the mother of the child belongs to Schedule tribe and 1 if the mother belongs to schedule caste. μ_j is the PSU (the main model has PSU fixed effects). X_i are controls for a set of demographic and socioeconomic characteristics that determine early-life health in India according to literature. They include mother's height, wealth index, education level of mother, gender of child (female = 1 if female, 0 otherwise), religion of child, birth order, whether the household has piped water (1 if yes, 0 if no). I choose these covariates because considerable evidence suggests that variables capturing economic status and education predict child development outcomes, likely in part because they are correlated with the quality and quantity of food and other health inputs that young children need to grow (Case et al. 2002; Desai and Alva 1998). I also control for high caste fraction. Coffey et al. (2019) find that since SCs live in the vicinity of higher caste people, they are social rank orders are enforced against them, making them feel discriminated against and hence causing a lower child height. However, since STs stay in regions mostly secluded from other castes, they are not discriminated against, despite being the poorest caste in India. Robust standard errors are clustered at the PSU level.

I also run treatment by covariate interactions. However, the results do not change – instead show the same story.

D. Results

Table 2: Regression Output

Dependent variable	IMR	IMR	IMR	IMR	NMR	NMR	NMR	NMR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Schedule Caste	5.737***	6.243***	1.210	1.264	3.994***	4.126***	1.184	1.470
	(1.19)	(1.28)	(1.99)	(2.03)	(0.95)	(1.04)	(1.64)	(1.70)
Mom's Height	No	-1.284***	No	-1.096***	No	-1.022***	No	-0.878***
		(0.09)		(0.10)		(0.08)		(0.08)
Wealth Index	No	-6.674***	No	-5.257***	No	-4.403***	No	-3.127***
		(0.55)		(0.71)		(0.46)		(0.58)
Education Level	No	-1.695***	No	-1.308***	No	-1.212***	No	-1.090***
		(0.14)		(0.16)		(0.11)		(0.13)
Female	No	-9.606***	No	-9.500***	No	-10.208***	No	-10.183***
		(0.94)		(0.98)		(0.77)		(0.80)
Piped Water	No	-4.472**	No	-2.544	No	-3.687**	No	-1.190
		(1.68)		(2.24)		(1.37)		(1.76)
High Caste Fraction	No	4.356*	No	0.000	No	4.643**	No	0.000
		(2.02)		(.)		(1.64)		(.)
Birth Order Dummy	No	Yes	No	Yes	No	Yes	No	Yes
Religion Dummy	No	Yes	No	Yes	No	Yes	No	Yes
PSU fixed Effects	No	No	Yes	Yes	No		Yes	Yes
Observations	524938	524938	524938	524938	520406	520406	520406	520406

"* p<0.05 ** p<0.01 *** p<0.001"

Notes: The table reports results from OLS regressions. The dependent variable in columns 1 through 4 is infant mortality (year1). The dependent variable in columns 5 through 8 is neonatal mortality (month1). Mortality variables are scaled as described in the text to generate coefficients that indicate impacts on rates × 1,000 (deaths per 1,000) children. Columns 1 and

5 do not control for SES variables or PSU fixed effects. Columns 2 and 6 control for SES variables, The SES variables are a set of controls including mother's height, mother's educational attainment, household wealth, birth order, an indicator for child's sex, an indicator for piped water and high caste fraction (Coffey et al., 2019). Columns 3 and 7 only have PSU fixed effects. Columns 4 and 8 control for SES variables and I add PSU fixed effects. Observations are children (live births). Probability weights are given to regression. Standard errors are clustered at the PSU level.

In Table 2, Model 1 (the naïve estimate) tells us that on average 6 of 1000 more Schedule Caste children are likely to die as compared to Schedule Tribe children before their first birthday, of which four SC children would die before the 1st month of their birth. These results are significant at the 95% Confidence Interval. When I add Socioeconomic status controls in column 2 and 6, I see that more Schedule Caste children are likely to die before their 1st year and first month of birth. These results are also significant. However, we should be careful to interpret that these columns only reiterate the puzzle, stating that Schedule Castes fare better in comparison to Schedule tribes when it comes Socio-economic status. In Columns 3 and columns 7, I add PSU fixed effects. Model 3 and Model 7 tell us that adding PSU fixed effects and urban/rural goodness, SCs are more likely to stay in areas that are worse for infant and neonatal mortality as compared to the Schedule tribes. In the main Models - Model 4 and Model 8, I add socio-economic-demographic covariates and PSU fixed effects. These models tell me that comparing live births in the same PSUs and matched on observably similar socio-economic status, SC children are more likely to die than ST children. However, the results are insignificant at the 95% confidence interval. The main models – Model 4 and Model 8 show that when we add PSU fixed effects, being a Schedule Caste or Schedule Tribe does not matter for Infant and Neonatal Mortality. What matters is where one stays. Controlling for PSU fixed effects, with or without the SES controls, renders the coefficient on caste insignificant.

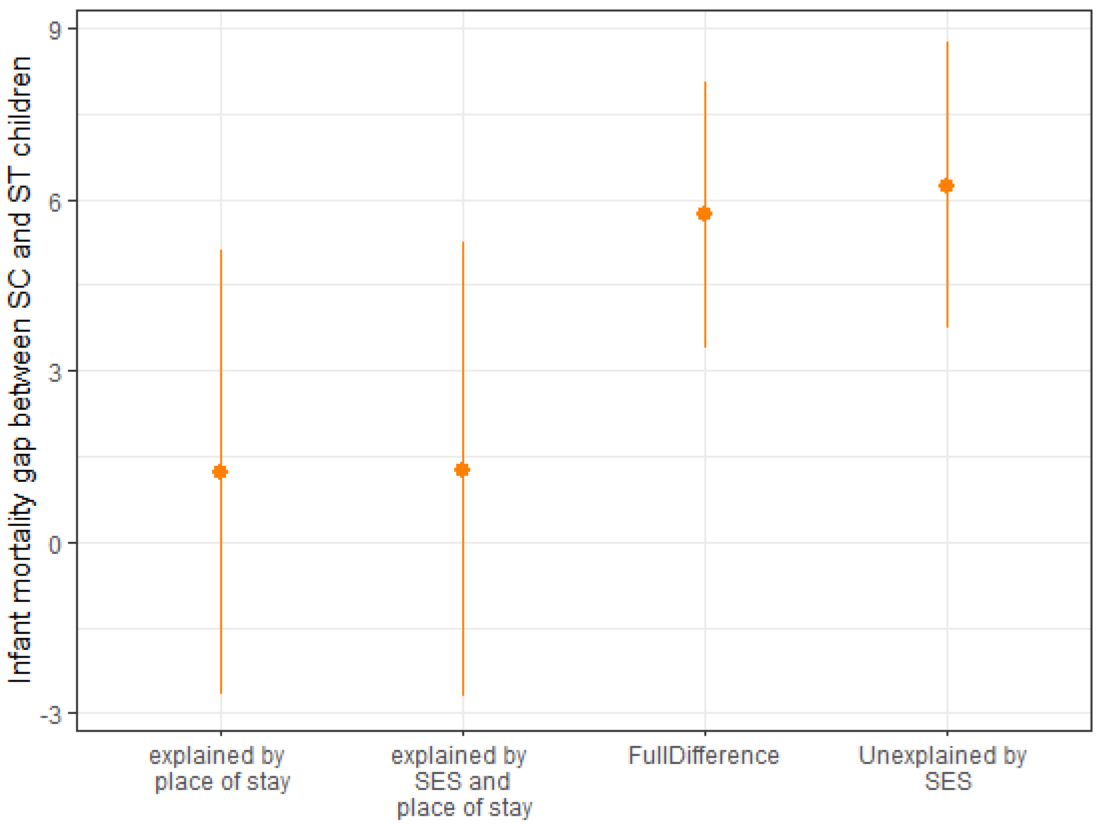

FIGURE 2: PLACE OF STAY CAN EXPLAIN THE PUZZLE

Figure 2 graphically depicts the regression output in table. One thing to notice is that the Infant mortality puzzle could be tried and explained depending on “place of stay” as our confidence intervals overlap with zero. That provides me the option of coming up with alternative hypothesis that could probably explain the anomaly.

IV. Conclusion/Scope for further research

I have laid out the puzzle in the paper stating that even though Schedule Castes in India are better off that Schedule Tribes in terms of health, wealth and educational status, the schedule caste children are more likely to die than schedule tribe children before their first birthday. This is puzzle because all main determinants in the health and demography literature that cause mortality, seem to be working in the opposite direction. I specifically look at two confounding variables from the literature – Open Defecation and High Caste Fraction that might lead to increased infant and neonatal mortality rates among SCs. I control for "Open defecation" based on a recent paper by Geruso and Spears (2018). SC population in India is mostly Hindu, while STs belong to either the Muslim or Christian religion. Geruso and Spears (2018) find that Hindus have a higher infant mortality rate than Muslims despite having a higher wealth, health and education because of the Hindu practice of defecating in the open. This practice leads to increased rate of infection among young rural children, causing higher deaths. However, when I control for open defecation, I do not find that open defecation is the cause behind the SC/ST mortality puzzle.

When I add PSU fixed effects, I see that it is where SCs live is what matters for infant and neonatal mortality. Trying to explore as to what about "where SCs/STs stay" could lead to this puzzle, I graph a map as to the geographical location of where STs and SCs stay. Figure 3 shows that while SCs are spread throughout the country, STs are clustered together in certain parts of the country and almost non-existent in the others.

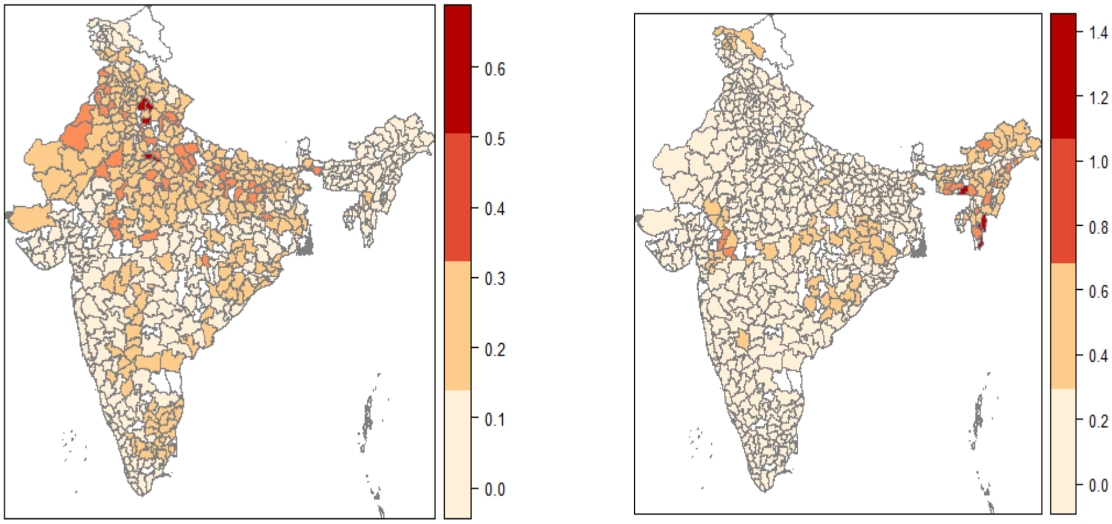

FIGURE 3: The left map shows the proportion of SC population of the total ST and SC population. The right map shows the proportion of ST populations of the total ST and SC population. The scales on the two maps are different because STs are less in number as compared to SCs. The figure shows that while SCs are spread throughout the country, STs are clustered together in certain parts of the country and almost non-existent in the others.

Then based on a recent paper, I control for high caste fraction – controlling for "whether higher caste live in surrounding areas", causing SCs to be discriminated against and hence causing higher infant mortality among SCs. I find that clearly, SCs tend to stay in regions with higher caste groups and STs do not.

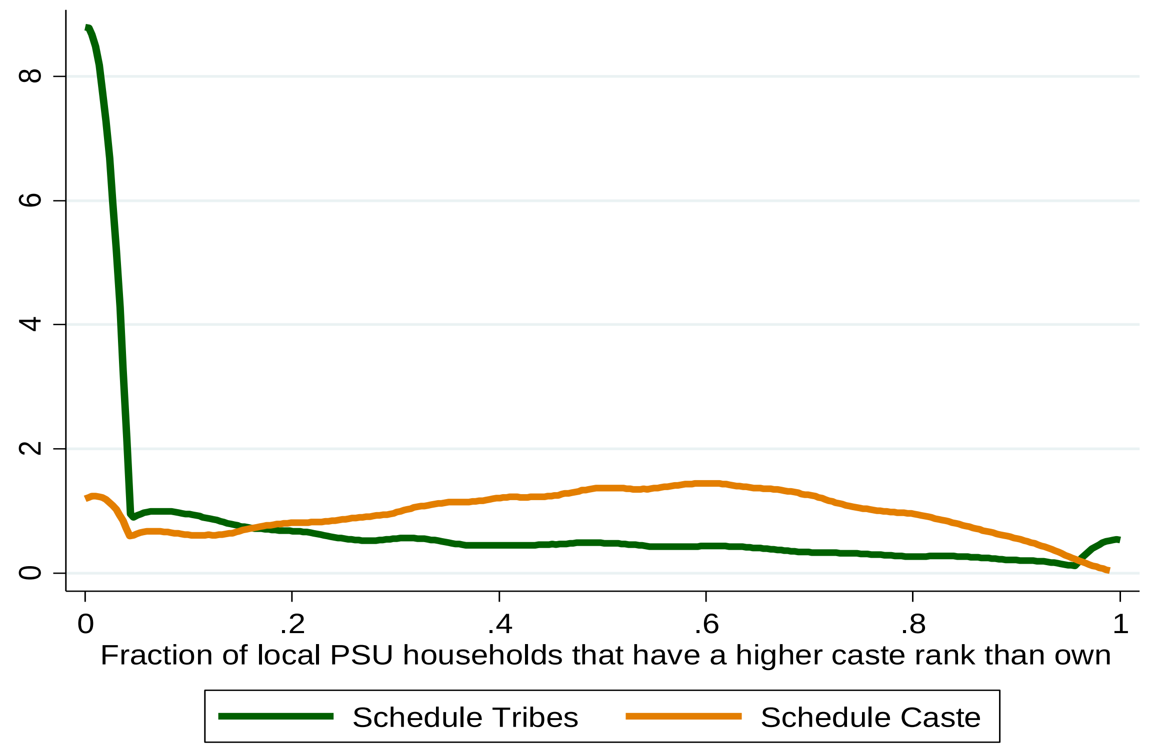

FIGURE 4: SCs TEND TO STAY IN REGIONS WITH HIGHER CASTE GROUPS COMPARED TO STs.

However, when I control for it in my regression (methodologically similar to Coffey et al. 2019), I do not find that "high caste fraction" (proxy for discrimination) is the causal mechanism behind the puzzle.

I also find that overall, these regions where STs stay are poorer regions with less access to health facilities. Therefore, the question remains - What about the PSUs leads to this difference? This is an extremely important question for policy makers. Until we do not know the cause, specific to where SCs or STs stay that leads to higher infant and neo natal mortality among the SCs, we will not be able to make policies to reduce the mortality rate. Health of a child is extremely important not just via the instrumentalist approach – that we need healthy children for economic development, but also for the equity argument. We need healthy children because they deserve to be healthy. To be healthy, we need to know what causes mortality and research in this area can lead to important health policy to reduce infant and neonatal mortality in India.

References

1. Ram, Bali, Abhishek Singh, and Awdhesh Yadav. "The persistent caste divide in India's infant mortality: A study of Dalits (ex-untouchables), Adivasis (indigenous peoples), Other Backward Classes, and forward castes." Canadian Studies in Population [ARCHIVES] 43, no. 3-4 (2016): 249-63.
2. Bora, Jayanta Kumar, Rajesh Raushan, and Wolfgang Lutz. "The persistent influence of caste on under-five mortality: Factors that explain the caste-based gap in high focus Indian states." PloS one 14, no. 8 (2019): e0211086.
3. Case, Anne, and Christina Paxson. "Stature and status: Height, ability, and labor market outcomes." Journal of political Economy 116, no. 3 (2008): 499-532.
4. Coffey, Diane, Ashwini Deshpande, Jeffrey Hammer, and Dean Spears. "The Impact of Poverty and Discrimination on Child Height in India." PRC Research & Policy Brief Series (2019).
5. Dommaraju, Premchand, Victor Agadjanian, and Scott Yabiku. "The pervasive and persistent influence of caste on child mortality in India." Population Research and Policy Review 27, no. 4 (2008): 477-495.
6. Filmer, Deon, and Lant H. Pritchett. "Estimating Wealth Effects Without Expenditure Data—or Tears: An Application to Educational Enrollments in States of India." Demography 38, no. 1 (2001): 115–32.
7. Geruso, Michael, and Dean Spears. "Neighborhood sanitation and infant mortality." American Economic Journal: Applied Economics 10, no. 2 (2018): 125-62.
8. Greenstone, Michael, and Rema Hanna. "Environmental regulations, air and water pollution, and infant mortality in India." American Economic Review 104, no. 10 (2014): 3038-72.

9. Kishor, Sunita. "May God give sons to all": gender and child mortality in India." American sociological review (1993): 247-265.
10. Million Death Study Collaborators. "Causes of neonatal and child mortality in India: a nationally representative mortality survey." The Lancet 376, no. 9755 (2010): 1853-1860.
11. Murthi, Mamta, Anne-Catherine Guio, and Jean Dreze. "Mortality, fertility, and gender bias in India: A district-level analysis." Population and development review (1995): 745-782.
12. Pande, Rohini P. "Selective gender differences in childhood nutrition and immunization in rural India: the role of siblings." Demography 40, no. 3 (2003): 395-418.
13. Steckel, Richard H. "Heights and human welfare: Recent developments and new directions." Explorations in Economic History 46, no. 1 (2009): 1-23.