Caste Differences in Child Growth: Disentangling Endowment and Investment Effects

Joseph Cummins* Jingyan Guo* Neha Agarwal† Anaka Aiyar‡
Vaishali Jain* Andrew Bergmann§

May 26, 2023

Abstract

Using the fourth round of the Indian National Family Health Survey (NFHS-4), and subsequently replicating our results using the fifth round (NFHS-5), we document differential child physical growth patterns across caste groups in India, demonstrating that lower caste children are born shorter and grow less quickly than children from higher-caste households. We then show that, in line with work from previous rounds of the NFHS, these differences are largely explainable by observable covariates, particularly maternal characteristics and household wealth variables. Our research also reveals a previously un-documented dynamic, that the influence of these variables changes as children develop, and suggests that caste-gaps are the result of multiple mechanisms impacting the child growth process at different stages of development. Using age-disaggregated decomposition methods, we demonstrate that health endowment related variables largely explain birth length gaps, and that investment related variables become increasingly influential as children age. Children from lower caste households thus face two margins generating height gaps as they age: a persistent endowment disparity present from birth, and a post birth investment differential that exacerbates the initial deficit.

JEL classification: I14, I15, O10, O15, O53

Keywords: Height for Age Z scores, Child health, Health Disparities, Health production, India, Caste, National Family Health Surveys

*University of California, Riverside
†University of Otago
‡University of Nevada, Reno
§Ohio State University

1 Introduction

We document large disparities in child height for age z-scores (HAZ) across Indian caste groups using data from the fourth Indian National Family Health Survey (NFHS-4) and replicate the results using the fifth round (NFHS-5). Scheduled Caste (SC) and Scheduled Tribe (ST) children are, on average, around 0.4 and 0.5 standard deviations (sd) shorter than Upper Caste (UC) children in the first six-months of life, while children from Other Backwards Classes (OBC) are about 0.2sd shorter. By age five, caste HAZ differentials have increased by an additional 0.1-0.2sd in each group.

Our results are fully consistent with previous research on the topic. What separates our work from previous analyses of child HAZ gaps across caste groups is our focus on the biological process of human growth. While both Van de Poel and Speybroeck (2009) and Coffey et al. (2019) are concerned with the social processes generating caste HAZ disparities, we are interested in the inter-generational and contemporary disparities in health inputs that generate them. Towards that end, we provide evidence that not only are the caste gaps in HAZ explainable at each age, but the extent to which different variables explain the gaps also change. Our results are easily interpretable within a health capital accumulation framework. Variables affecting child health endowments explain birth length differences and have persistent effects across childhood. Variables affecting health investments in children, post-birth, do not explain birth length differences, but grow increasingly influential on HAZ disparities as children grow.

It is, of course, well understood that early life health is determined by factors that affect child health both before and after children are born (Danaei et al., 2016; De Onis and Branca, 2016; Prendergast and Humphrey, 2014). At a demographic level, Victora et al. (2010) documents overarching patterns in HAZ-differentials across socioeconomic groups, but does not quantify the dynamic explanatory power of child health determinants. Rieger and Trommlerová (2016) and Alderman and Headey (2018) document changes in the influence of various socioeconomic variables as children age but do not focus directly on contexts of social discrimination. The former compares non-parametric HAZ-age profiles across socioeconomic categories (wealth quintile and mother's education etc.), and the latter estimates the coefficients on important household variables separately for children under- and over-2 years of age. Neither provide scaled estimates of the relative explanatory power of the various inputs across age, a key component in quantifying and understanding the dynamic effects of child health inputs that are the biological root causes of child HAZ disparities. In relation to the above work our results can be understood as quantifying the theorized dynamics of child development in a context of explaining disparities: lower caste children are born with measurable health deficits relative to higher caste children, and these deficits are exacerbated by measurable deficits in post-birth health inputs as children age. In that sense our results reflect the biological nature of child growth faltering, showcasing the dynamics of child HAZ caste gaps.

While some of the dynamics we describe have been explored in the broader literature on growth faltering, there is surprisingly little evidence documenting the age dynamics of childhood HAZ disparities, in India or other countries. Moreover, we find no previous work in the economic demography literature examining how the changing influence of household variables relates to understanding caste HAZ disparities. Previous analyses of child caste HAZ disparities (Van de Poel and Speybroeck, 2009; Coffey et al., 2019) have focused instead on understanding the discriminatory practices that generate disparities in child HAZ. Van de Poel and Speybroeck (2009) attempts to quantify the role of unobserved, contemporaneous discrimination in caste HAZ gaps, while Coffey et al. (2019) relies on an economic interpretation of social hierarchy to investigate and interpret regional heterogeneity in caste HAZ gaps.

We focus on an economic interpretation of the biological processes underlying these caste HAZ disparities, locating the appearance and describing the dynamics of these

HAZ gaps in the child development process itself. The goal of this work is to provide researchers and policy makers with economic insights into the biological (in contrast to social) processes behind the disparate and diverging child growth trajectories across castes. Our perspective shifts the policy debate away from the social determinants of caste discrimination and towards the biological effects of such discrimination. Our results speak to the timing and subject of any such policies or interventions aimed at reducing caste disparities in child health.

Our results thus reinforce an important overarching insight for both researchers and policy makers - understanding or ameliorating caste disparities in child HAZ will require focus on both contemporary disparities in wealth and access to health services, as well as inter-generational disparities in parental health. While such arguments have appeared before in terms of understanding global child stunting broadly (Danaei et al., 2016; De Onis and Branca, 2016; Prendergast and Humphrey, 2014), our analysis shows that multiple channels (parental, child, and pre- and post-birth) matter for understanding differential growth faltering in addition to social discrimination within countries as well.

1.1 Overview of Results

Figure 1 summarizes our main empirical focus and findings, graphing mean HAZ across child age in months for each caste group. The overall pattern across all groups is similar and mirrors the HAZ-age profile shape common across the developing world (Victora et al., 2010). Indian children are, on average, born below the WHO reference population by between 0.2sd and 0.5sd. Over the first two years of life, Indian children then grow too slowly compared to the reference population, reflected in a decreasing mean HAZ over this period.

The caste differentials themselves display similar age-dynamics. SC, ST and OBC children are, on average, born shorter than UC children. These gaps then widen over

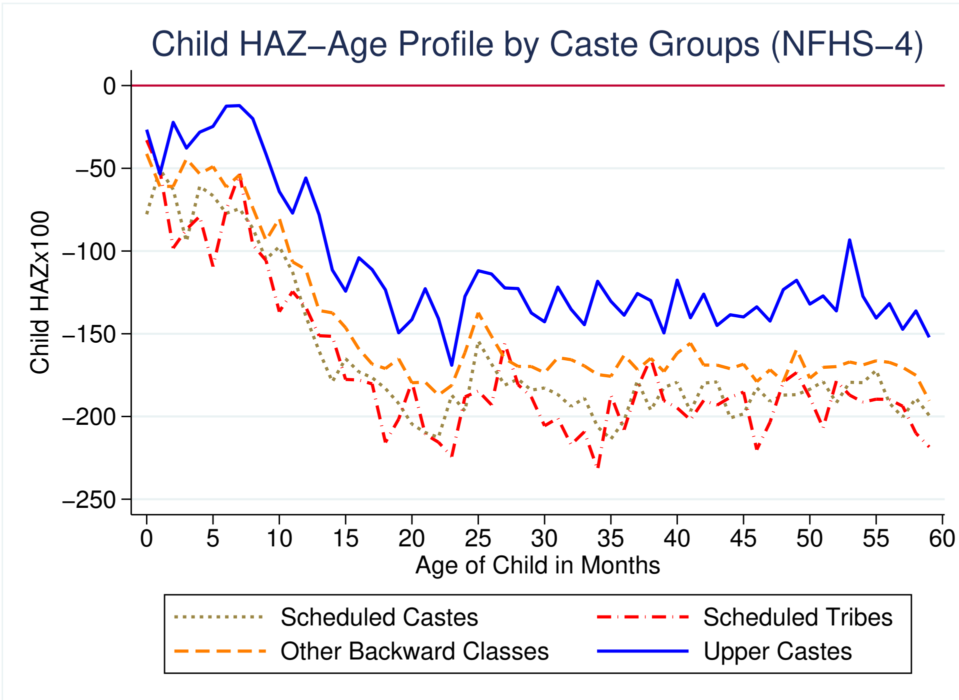

Figure 1: Child HAZ by Caste Groups

This figure graphs mean child HAZ score (x100) by caste groups for children from 0 to 5 years. The x-axis represents child age in months and the y-axis is mean weighted HAZ(x100). Median child HAZ for reference population children is 0. The mean child HAZ in this sample is -1.49. A child with HAZ between -2 and -3 indicates moderate chronic malnutrition (stunted), and HAZ below -3 indicates severe chronic malnutrition. The results are weighted by sample weights.

the first two years of life, and subsequently remain relatively constant or slightly decrease between the ages of 2 and 5.

Caste differentials also largely disappear, or at least greatly attenuate in magnitude, after adjusting for a broad set of household and community variables in a regression model, a finding common in studies analyzing data from previous rounds of the NFHS as well (Van de Poel and Speybroeck, 2009; Coffey et al., 2019). Conditional on a rich set of covariates, point estimates become, for the most part, statistically insignificant at every age, and the few remaining age-caste group coefficients that remain statistically significant are all below 0.2sd in size.

We interpret these results from a theoretical perspective focused on these dynamics of child growth and grounded in a dynamic health capital accumulation interpretation of

HAZ (Grossman, 1972; Becker, 1962). Health capital theory considers the realized health of a person at any age as the result of two distinct mechanisms: a health endowment provided to a child at birth, and a stream of subsequent health inputs consumed by the child after birth. Following Aiyar and Cummins (2021), we interpret HAZ at birth as a measure of child health endowments, and the rate of loss of HAZ as a measure of the interaction between this health endowment and the subsequent stream of health inputs consumed by the child.

We then investigate the extent to which these caste HAZ gaps can be explained by observable variables related to the child's health endowment (e.g., maternal health variables) and/or to post-birth inputs in the form of private investments (e.g., household wealth) and public goods (e.g., sanitation), and how the relative influence of variables change as children age. The health capital theory prediction is clear: endowment-related variables should predominantly explain birth length, and investment related variables should become increasingly influential as children age.

Our decomposition results suggest precisely this: an initial HAZ-deficit at birth that is statistically explained largely by differences in variables related to the health endowment; and an effect of health-investment related variables that is small at birth and grows in relative importance as children age. By the age of 5, we estimate that around half of the caste-HAZ deficit is due to differences in health endowments, and around half due to differences in post-birth child health investments. We conclude that initial endowment deficits across caste persist through early childhood, while investment effects accumulate over the first few years of life.

Our paper proceeds as follows. In Section 2 we provide an overview of economic and health disparities across caste groups in India. In Section 3 we discuss the economic and biological underpinnings of the health capital framework to explain how disparities in health endowments and post-birth health inputs impact a child's health at any given point in their lives. We use these insights to develop explanations on why caste disparities may come to exist, accumulate and persist over time. In Section 4 we provide a discussion of the data used in our analysis and in Section 5 we propose and discuss empirical methods we employ to investigate the hypotheses developed in Section 2. Section 6 presents main results using the NFHS-4 and Section 7 replicates our methods and results using the NFHS-5. We conclude our analysis in Section 8 with a brief discussion on the implications of our research for understanding the age-dynamics of the determinants of health disparities across socioeconomic groups.

2 Caste Disparities

Caste disparities in India exist across almost every meaningful human welfare measure including household earnings, educational attainment and life-course health outcomes (Deshpande, 2000, 2001; Borooah, 2005; Kijima, 2006; Munshi and Rosenzweig, 2006; Subramanian et al., 2006; Deshpande, 2007; Deshpande and Newman, 2007; Asher et al., 2018; Van de Poel and Speybroeck, 2009; Perkins et al., 2011; Zacharias and Vakulabharanam, 2011; Kumar, 2013; Deshpande and Sharma, 2016; Maity, 2017; Deshpande and Ramachandran, 2019; LoPalo et al., 2019; Munshi, 2019; Blunch and Gupta, 2020; Vyas et al., 2022; Goraya, 2023). Health outcome disparities across caste are well documented across the lifespan: lower caste children are more likely to experience stunted growth. Lower caste men and women are also shorter and less healthy, and people from lower castes die younger (Kijima, 2006; Subramanian et al., 2006; Van de Poel and Speybroeck, 2009; Perkins et al., 2011; Maity, 2017; Vyas et al., 2022). The literature on these caste disparities in health outcomes has largely focused on quantifying the extent of current discrimination against lower caste members, and tracing out the mechanisms through which the long history of caste discrimination has translated religious social hierarchies into political and economic hierarchies that generate large health disparities (Coffey et al., 2019; LoPalo et al., 2019; Blunch and Gupta, 2020; Ramachandran and Deshpande, 2021).

Both sets of discriminatory practices - contemporaneous and historical - are potential explanations of child HAZ disparities across castes that we present. Caste-based occupational sorting that arises due to closely-knit caste networks leads to exclusion of SC and ST communities from lucrative livelihood opportunities (Deshpande, 2000, 2001; Ito, 2009; Siddique, 2011; Munshi, 2019), generating income disparities that directly decrease the ability of households to invest in their children's consumption. SC/ST and OBC groups are thus over-represented among the lower economic status groups. It is uncontroversial to argue that, in India at least, higher incomes correlate with higher consumption of nutrition and health inputs, and are thus correlated with subsequent child health outcomes directly. Furthermore, discriminatory practices lower access to health inputs for SC, ST, and OBC children both directly (Coffey et al., 2019; LoPalo et al., 2019; Spears and Thorat, 2019; Ramachandran and Deshpande, 2021) and indirectly (Debnath and Jain, 2020; Blunch and Gupta, 2020).

Additionally, lack of economic mobility reinforces health disparities across generations. Many studies have documented that poor maternal health leads to worse health outcomes for children (Addo et al., 2013; Aizer and Currie, 2014; Chakrabarti et al., 2021). Furthermore, maternal health is correlated with household wealth, which is strongly differential across castes. Together, the health-wealth correlation for parents generates a channel for the inter-generational transmission of poor health within lower caste households.

In the next section, we offer a biologically-focused economic perspective on the process through which health disparities come to exist across children in different caste groups. To do so, we model the biological process of child growth and the economic factors that influence this process, providing clarity on the potential mechanisms through which disparities in child health come to persist across castes in India.

3 Theoretical Framework

We approach the question of child growth differentials across caste from the perspective of health capital accumulation. In this section we sketch out the structure of health capital accumulation theory in order to define the roles of health endowments and investment streams in early life development (Grossman, 1972). We then discuss how HAZ is an especially appropriate measure of early life health capital and how the HAZ-age profile reflects the process of early life health capital accumulation in poor countries.

3.1 Health Capital

Households have preferences over the consumption and health outcomes of their members, and optimize an inter-temporal lifetime utility function representing those preferences subject to a budget constraint for their expenditures (Grossman, 1972; Becker, 1962). Households can purchase consumption and health investments at a cost that does not exceed the available budget in a period. We consider a household with a single newborn child.

Child health capital at birth is defined as H_0, a health endowment that is bestowed upon, and not chosen by, the newborn. In any given subsequent period, health at age 'a', H_a, is the result of a production function $f(H_{a-1}, I^*_{a-1})$ which takes as its inputs the health capital of the previous period (H_{a-1}) and optimally chosen health investments (I^*_{a-1}) in the previous period.

Households choose optimal consumption and health investment by maximizing their utility function subject to the budget constraints and the health capital production function. Choosing optimal health investment implies choosing an optimal health level, given health in the previous period (H_{a-1}). Furthermore, since health iterates from period to period beginning at birth, H_a can be expressed as $H^*_a(H_0, I^*)$, where I^* represents the stream of child health investments provided from birth up until a child is observed.

3.2 Endowments, Private Investments, and Public Health

The biological content of health capital theory resides in the health production function $f(H_0, I^*)$, and differences in health are attributed to differences in H_0 and I^*. That is, health capital theory conceives of two sets of potential causal channels that determine a child's realized health at any given point in their lives: health endowments (H_0) and health input streams (I^*). The legacy and contemporary reality of caste socioeconomic disparities guarantee that both of these causal channels are likely to operate on caste HAZ disparities: on average, UC women are healthier when they give birth, and have more resources to provide their children after birth.

The health capital model defines the health endowment as the health of a child at birth, an initial store of health (energy, potential, genetics) bequeathed (pre-birth) to a child as they enter the world. We proxy for this theoretical construct using caste-level estimates of implied birth length and very early life HAZ. Determinants of birth length, birth weight, or other very early life health measurements include maternal height, weight, age, and birth order (Currie, 2009; Ahmed and Ray, 2018; Chari et al., 2017; Özaltin et al., 2010; Maertens, 2013; Swaminathan et al., 2019; Von Grafenstein et al., 2023). Timely interventions on children born in the hospital can also play a role in ensuring that children are healthier at birth (Godlonton and Okeke, 2016; Daysal et al., 2015). We consider maternal health variables, birth order, and pre-natal and delivery care as observable characteristics related to the determinants of the health endowment, and thus a child's length at birth.

After-birth, health capital is determined jointly by this initial health endowment and the subsequent stream of health inputs a child experiences up until they are measured. Empirically, we thus consider differences in the rate of growth of child HAZ as evidence of interactions between initial health endowments and the subsequent stream of health inputs provided to the child. This investment stream can be conceptually divided into two sub-types: private investments made by the household, and public investments affecting the availability of public goods and the health environment. Private investments are made by individual households, and can be proxied most directly by household wealth or assets (land ownership, bathroom facilities) or parental characteristics that allow parents to productively invest their time in children (education) (Swaminathan et al., 2019; Attanasio et al., 2020; Bhalotra et al., 2020). Alternatively, public goods like sanitation and health environment affect all children in an area (Spears, 2020; Geruso and Spears, 2018). We conceptually and empirically separate these two investment streams to measure their relative importance in determining caste HAZ gaps.

Health capital is thus a cumulative, lifelong process that, at any given age, contains within it an initial endowment component (H_0) and a subsequent investment component related to the cumulative effect of the stream of privately and publicly provided health investments (I^*) a child has received up until that point in their lives. We acknowledge that this categorization of household variables as exclusive determinants of either health endowments or health investments, while not arbitrary, is to some degree both ad hoc and imprecise. It is clear that many variables, such as maternal height or household wealth, will affect both child health endowments and the subsequent stream of child health inputs (at least indirectly). We discuss this concern in detail in Section 4.2

3.3 Height-for-Age Z-score (HAZ)

A good measure of health capital that is a cumulative measure of both health at birth and the subsequent investment stream is HAZ.

Child height-for-age Z-score is an age- and gender-normalized measure of child length. The Z-score measures we employ have been standardized across developed and developing countries by the World Health Organization (WHO) using a reference population of relatively well-nourished children (WHO, 2006), with part of the sample coming from an affluent neighborhood in New Delhi. These well-nourished children grow, on average, at the same rates across the sampled countries, implying that deviation from these standards is not determined by genetic origin, but instead by the circumstances of a child's birth and growth. This normalization allows for international comparison of population average child health outcomes and the estimation of the impacts of interventions targeted at improving child health.

Children are born with a certain length, and their growth trajectory is then determined by the interaction of this initial birth length and the stream of nutritional and medical health inputs the child experiences from that point on. That is, HAZ contains within it information on the cumulative health experience of children from birth until they are measured, in the same manner as health capital.

Other anthropometric measures of child health, such as those based on weight or arm circumference and which change rapidly in the short term, cannot claim to be such acceptable proxies for health capital and its nature as a cumulative stock accrued as people age.

3.4 HAZ-age Profile

Victora et al. (2010) were the first to document the age-dynamics of HAZ as a consistent feature of child health in developing countries. They show that, across the developing world, children are born with HAZ that is slightly lower than healthy populations (slightly below 0). These children then grow much less quickly than the median child in the reference population up until the child is around 2 years old. Further work has shown that this pattern is consistent across social groups, and is not explained by observable covariates (Rieger and Trommlerová, 2016; Roth et al., 2017; Alderman and Headey, 2018).

Given this consistent and stable pattern in global child growth patterns, Aiyar and Cummins (2021) develop models to quantify changes not in child HAZ, but in the shape of the HAZ-age profile. They outline the use of regression models to capture the effects of a key variable of interest on changes in the shape of the HAZ-age profile. They also develop the health capital accumulation interpretation we employ, relating changes in the HAZ-age profile intercept at age 0 (implied birth length z-score) to differences in health endowments, and changes in the slope of the HAZ-age profile to differences in the interaction between the health endowment and the post-birth health investment stream.

Our work extends this age-profile empirical perspective aimed at disentangling the determinants of differential child growth trajectories to the realm of decomposition methods. While Aiyar and Cummins (2021) were interested in estimating correlations between one key variable and changes in the shape of the HAZ-age profile, here we are interested in how suites of variables, defined in relation to theoretical economic mechanisms, can explain differences in two HAZ-age profiles solely distinguished by a child's caste.

4 Data

4.1 Data Source

Our primary child-level dataset of outcomes and covariates come from the NFHS 2015-16 (NFHS-4), the fourth of the Indian NFHS series.[1] The dataset is population representative at the district level. Our main estimation sample consists of information on 146,778 Hindu children below the age of five. We utilize the definitions of caste as recognized

[1]In Appendix Section 8 we present a description of the data used in our analysis of NFHS-5 for Section 6. Individual level information in this dataset was made public after our analysis of the NFHS-4 data was complete and disseminated. Hence all analyses in the results section uses data from NFHS-4. In Section 6, we use the same methods to compare results from our analysis of NFHS-4 and NFHS-5.

by the NFHS-4[2]. Table 1 provides summary statistics. Indian children are on average -1.49sd shorter than the WHO reference population median. Caste groups follow IPUMS-DHS recoding to include: Scheduled Caste (SC), Scheduled Tribe (ST), Other Backward Classes (OBC), and Upper Caste (UC). Overall, in our sample, 23.7% are SC, 14.6% are ST, 43.7% are OBC, and 17.9% are UC. As expected, lower caste and tribal children have lower HAZ scores on average than their UC counter parts. SC children are 1.67sd below, OBCs are 1.49sd below, and ST children are around 1.72sd below the reference height for their sex and age, while UC children are only 1.12sd below.

4.2 Variable Groups

The NFHS-4 provides information on a large number of child, parent, household and community level characteristics. Motivated by health capital theory, we separate our observed covariates into three groups: endowment variables, private investment variables, and public health variables. A table of representative summary statistics is provided in Table 1, and the full table including all variables is available in the Appendix Table A.1.

As mentioned in our discussion of health capital theory, our categorization of household variables as exclusive determinants of either health endowments or health investments, while not arbitrary, is imprecise. It is reasonable to argue that many variables, such as maternal height or household wealth, will affect both child health endowments and the subsequent stream of child health inputs (at least indirectly). We stress that our goal is not to estimate the causal impact of any particular variable in explaining these caste disparities but rather to highlight the two distinct channels (endowment and investment) as determinants of the child HAZ gaps. An empirically equivalent definition of "endowment" and "investment" related variables would be variables that could (in theory) be

[2]Our main estimates compare SC/ST/OBC children to UC children, and children from other religions are dropped. The majority of children of other religions are Muslim, and in the Appendix we provide an analysis comparing Muslim Indian children with UC Hindu Indian children.

acted upon either before (endowment) or after (investment) a child is born. Child birth order may have effects on the within-household distribution of resources post-birth, and maternal height may continue to operate post-birth, but by definition birth order cannot be altered after birth and maternal growth spurts after birth are unlikely to be largely determinative of child growth rates. Alternatively, household wealth is likely to have cumulative effects over the life-course by affecting the stream of post-birth child health inputs, and thus post-birth intervention or manipulation that alters the household budget constraint may be effective at improving child growth. When we frame the distinction in this manner, our theoretical distinction among metaphysical objects (endowments and investments) becomes an operational distinction regarding the locus of any potential intervention aimed at alleviating caste HAZ disparities (before or after a child is born). With this caveat in mind, we now proceed with our three families of variables in explaining the caste disparities in child HAZ.

4.2.1 Health Endowments

Endowment variables include birth order, maternal age at child's birth, maternal HAZ and WAZ, and delivery care. From Table 1, we see that Indian mothers are on average 24.23 years old, have around 2.14 children and are 2.03sd shorter than their age-gender median healthy height across the world. UC Mothers are on average healthier than lower caste mothers. They are slightly less likely to have more than two children, and more likely to report having given birth in a hospital. ST mothers are only slightly less tall than SC mothers but are much less likely to give birth in a hospital. SC and OBC mothers share similar baseline health and are about half as likely to report having their deliveries in a hospital while 71% of UC mothers deliver in hospitals.

4.2.2 Private Investments

Private investment variables include wealth, maternal education, motorcycle ownership, ownership of agricultural land, access to treated water, owning a private toilet or shared toilet, vaccination status and post-natal care. 71% of mothers have at least a primary education. About 32% of the entire sample had access to clean drinking water, and 13% of villages had sewer system access. Agricultural land is owned by 45% of households, while 24% live in urban settings. On average, UC members tend to be better off in terms of private investments. UC mothers are more likely to have some education (89%) and are less likely to be among the poorest asset quintile and more likely to be among the richest in the sample.

Among the disadvantaged groups, SC and ST households are much more likely to be among the poorest quintile in asset ownership and less than 10% are in the richest quintile. OBC households, on the other hand, are slightly more likely than SC/ST households to be among the richest, though women from the UC dominate the richest category. Around one-third of lower caste and tribal woman have no education, and less than 10% of mothers have higher education. Interestingly, OBCs own assets like motorcycles and land at a similar rate as UC mothers. ST households are more likely to own agricultural land but less likely to own means of transport than SC households. Treated drinking water is more prevalent for ST (38%) and UC (44%), while SC lag at 22% of households. About half of SC, ST and OBC mother have no access to household toilets. Nearly 90% of all children in these groups have completed their Bacille Calmette-Guérin (BCG) vaccinations for tuberculosis, with measles vaccine take-up at the lowest of around 70%. In this dimension, children among ST groups fare the worst in terms of vaccination completion but all other groups are comparable.

4.2.3 Public Investments

Public health variables include access to a sewer system in the primary sampling unit (PSU), state-urban dummies, and whether distance to health facility is a barrier to health care access. On average, 13% have a sewer system, 24% live in urban areas, and 65% face barriers to access a health facility. UCs have more access to sewer systems and are more likely to live in urban areas while STs are the least likely to have access to the system and also half as likely to live in urban areas. The geographical distribution of SCs and OBCs mothers are similar to the national average. Regional effects (dummy variables for state-by-urban) are also considered effects of public health goods. Location indicators at the state-by-urban level capture differences in environment, markets, resources, infrastructure and access to services that are locally important but unobservable in our data.

Overall, we see large caste differences on most traits related to endowment, private investment and public health, and smaller differences on mother's age at child birth, vaccination completion for children, postnatal care by health worker/facility, and distance of the health facility.

Table 1: Summary Statistics 1 (NFHS-4)

	(1) SC mean/sd	(2) ST mean/sd	(3) OBC mean/sd	(4) UC mean/sd	(5) All mean/sd
HAZ (x 100)	-166.79	-171.84	-148.69	-112.61	-149.29
	(164.41)	(169.98)	(166.18)	(160.43)	(166.28)
Endowment Variables					
Mother's Age at Child's Birth	24.21	24.03	24.21	24.44	24.23
	(4.78)	(4.95)	(4.53)	(4.52)	(4.64)
Mother's HAZ (x 100)	-217.65	-211.85	-201.27	-182.60	-203.18
	(93.00)	(89.77)	(95.68)	(94.38)	(94.87)
Birth Order	2.29	2.29	2.12	1.86	2.14
	(1.46)	(1.43)	(1.31)	(1.04)	(1.33)
Private Investment Variables					
Poorest	0.33	0.50	0.23	0.08	0.26
	(0.47)	(0.50)	(0.42)	(0.28)	(0.44)
Maternal Educ at least Primary	0.65	0.55	0.71	0.89	0.71
	(0.48)	(0.50)	(0.45)	(0.31)	(0.45)
Treat Drinking Water	0.22	0.38	0.31	0.44	0.32
	(0.41)	(0.49)	(0.46)	(0.50)	(0.47)
Owns Agricultural Land	0.33	0.55	0.48	0.49	0.45
	(0.47)	(0.50)	(0.50)	(0.50)	(0.50)
Public Health Variables					
PSU has sewer system access	0.10	0.07	0.12	0.23	0.13
	(0.31)	(0.26)	(0.33)	(0.42)	(0.34)
Urban = 1	0.23	0.11	0.25	0.39	0.26
	(0.42)	(0.32)	(0.43)	(0.49)	(0.44)
N	34913	21441	64206	26218	146778

This table shows the summary statistics for endowment, private investment, and public health variables. It contains data from NFHS-4/2015 IPUMS-DHS in India that are used in the analyses. Results are weighted by sample weights.

5 Empirical Methods

We employ three main empirical methods to estimate and explain caste HAZ disparities. We use standard linear regression techniques, disaggregated by child age group, to estimate both the magnitude and age-dynamics of unconditional and conditional HAZ differenctials across caste groups. To capture measures more closely aligned to our health capital concepts of endowment and investment effects, we then augment these non-parametric estimates with the parameterized age-profile methods described in Aiyar and Cummins (2021) that estimate caste gaps in (implied) birth length and the rate of growth of young children. Finally, we decompose the caste gaps using simple Oaxaca-Blinder (OB) decompositions to quantify the contribution of the relative influence of the three theoretical health capital channels toward explaining the HAZ gaps as children age: endowments, private investments, and the public health environment.

5.1 Estimating Unconditional and Conditional Caste Gaps

We first estimate age-specific mean HAZ differences across caste group (relative to UC children) using a standard linear regression model:

$$Y^a_{irvg} = \delta^a_g * Caste^a_g + X'^a_{1irvg}\beta^a_1 + X'^a_{2v}\beta^a_2 + \lambda^a_r + \epsilon^a_{irvg} \tag{1}$$

where Y^a_{irvg} is the HAZ of child i, aged a, living in state r in primary sampling unit *(PSU)* v and belonging to caste g. $Caste^a_{irvg}$ is a vector of indicator variables representing the caste to which the child belongs: SC, ST or OBC (with UC comprising the omitted reference group). X_1 and X_2 are individual-level explanatory variables and village level variables respectively, and λ^a_r is a vector of indicator variables for each DHS state-by-urban location, each specific to child age group, a. The coefficients of interest, δ^a_g, are age-specific estimates of the HAZ gap between UC children and the other caste groups.

We always estimate this equation separately for each 6-month child age group, allowing δ, β, and λ to vary by age.

We first estimate Equation 1 omitting X_1, X_2, and λ to estimate the unconditional HAZ gap between UC and other castes. The regressions are weighted by individual survey weights and standard errors are clustered at the primary sampling unit level. The resulting coefficients constitute our estimate of the true population gap in HAZ. We then include X_1, X_2, and λ and re-estimate Equation 1 to estimate the adjusted caste HAZ gaps given observable covariates of households and children. The estimates here are interpreted as estimates of the height gap that would exist in the population if the observed covariates included in the regression model were evenly distributed across the caste groups.

5.2 Estimating Proxy Parameters: Implied birth HAZ (α) and rate of loss of HAZ (β)

Aiyar and Cummins (2021) propose an alternative method to estimating how covariates of interest affect child health endowments and investments. As a complementary approach to estimating age-specific coefficients of the same model to trace out HAZ disparities across age, they attempt to more directly estimate the determinants of the HAZ-age profile using a two step quasi-structural approach. They first estimate the group-level parameters of a stylized structural HAZ-age profile, specifically the intercept (birth length) and slope (rate of growth) of the average outcome-age profile for that group. Then, in a second stage regression, these estimates are used as observations to estimate the determinants of the group-level parameters themselves.

This allows the model to focus on two particular features of the HAZ-age profile: the y-axis intercept (the implied group-average birth length), which is an empirical counterpart to the health endowment; and the slope of the HAZ-age profile over the first two years of life (rate of loss of HAZ), which is related to the interaction of the health endowment and the subsequent stream of health inputs provided to a child. Furthermore, when we bin regressions by age, we lose precision on the intercept itself (we estimate a single coefficient for 0-6 month olds) and on the rate of child growth (by failing to borrow information from observations across age-bins). By parameterizing the group-level HAZ-age profile as an intercept and a slope across age, we can potentially improve precision of parameter estimates, and simultaneoulsy produce estimates that map more closely to the mechanics and predictions of health capital theory.

The model begins with the following equation, estimated only on children aged 0-2. We restrict these regressions to children under the age of 2 to capture the characteristic loss of HAZ over the first two years of life (Victora et al., 2010).

$$Y_i^{rg} = \alpha^{rg} + \beta^{rg} * Age_i^{rg} + \epsilon_i^{rg} \tag{2}$$

Y_i^{rg} is the HAZ of child i from state r and caste-group g. $\hat{\alpha}^{rg}$, the estimate of the implied (sub-group) average birthlength z-score, is interpreted as a measure of the average health endowment for caste group g in geographic regions r, defined in our main specifications as state-by-urban groups. Similarly, $\hat{\beta}^{rg}$ estimates the rate of loss of HAZ over the first two years of life for caste group g in region r. Slope coefficients are interpreted as the result of the (subgroup average) interaction between the initial health endowment and the subsequent stream of consumed health inputs. Regressions in this first stage are weighted by survey sampling weights.

We then estimate caste level differences in $\hat{\beta}$ and $\hat{\alpha}$ with the following regression equation:

$$\hat{Z}_{rg} = \delta^g * Caste_g + Endowment'_{rg}\beta_1 + Private'_{rg}\beta_2 + Public'_{rg}\beta_3 + \epsilon_{rg} \tag{3}$$

Here, $\hat{Z}_{rg}$ is either the health endowment, $\hat{\alpha}_{rg}$, or rate of loss of HAZ, $\hat{\beta}_{rg}$ from Equation 2. Again δ^g is the estimated caste gap for health endowment or rate of growth.

As above in the individual-level regressions, we estimate unconditional versions of Equation 3 and conditional regressions that adjust the gap for observable covariates. For the conditional regressions, we include the state-urban-caste-group mean of our endowment and private investment variables. Public health variables are captured by state-urban fixed effects. These regressions are run at the state-urban-by-caste level and are weighted by cell-size of each state-urban-caste group. Standard errors are clustered by state-urban-caste group.

5.3 Decomposition

In previous studies and our own, the unconditional caste gap estimates are large, but the conditional caste gap differences are small. This motivates a decomposition exercise that attempts to quantify the contributions of the different covariates towards explaining the unconditional disparity estimates. We employ a Oaxaca-Blinder (OB) decomposition to answer that question, and provide in the Appendix an alternative decomposition based on Gelbach (2016).

In the OB decomposition framework, the difference in mean HAZ between caste group g and the UC (baseline) for a specific age group a can be represented as:

$$\overline{HAZ}^{g} - \overline{HAZ}^{UC} = \overline{\beta}^{g}\bar{X}^{g} - \overline{\beta}^{UC}\bar{X}^{UC} \tag{4}$$

The Oaxaca-Blinder decomposition then re-organizes those terms into:

$$\overline{HAZ}^{g} - \overline{HAZ}^{UC} = \hat{\beta}^{g}(\overline{X}^{g} - \overline{X}^{UC}) + \overline{X}^{UC}(\hat{\beta}^{g} - \hat{\beta}^{UC}) \tag{5}$$

The first and the second terms on the RHS are commonly referred to as the explained and the unexplained portions of the HAZ gap, respectively. In our case, the explained effect tells us the difference in mean HAZ due to the differences in the average level of covariates between lower caste children and UC children. The unexplained effect is interpreted as the difference in the returns to the covariates. In other words, if caste group g was given the mean level of observed covariates as that of the UC group ($\bar{X}^{UC}$), the remaining difference in HAZ between the groups would be apportioned to differences in the returns to those covariates, and ascribed to the unexplained portion of variation. For ease of interpretation, we present the explained share of variation from the OB decompositions in percentage points, calculated as the explained variation (in HAZ units) divided by the total gap in the unconditional regressions (also in HAZ units).

6 Results

6.1 Unconditional and Conditional Estimates

Figure 2 (Top Panel) shows regression estimates from the unconditional and conditional model from Equation 1, tracing out mean HAZ differences across caste groups by age. Each point estimate is the coefficient of a caste group relative to the base category of UC children. The y-axis represents the coefficient estimates (and 95% confidence intervals) on the caste-group variables, and the x-axis separates the estimates into 6-month age-bins. In the unconditional model, all children in lower caste and tribal minority groups are born smaller than UC children and experience post-birth growth rates that leave them much shorter than UCs by the time they have completed early childhood.

OBC children start around 0.21sd below UC children at birth. OBC children then grow less quickly than UC children after birth and average differences increase to around 0.36sd by an OBC child's third birthday. These differences reduce to around 0.33sd by the 4th year, providing some evidence of catch up in later years.

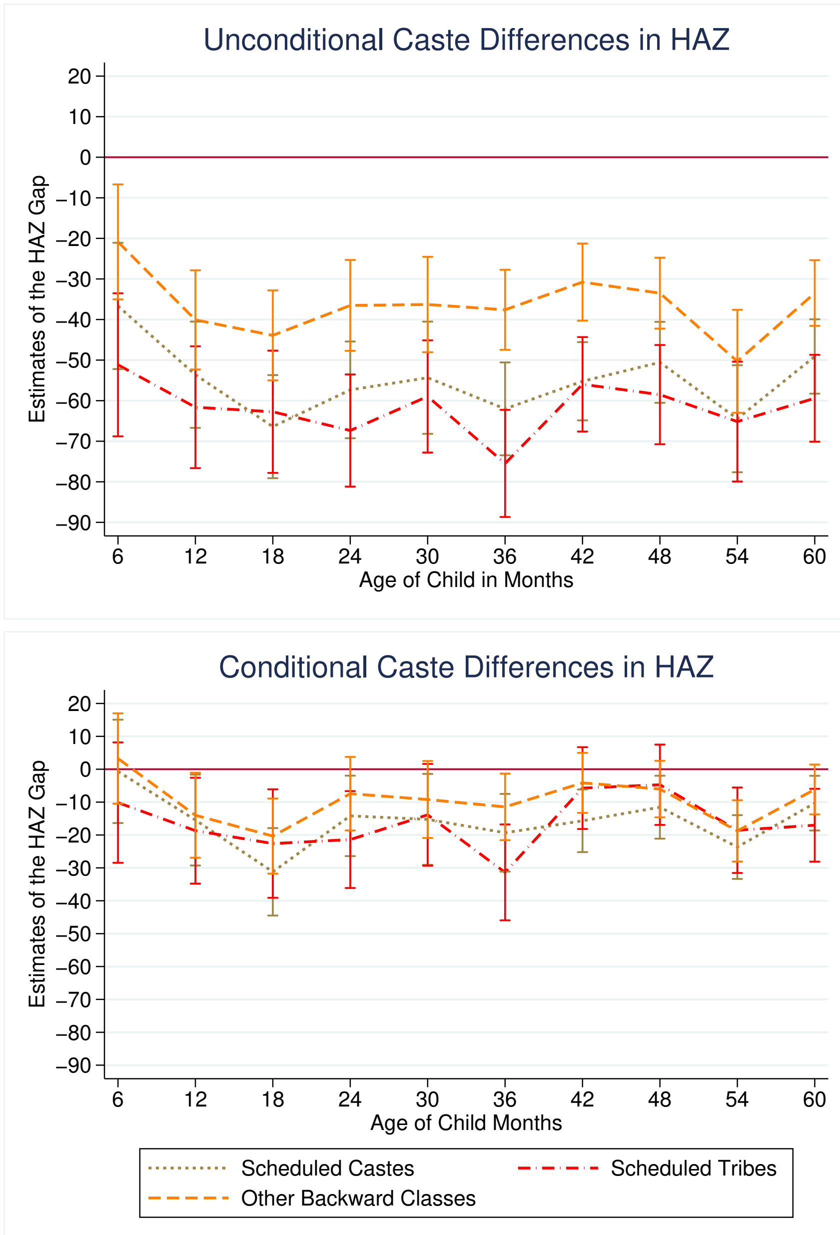

Figure 2: Regression Estimates of Caste HAZ Differentials by Age (0-60 months)

This figure presents regression estimates for both unconditional and conditional regressions in Equation 1. The y-axis provides the coefficient estimates (and confidence intervals) on the caste-group variables, and the x-axis separates the estimates into 6-month age-bins. The top panel presents regression estimates from the unconditional model. In the bottom panel, the conditional regression estimates include all controls related to health endowments, private investment, and public health. The regressions are weighted by survey weight and clustered at PSU level.

SC children are born with HAZ scores around 0.37sd below UC children. Over the first three years, they lose, on average, an additional 0.25sd in HAZ, leading to a 0.62sd difference by their third birthday. After that, heights improve slightly and by the child's fifth birthday they are on average 0.5sd below an UC child.

ST children experience the largest deficit of around 0.51sd at birth. They face a large dip in HAZ by their 3rd birthday where differences grow to around 0.75sd. By the time they are done with early childhood, children are 0.6 SD below UC children.

Figure 2 (Bottom Panel) provides conditional (regression adjusted) caste gap estimates controlling for covariates related to the health endowment, private investment, and public investment variables described above. Including covariate adjustment in the model reduces differences in HAZ differences across all caste groups.

STs, who face the largest unconditional HAZ gap relative to UCs, see reductions in the gap at birth by around 0.41sd once covariates are included. Adjusted differences in birth length differences are indistinguishable from zero. After six months, adjusted differences increase to 0.3 sd by a child's third birthday, before settling to 0.2 sd by a child's fifth birthday. These estimates are consistent with the observable covariates explaining the ST birth length gap in its entirety, and between 65-85% of the gap with UC children by age 5.

Adjusted for observed covariates, SC and OBC children are statistically insignificantly shorter than UC children at birth. However, as children age, differences begin to appear. Coefficients for OBCs children hover between 0 and 0.2sd across most ages, with estimates for only 4 of the 10 periods statistically distinguishable from 0 at 95% confidence level. Adjusted estimates for SCs show similar patterns with somewhat larger point estimates that hover between 0.05 to 0.31sd, with 9 of the 10 estimates being significantly different from 0. Even still, covariate adjustment reduces the height gaps between OBCs and STs with UCs by well over half.

6.2 Intercept (α) and Slope (β) Estimates

An alternative way to investigate the age dynamics of HAZ caste disparities is to focus on the location and shape of the HAZ-age profile itself, and then estimate the determinants of that shape (Aiyar and Cummins, 2021). Results from Equation 3, estimating our measures of implied birth length and rate of growth, are presented in Table 2. Columns 1 through 4 in Table 2 provide our baseline estimates of caste differentials in the intercept (α) and slope (β) of the HAZ-age profile over the first two years of life. The first two columns present estimates for caste gaps in α and the second two columns provide estimates for β. Columns 1 and 3 present the unconditional estimates, where the second stage regression (Equation 3) is estimated without covariates, and columns 2 and 4 provide estimates when state-caste-urban group level mean covariates are included in the second stage.

Table 2: Rate of HAZ Loss and Caste (NFHS-4)

	(1) α b/se	(2) α(| X) b/se	(3) β b/se	(4) β(| X) b/se	(5) β(| α) b/se	(6) β(| X,α) b/se
SC	-41.98***	-3.50	-1.13***	-0.85	-2.91***	-1.04**
	(6.58)	(12.30)	(0.39)	(0.73)	(0.38)	(0.44)
ST	-52.02***	-39.34**	-0.64	1.98*	-2.84***	-0.17
	(10.75)	(16.75)	(0.73)	(1.13)	(0.62)	(0.67)
OBC	-27.32***	-12.49	-0.78**	0.15	-1.94***	-0.54
	(5.44)	(8.72)	(0.35)	(0.53)	(0.28)	(0.33)
Mean	-29.2	-29.2	-7.5	-7.5	-7.5	-7.5
R Square	0.16	0.78	0.02	0.76	0.46	0.94
Weighted N	55679	55679	55679	55679	55679	55679
Real N	183	183	183	183	183	183

The results are weighted by numbers of individuals in each state-caste-urban cell and clustered at state-urban level. The covariates used include endowment, private investment, and public health variables. The state-urban fixed effects are included in the public health variables. Age Cutoff = 24 Months. p-values: * 0.10, ** 0.05, *** 0.01.

The results on birth length are similar to those in the first age-bins of the individual-level regressions. SC children are on average born a statistically significant 0.42sd shorter than the average UC child. Conditional on the inclusion of covariates, this difference reduces to 0.03sd and is statistically insignificant. OBC children have a statistically significant average unconditional birth length deficit of 0.27sd relative to UCs, a smaller gap compared to SC, but these differences reduce to 0.11sd and are statistically insignificant once we condition on household covariates. ST children have statistically significant birth length deficits both with and without conditional covariates. Unconditionally, ST children are born 0.52sd below UC. Conditionally, this gap tightens to a still significant 0.37sd.[3]

A similar picture emerges from the slope (β) estimates, where differences in growth rates are also largely attenuated when adjusted for observable group covariate means. Unconditional caste group point estimates in Column 3 are negative and statistically significant for SCs (-0.011sd/month) and OBCs (-0.0078sd/month), but not for STs. The adjusted estimates for SC and OBC estimates are insignificant and closer to 0. For STs, the adjusted estimates suggest that the rate of loss increases by 0.0198 sd/month.

Columns 5 and 6 in Table 2 provides a second set of estimates for β and represent the results of a slightly different motivation and thought experiment. We know from previous results that lower caste children are born shorter. If there is a natural relationship across all castes in which birth length affects rate of growth, then any attempt to capture a meaningful correlation between caste and relative rate of child growth (separate from an endowed birth weight effect) would need condition on the birth length. We thus present results in columns 5 and 6 of Table 2 including $\hat{\alpha}$ as a control variable in Equation 3. The model implicitly allows the slope of the HAZ-age profile to vary based on the intercept, in a manner common across caste groups.

Column 5 provides estimates from a regression of $\hat{\beta}$ on caste dummy variables and group-specific $\hat{\alpha}$ and column 6 presents estimates conditional on covariates. Conditional on intercept estimate, the slope estimates indicate generally large growth rate gaps be-

[3]These estimates come from restricting the estimation sample to children under the age of 24 months, to focus on the faltering of HAZ scores over the first two years of life. Figure A.1 provides estimates at alternative cutoffs and the results are robust to selection of age cutoff

tween UC children and SC (-0.029 sd/month), OBC (-0.019 sd/month) and ST (-0.028 sd/month) children. These gaps would produce a cumulative effect of between -0.45sd (for OBC) to -0.7sd (for SC) HAZ points by 24 months of age. Point estimates for STs and OBCs get very close to zero and become statistically insignificant with the inclusion of group level covariates, while the estimate for SCs remains large but much smaller than the unconditional estimates (-0.01 sd/month, or a cumulative effect of -0.24sd by age 2).

6.3 Decomposition Results

Figure 3 shows the Blinder-Oaxaca decomposition results, presented in percentage of unconditional height gap as described in Equation 4 and Equation 5. In the top left panel, child age measured in 6-month age groups is on the x-axis and the y-axis is the percent of the caste gap that can be explained by model covariates.

In general, well over half of the unadjusted caste HAZ differences can be explained by our observable covariate groups. Birth length is almost fully explained, but during the critical year following, when child HAZ drops most rapidly, the covariates lose some explanatory power. However, from 18 months onward, the model explains an increasing share of the caste HAZ gaps, so that over 80% of the gap is explainable by age 5. Of additional interest is the fact that the dynamic effect of covariate explanatory power across age is consistent across caste groups as well - the functioning of the covariates on HAZ appears to be stable across all groups of children.

The remaining panels of Figure 3 decompose this explained variation into the percent of the total caste gap explained by each of our three families of explanatory variables. As theory would predict, endowment variables explain the largest share of the explained variation in caste gaps for the youngest children in all caste groups. Private investment variables have almost no explanatory power over newborn caste HAZ gaps.

As children age, the explanatory power of endowment variables remains relatively

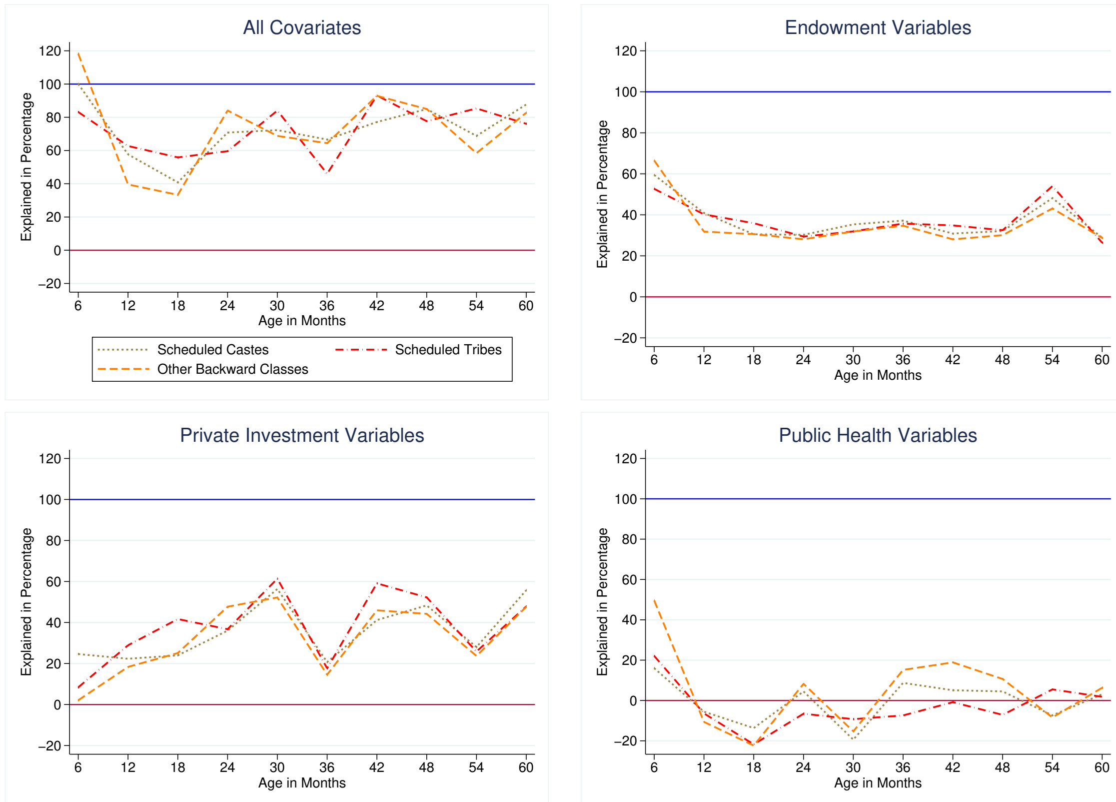

Figure 3: Blinder-Oaxaca Decomposition Results: Percent Explained

This figure presents Blinder-Oaxaca decomposition results. Child age measured in 6-month age groups is on the x-axis and the y-axis is the percent of the caste gap that can be explained by model covariates. The top left panel presents the percent of the HAZ gap that is explained when all control variables are included. The top right panel is the percent of the HAZ gap explained by endowment variables. The bottom left panel depicts the percent of the HAZ gap explained by the private investment variables. The bottom right panel presents the percent HAZ gap that is explained in percentage by public health variables.

stable, decreasing only slightly to about 30% by age 5. The influence of private investment variables, though, begins to increase as children age. By age 5, half of caste HAZ differentials can be explained by the private investment related variables.

Public health variables tend to have much smaller and inconsistent explanatory power over caste HAZ gaps. The only group for whom public health variables explain considerable fraction of the caste gap is for OBCs. The estimates, though, are difficult to interpret, given that location and public infrastructure effects are estimated to mostly exacerbate OBC deficits in the first two years, and mostly reduce deficits in later years. In general, we find surprisingly small associations between public health variables and child HAZ, despite the well-known importance of public goods like clean water and sanitation services. While public health environment is clearly an important determinant of child health, it is not, according to our estimates, important to explaining caste HAZ gaps.

6.3.1 Timing of Investments

One econometric concern about our estimates of the increasing influence of investment related variables is that, unlike endowment related variables such as maternal health, they are sometimes increasing in value, or probability, as children grow. Newborns do not receive many vaccinations, and children under the age of 6 months have only had so many health care visits. In order to determine the extent to which our increase in the explanatory power of private investment may be due simply to an increase in the magnitude or probability of age-determined covariates, we divide our Oaxaca-Blinder decomposition results for private investments into age-invariant and age-varying sub-types. We then graph out the relative contribution of each sub-type to the overall explanatory power of private investments across age in Figure 4.

The top row of Figure 4 graphs the outcome-age profile for two representative variables, one from each sub-type of private investment variables. The upper left panel graphs the age-profile of an important age-invariant private investment, the fraction of children in households at each asset quintile across child age. Household wealth is not correlated with child age, at least for very young children.

On the other hand, one can essentially infer the vaccination schedule for children from the top right panel, which graphs the age-profile of various vaccine take-ups, a representative age-varying private investment. If vaccinations are driving the explanatory power of private investments, then the apparent increase in the explanatory power of private investments variables could, in theory, be simply the result of the increasing take-up of vaccines as children age. That is, the estimates of increasing explanatory power of

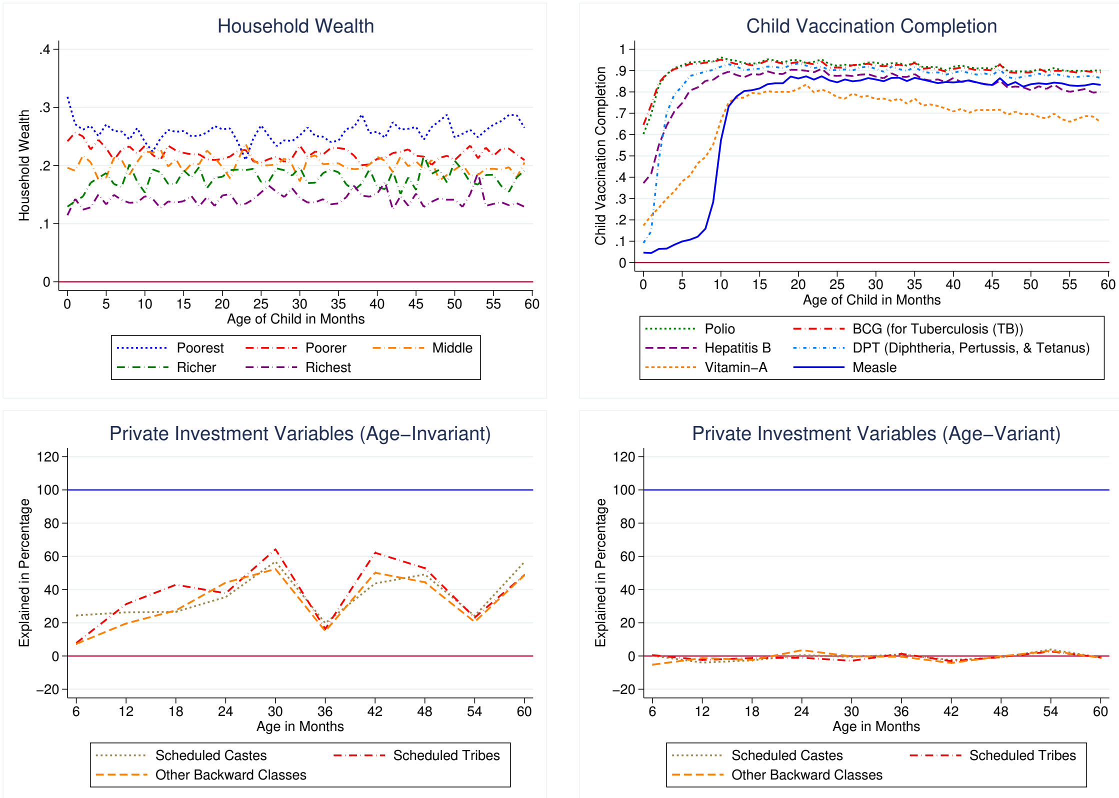

Figure 4: Age-Specific and Non-Age-Specific Private Investments

The top two panels of the figure shows the frequency of two particular private investment variables for children from 0 to 5 years. The top left panel shows the age-invariant nature of the household wealth measure of private investments across child age, while the top right panel shows the age-specific frequency of vaccination status by child age. The bottom row displays Blinder-Oaxaca decomposition results separately for age-varying and age-invariant private investment variables. Child age measured in 6-month age groups is on the x-axis and the y-axis is the percent of the caste gap that can be explained by model covariates. The bottom left panel presents the percent of the HAZ gap explained by the private investment variables that are age-invariant. The bottom right panel depicts the percent of the HAZ gap explained by the private investment variables that vary by age.

private investments would be driven by some of them simply turning on.

The bottom row of Figure 4 shows the Oaxaca-Blinder decomposition for the age-invariant and age-varying private investment variables. The left panel shows that the private investment effects are mostly driven by age-invariant private investment and that age-varying variables such as vaccinations and prenatal care do not explain much of the HAZ gaps.

6.3.2 Re-Scaled Decomposition

The decompositions above are based on the explanatory power of our families of covariates relative to the unconditional HAZ gap between groups. This makes sense as a baseline counterfactual to the extent we are interested in quantifying our ability to account for the health disparity itself. However, we might also be interested in the extent to which the covariate families explain the change in our regression estimates themselves, from conditional to unconditional, without regards to the underlying health disparity.

We provide these re-scaled estimates in the Appendix Table 8, using the decomposition method proposed in Gelbach (2016). The results are qualitatively and quantitatively similar, and the re-scaling does not affect our inferences or conclusions from the OB decompositions.

6.4 Additional Analyses

We provide two additional analyses speaking to issues previously addressed in the literature. First, following Bhalotra et al. (2010), we present results comparing Muslim children with UC Hindu children. Second, following Coffey et al. (2019), we compare estimates across regions with higher and lower concentrations of UC Hindus.

6.4.1 Muslim Children

Our analysis thus far has focused exclusively on Hindu children. A natural question, though, is the extent to which the patterns we see in HAZ disparities among Hindu children resemble patterns in HAZ disparities with another marginalized group of Indians: Muslims. We thus estimate our regression and decomposition exercises comparing Muslim children with UC Hindu children to compare the dynamics and determinants of UC-Muslim HAZ gaps.

Compared to UC-Hindus, Muslim children are 0.2 sd shorter at birth and experience a slower growth rate, generating a cumulative difference of 0.3sd by age 5 (Figure A.3). These results are in line with Bhalotra et al. (2010) who find that Muslim children are more likely to be stunted when compared to UC-Hindus. The decomposition results comparing Muslim and UC-Hindu children lead to the same broad conclusion as comparisons among different caste groups of Hindu children - endowments matter early on, the effect of private investment variables accumulates over time, and public health variables don't seem to explain much of HAZ differences between UC and Muslim children (Figure A.5). We elaborate further on these results in Appendix 8.

6.4.2 Concentration of Upper Caste Hindus

Coffey et al. (2019) and LoPalo et al. (2019) highlight the role of exclusionary social practices like untouchability in preventing access to adequate sanitation or health services among SCs. Following Coffey et al. (2019), we divide our sample into those living in regions with high or low UC share, a variable they use to proxy for local discrimination.

Our results, presented in Figures A.7 and A.8, do not significantly differ across these groups of regions. Children from both sub-samples, regions with high share of UC and low share of UCs, experience the same endowment and investment disadvantages.

7 Replication Using NFHS-5

We replicate the analyses above using the most recent wave of the NFHS conducted in 2019-20 (NFHS-5), which provides an independent sample of over 250,000 children. Initial choices over sample selection and regression models employed in this paper, along with the economic interpretation of our results, were developed in 2021 and were presented at several conferences, using data exclusively from NFHS-4[4]. Subsequently, data from the

[4]The NFHS-5 was released in May 2022. Prior to this release, our preliminary work had been presented at seminars at University of Nevada, Reno and the University of California, Riverside, and at the Pacific Development Conference 2022, where a link to a full draft of the initial analysis on NFHS-4 data can be

next round of the NFHS-5 was made available to researchers. We then replicated the results from the NFHS-4 using the same methods and (to the extent possible) the same regression specifications, but applying them to the data from the NFHS-5. That is, we used the methods, specifications and interpretation from our results using NFHS-4 as a kind of *de facto* pre-analysis plan for the NFHS-5 data.

We choose to present the main analysis of the paper using NFHS-4 because the data from the NFHS-5 wave is not complete for some of the variables. Specifically, NFHS-5 only provides information on health inputs (e.g. vaccination and pre-natal check status) for children up to the age of 3 years whereas we have complete information for children up to the age of 5 years in NFHS-4 from IPUMS. The sample that answered the vaccination (Polio, DPT, Hepatitis B, Measles, BCG) question also changed in NFHS-5. These variables are important explanatory variables in our decomposition analysis. Nevertheless, using this newer dataset, we can replicate a majority of the analyses conducted on the NFHS-4.

Figure 5 presents a comparison between the results from the NFHS-4 (left column) and NFHS-5 (right column) analysis. The left column reproduces our main figures using the NFHS-4 for comparison to results from NFHS-5 in the right column. The patterns in the HAZ-profile are quite similar between the two waves. Children who are among the lower caste are shorter than upper caste children at birth and these differences get larger as children age. Similar to NFHS-4, SC and ST children in the NFHS-5 face the largest HAZ disparities.

found here: https://cega.berkeley.edu/pacdev-2022-conference-schedule/.

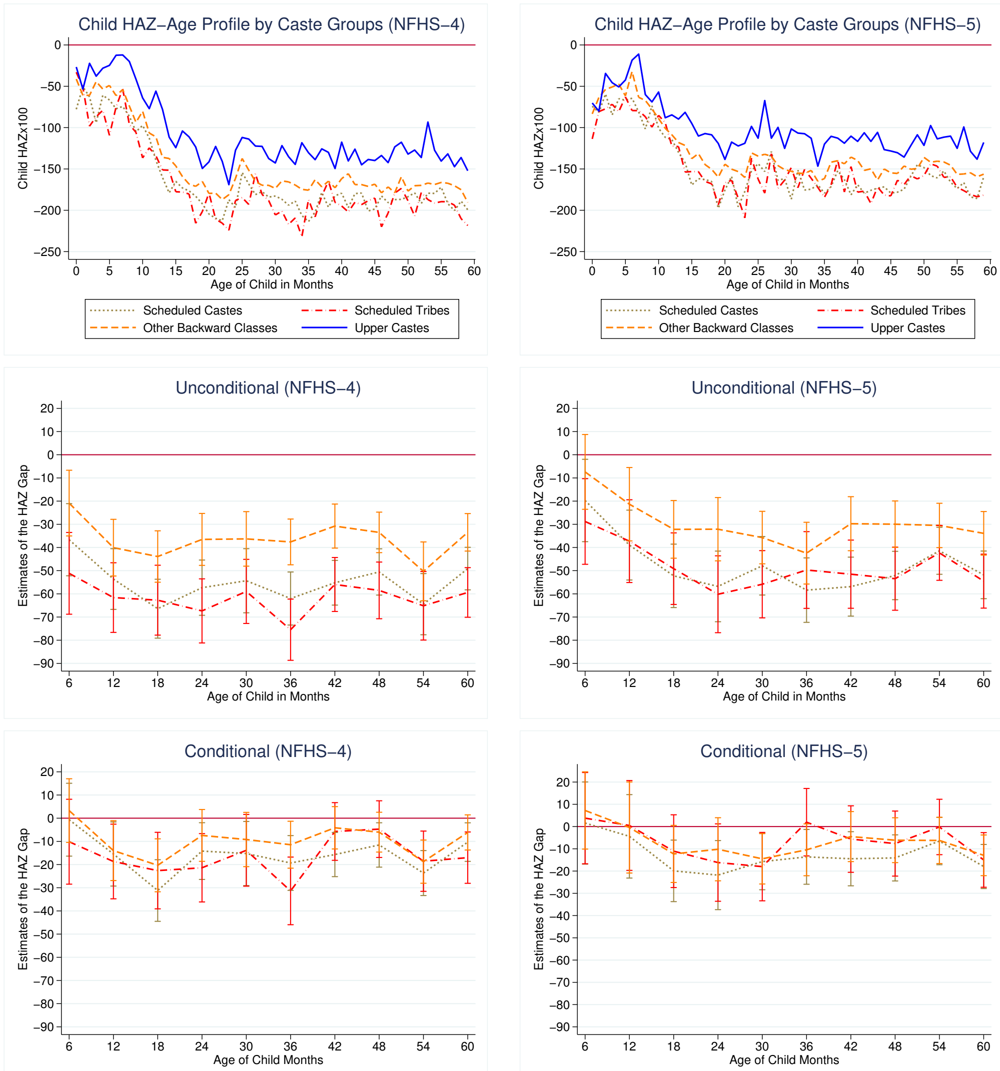

Figure 5: Replication of Individual-Level Results: NFHS-4 versus NFHS-5

The first row graphs mean child HAZ score (x100) by caste groups for children from 0 to 5 years. The left column shows estimates using the NFHS-4 (2015-2016) and the right columns shows the results from the NFHS-5 (2019-2021). The x axis represents age in months and the y axis is the weighted HAZ. The second and third rows graph the regression estimates for both unconditional and conditional regressions as well as the explanatory power. For the bottom two rows, the y-axis is the coefficient estimates (and confidence intervals) on the caste-group variable, and the x-axis separates the estimates into 6-month age-bins. The middle row presents regression estimates from the unconditional model. In the bottom row, the conditional regression estimates adjust for all controls related to endowment, private investment, and public health. The results are weighted by survey weight and clustered at PSU level.

In the second row, we present the unconditional regression estimates of caste group differences in HAZ. Compared to NFHS-4, there seem to be no differences in height at birth between OBC children and UC children. SC and ST children do face a birth length deficit and this increases as the children age. These raw differences in HAZ are slightly smaller than those from the NFHS-4 but show similar dynamics across child age. Conditional estimates, presented in the bottom row, again indicate that caste differences in HAZ can be explained by observable characteristics of households, similar to the NFHS-4 results above and results from other research using NFHS-2 and NFHS-3 (Van de Poel and Speybroeck, 2009; Coffey et al., 2019).

Similarly, in Table 3, we replicate our earlier findings on implied birth length and rate of growth from Table 2 with results from NFHS-5 data. By and large, the estimates are similar in sign and magnitude. Children among lower caste groups start off with lower HAZ at birth (column 1) but these differences can be explained by differences in endowment-related variables (column 2). In Column 3, we see that lower caste children experience slower growth in the first two years of their lives. After conditioning on covariates, the caste gaps in rate of HAZ loss are no longer statistically significant. In columns 5, the rate of loss of HAZ is conditioned on the initial group-level birth length estimate ($\hat{\alpha}$). Children from SC and ST groups again show evidence of slower growth in the NFHS-5, as in the NFHS-4. And again similar to the NFHS-4 results, Column 6 shows that after controlling for covariates and the initial birth length differences, there are no differences in the child growth rates[5].

[5] All other results for NFHS-5, including decomposition results, convey a similar accordance with results from NFHS-4 and can be made available on request. More information on our NFHS-5 analysis, and the compatibility of the data with observations from NFHS-4, can be found in Appendix Figure 8.

Table 3: Rate of HAZ Loss and Caste (NFHS-5)

	(1) α b/se	(2) $\alpha(\mid X)$ b/se	(3) β b/se	(4) $\beta(\mid X)$ b/se	(5) $\beta(\mid \alpha)$ b/se	(6) $\beta(\mid X,\alpha)$ b/se
SC	-25.29***	2.65	-1.68***	-0.49	-3.28***	-0.33
	(5.79)	(17.81)	(0.55)	(1.16)	(0.46)	(0.51)
ST	-31.64***	10.57	-1.46***	-0.19	-3.47***	0.44
	(8.01)	(22.18)	(0.51)	(1.64)	(0.47)	(0.75)
OBC	-14.92**	0.24	-1.03*	-0.37	-1.98***	-0.35
	(6.10)	(12.18)	(0.56)	(0.77)	(0.40)	(0.30)
Mean	-40.2	-40.2	-5.8	-5.8	-5.8	-5.8
R Square	0.07	0.58	0.03	0.65	0.65	0.90
Weighted N	52244	52244	52244	52244	52244	52244
Real N	184	184	184	184	184	184

The results are weighted by numbers of individuals in each state-caste-urban cell and clustered at state-urban level. The covariates used include endowment, private investment, and public health variables. The state-urban fixed effects are included in the public health variables. Age Cutoff = 24 Months. p-values: * 0.10, ** 0.05, *** 0.01.

The striking similarity of results over the two waves indicates an amount of temporal stability in the child growth dynamics we observed in data from both 2015 and 2019, even when applying a de facto pre-defined regression model to the newer dataset. This eases concerns related to cherry picking results or cherry picking regression model specifications. This temporal stability exists over and above the similarities in dynamics across caste groups within each wave. Both of these layers of stability - within and between survey waves - lend credence to our argument that the differential growth processes we identified in the 2015 data are meaningful beyond just a particular sample at a particular moment in time.

8 Conclusion

HAZ gaps in children born across various castes in India are significant. The caste HAZ gap dynamics that we estimate indicate a disparity that is present at birth and increasing over the first few years of life. Health capital theory provides us an intuitive way to interpret this result: if social, economic and health disparities exist prior to birth, these differences will be reflected in differences in the health endowments of children (Currie and Almond, 2011). These endowment effects persist in children, but as children age they are exacerbated by disparities in the stream of child inputs consumed after birth.

In this paper, our age-dynamic estimates provide a number of insights into the nature of caste HAZ differentials. First, in line with previous literature, we find the gaps are both large and the explained share of variation is high. Coffey et al. (2019), LoPalo et al. (2019), Van de Poel and Speybroeck (2009) and Ramachandran and Deshpande (2021) all document caste HAZ gaps, and all similarly find that observables explain a large share of the differentials across caste groups.

Second, our framework reveals new features of these caste gaps. We document that the caste gaps are present at birth and grow in size, particularly over the first two years of life.

Our decomposition results suggest that both health-endowment variables and (private) health-investment variables matter significantly in explaining child development, in a manner that changes as children age. Endowment related variables largely explain birth length HAZ gaps across castes, and private investments become increasingly important to explaining those gaps as children age. These patterns are remarkably similar in the fourth and fifth waves of the NFHS that we analyze.

While previous literature documents the HAZ differentials across age for different socioeconomic groups (Rieger and Trommlerová, 2016) and estimates the associations between household variables and child HAZ for children under and over the age of 2 (Alderman and Headey, 2018; Coffey et al., 2019; LoPalo et al., 2019; Ramachandran and Deshpande, 2021), ours is the first study that provides the relative contribution of child health inputs across age. Additionally, we use this empirical framework in the context of caste disparities in child HAZ and highlight the biological processes leading to the observed caste gaps. In doing so, we contribute to the literature that in its current state focuses on the social discriminatory mechanisms that generate these disparities (Van de Poel and Speybroeck, 2009; Coffey et al., 2019; Ramachandran and Deshpande, 2021).

Our findings of the persistent effects of endowment related variables as children age imply that historical factors affecting maternal physiology are likely to be nearly as important in generating caste HAZ gaps as contemporary factors. Simultaneously, the increasing importance of post-birth child investments as children age indicates that lower caste children face real disadvantages that disproportionately and negatively affect their growth and development even today. And while our results are consistent with both discriminatory and non-discriminatory, and both contemporary and historical factors, they also cast doubt on the ability of public policy to remediate HAZ caste gaps in the short-term. Health endowments don't change after birth, and remedying the health endowment disparities across castes is likely to be a multi-generational process.

References

Addo, O. Y., Stein, A. D., Fall, C. H., Gigante, D. P., Guntupalli, A. M., Horta, B. L., Kuzawa, C. W., Lee, N., Norris, S. A., Prabhakaran, P., et al. (2013). Maternal height and child growth patterns. *The Journal of pediatrics*, 163(2):549–554.

Ahmed, S. and Ray, R. (2018). Do in utero shocks have adverse effects on child health outcomes and can welfare schemes ameliorate such effects? evidence from andhra pradesh, india. *Journal of Biosocial Science*, 50(6):770–799.

Aiyar, A. and Cummins, J. R. (2021). An age profile perspective on two puzzles in global child health: The Indian Enigma & economic growth. *Journal of Development Economics*, 148:102569.

Aizer, A. and Currie, J. (2014). The intergenerational transmission of inequality: Maternal disadvantage and health at birth. *science*, 344(6186):856–861.

Alderman, H. and Headey, D. (2018). The timing of growth faltering has important implications for observational analyses of the underlying determinants of nutrition outcomes. *PloS one*, 13(4):e0195904.

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

Attanasio, O., Meghir, C., and Nix, E. (2020). Human capital development and parental investment in india. *The Review of Economic Studies*, 87(6):2511–2541.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70(5, Part 2):9–49.

Bhalotra, S., Valente, C., and Van Soest, A. (2010). The puzzle of muslim advantage in child survival in india. *Journal of Health Economics*, 29(2):191–204.

Bhalotra, S. R., Delavande, A., Gilabert, P., and Maselko, J. (2020). Maternal investments in children: the role of expected effort and returns.

Blunch, N.-H. and Gupta, N. D. (2020). Mothers' health knowledge gap for children with diarrhea: A decomposition analysis across caste and religion in india. *World Development*, 126:104718.

Borooah, V. K. (2005). Caste, inequality, and poverty in india. *Review of development economics*, 9(3):399–414.

Chakrabarti, S., Scott, S. P., Alderman, H., Menon, P., and Gilligan, D. O. (2021). Intergenerational nutrition benefits of india's national school feeding program. *Nature Communications*, 12(1):1–10.

Chari, A., Heath, R., Maertens, A., and Fatima, F. (2017). The causal effect of maternal age at marriage on child wellbeing: Evidence from india. *Journal of Development Economics*, 127:42–55.

Coffey, D., Deshpande, A., Hammer, J., and Spears, D. (2019). Local social inequality, economic inequality, and disparities in child height in india. *Demography*, 56(4):1427–1452.

Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *JoUrnal of EconomIc LiteratUre*, 47(1):87–122.

Currie, J. and Almond, D. (2011). Human capital development before age five. In *Handbook of Labor Economics*, volume 4, pages 1315–1486. Elsevier.

Danaei, G., Andrews, K. G., Sudfeld, C. R., Fink, G., McCoy, D. C., Peet, E., Sania, A., Smith Fawzi, M. C., Ezzati, M., and Fawzi, W. W. (2016). Risk factors for childhood stunting in 137 developing countries: a comparative risk assessment analysis at global, regional, and country levels. *PLoS medicine*, 13(11):e1002164.

Daysal, N. M., Trandafir, M., and Van Ewijk, R. (2015). Saving lives at birth: The impact of home births on infant outcomes. *American Economic Journal: Applied Economics*, 7(3):28–50.

De Onis, M. and Branca, F. (2016). Childhood stunting: a global perspective. *Maternal & child nutrition*, 12:12–26.

Debnath, S. and Jain, T. (2020). Social connections and tertiary health-care utilization. *Health Economics*, 29(4):464–474.

Deshpande, A. (2000). Does caste still define disparity? a look at inequality in kerala, india. *American Economic Review*, 90(2):322–325.

Deshpande, A. (2001). Caste at birth? redefining disparity in india. *Review of Development Economics*, 5(1):130–144.

Deshpande, A. (2007). Overlapping identities under liberalization: Gender and caste in india. *Economic Development and Cultural Change*, 55(4):735–760.

Deshpande, A. and Newman, K. (2007). Where the path leads: The role of caste in post-university employment expectations. *Economic and Political Weekly*, pages 4133–4140.

Deshpande, A. and Ramachandran, R. (2019). Traditional hierarchies and affirmative action in a globalizing economy: Evidence from india. *World Development*, 118:63–78.

Deshpande, A. and Sharma, S. (2016). Disadvantage and discrimination in self-employment: Caste gaps in earnings in indian small businesses. *Small Business Economics*, 46:325–346.

Gelbach, J. B. (2016). When do covariates matter? and which ones, and how much? *Journal of Labor Economics*, 34(2):509–543.

Geruso, M. and Spears, D. (2018). Neighborhood sanitation and infant mortality. *American Economic Journal: Applied Economics*, 10(2):125–62.

Godlonton, S. and Okeke, E. N. (2016). Does a ban on informal health providers save lives? evidence from malawi. *Journal of Development Economics*, 118:112–132.

Goraya, S. S. (2023). How does caste affect entrepreneurship? birth versus worth. *Journal of Monetary Economics*.

Grossman, M. (1972). On the concept of health capital and the demand for health, 80 j. *Pol. Econ*, 223(10.2307):1830580223.

Ito, T. (2009). Caste discrimination and transaction costs in the labor market: Evidence from rural north india. *Journal of development Economics*, 88(2):292–300.

Kijima, Y. (2006). Caste and tribe inequality: evidence from india, 1983–1999. *Economic Development and Cultural Change*, 54(2):369–404.

Kumar, S. M. (2013). Does access to formal agricultural credit depend on caste? *World Development*, 43:315–328.

LoPalo, M., Coffey, D., and Spears, D. (2019). The consequences of social inequality for the health and development of india's children: the case of caste, sanitation, and child height. *Social Justice Research*, 32(3):239–254.

Maertens, A. (2013). Social norms and aspirations: age of marriage and education in rural india. *World Development*, 47:1–15.

Maity, B. (2017). Comparing health outcomes across scheduled tribes and castes in india. *World Development*, 96:163–181.

Munshi, K. (2019). Caste and the indian economy. *Journal of Economic Literature*, 57(4):781–834.

Munshi, K. and Rosenzweig, M. (2006). Traditional institutions meet the modern world: Caste, gender, and schooling choice in a globalizing economy. *American Economic Review*, 96(4):1225–1252.

Özaltin, E., Hill, K., and Subramanian, S. (2010). Association of maternal stature with offspring mortality, underweight, and stunting in low-to middle-income countries. *Jama*, 303(15):1507–1516.

Perkins, J. M., Khan, K. T., Smith, G. D., and Subramanian, S. (2011). Patterns and trends of adult height in india in 2005–2006. *Economics & Human Biology*, 9(2):184–193.

Prendergast, A. J. and Humphrey, J. H. (2014). The stunting syndrome in developing countries. *Paediatrics and international child health*, 34(4):250–265.

Ramachandran, R. and Deshpande, A. (2021). The impact of caste: A missing link in the literature on stunting in india.

Rieger, M. and Trommlerová, S. K. (2016). Age-specific correlates of child growth. *Demography*, 53(1):241–267.

Roth, D. E., Krishna, A., Leung, M., Shi, J., Bassani, D. G., and Barros, A. J. (2017). Early childhood linear growth faltering in low-income and middle-income countries as a whole-population condition: analysis of 179 demographic and health surveys from 64 countries (1993–2015). *The Lancet Global Health*, 5(12):e1249–e1257.

Siddique, Z. (2011). Evidence on caste based discrimination. *Labour Economics*, 18:S146–S159.

Spears, D. (2020). Exposure to open defecation can account for the indian enigma of child height. *Journal of Development Economics*, 146:102277.

Spears, D. and Thorat, A. (2019). The puzzle of open defecation in rural india: evidence from a novel measure of caste attitudes in a nationally representative survey. *Economic Development and Cultural Change*, 67(4):725–755.

Subramanian, S. V., Nandy, S., Irving, M., Gordon, D., Lambert, H., and Davey Smith, G. (2006). The mortality divide in india: the differential contributions of gender, caste, and standard of living across the life course. *American Journal of Public Health*, 96(5):818–825.

Swaminathan, H., Sharma, A., and Shah, N. G. (2019). Does the relationship between income and child health differ across income groups? evidence from india. *Economic Modelling*, 79:57–73.

Van de Poel, E. and Speybroeck, N. (2009). Decomposing malnutrition inequalities between scheduled castes and tribes and the remaining indian population. *Ethnicity & Health*, 14(3):271–287.

Victora, C. G., de Onis, M., Hallal, P. C., Blössner, M., and Shrimpton, R. (2010). Worldwide timing of growth faltering: revisiting implications for interventions. *Pediatrics*, pages peds–2009.

Von Grafenstein, L., Klasen, S., and Hoddinott, J. (2023). The indian enigma revisited. *Economics & Human Biology*, 49:101237.

Vyas, S., Hathi, P., and Gupta, A. (2022). Social disadvantage, economic inequality, and life expectancy in nine indian states. *Proceedings of the National Academy of Sciences*, 119(10):e2109226119.

WHO (2006). *WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development*. World Health Organization.

Zacharias, A. and Vakulabharanam, V. (2011). Caste stratification and wealth inequality in india. *World Development*, 39(10):1820–1833.

Appendix

Detailed Summary Statistics

Table A.1: Summary Statistics 2

	(1) SC mean/sd	(2) ST mean/sd	(3) OBC mean/sd	(4) UC mean/sd	(5) All mean/sd
HAZ (x 100)	-166.79	-171.84	-148.69	-112.61	-149.29
	(164.41)	(169.98)	(166.18)	(160.43)	(166.28)
Endowment Variables					
Birth Order	2.29	2.29	2.12	1.86	2.14
	(1.46)	(1.43)	(1.31)	(1.04)	(1.33)
Percent BO = 1	0.36	0.36	0.38	0.45	0.39
	(0.48)	(0.48)	(0.49)	(0.50)	(0.49)
Percent BO = 2	0.31	0.31	0.34	0.36	0.33
	(0.46)	(0.46)	(0.47)	(0.48)	(0.47)
Percent BO (2-6)	0.31	0.32	0.27	0.19	0.27
	(0.46)	(0.47)	(0.44)	(0.39)	(0.44)
Percent BO > 6	0.02	0.02	0.01	0.00	0.01
	(0.14)	(0.13)	(0.11)	(0.06)	(0.11)
Mother's Age at Child's Birth	24.21	24.03	24.21	24.44	24.23
	(4.78)	(4.95)	(4.53)	(4.52)	(4.64)
Mother's Age Squared	609.21	601.84	606.47	617.97	608.72
	(256.43)	(266.47)	(241.90)	(238.78)	(247.96)
Mother's HAZ (x 100)	-217.65	-211.85	-201.27	-182.60	-203.18
	(93.00)	(89.77)	(95.68)	(94.38)	(94.87)
Mother's WAZ (x 100)	-113.47	-142.25	-102.61	-68.64	-103.63
	(111.92)	(99.46)	(114.19)	(121.94)	(115.34)
Delivery Care	0.51	0.47	0.56	0.71	0.56
	(0.50)	(0.50)	(0.50)	(0.45)	(0.50)

Private Investment Variables					
Poorest	0.33	0.50	0.23	0.08	0.26
	(0.47)	(0.50)	(0.42)	(0.28)	(0.44)
Poorer	0.26	0.25	0.22	0.16	0.22
	(0.44)	(0.43)	(0.41)	(0.36)	(0.41)
Middle	0.20	0.14	0.21	0.21	0.20
	(0.40)	(0.35)	(0.41)	(0.40)	(0.40)
Richer	0.14	0.07	0.20	0.25	0.18
	(0.35)	(0.25)	(0.40)	(0.43)	(0.38)
Richest	0.08	0.04	0.14	0.30	0.14
	(0.26)	(0.19)	(0.35)	(0.46)	(0.35)
Maternal Educ at least Primary	0.65	0.55	0.71	0.89	0.71
	(0.48)	(0.50)	(0.45)	(0.31)	(0.45)
Percent No Educ (mother)	0.35	0.45	0.29	0.11	0.29
	(0.48)	(0.50)	(0.45)	(0.31)	(0.45)
Percent Primary Educ (mother)	0.16	0.16	0.13	0.09	0.13
	(0.37)	(0.37)	(0.33)	(0.29)	(0.34)
Percent Secondary Educ (mother)	0.42	0.35	0.47	0.58	0.47
	(0.49)	(0.48)	(0.50)	(0.49)	(0.50)
Percent Higher Educ (mother)	0.07	0.04	0.11	0.22	0.11
	(0.25)	(0.19)	(0.32)	(0.41)	(0.31)
Owns Motorcycle	0.29	0.27	0.44	0.54	0.40
	(0.45)	(0.44)	(0.50)	(0.50)	(0.49)
Owns Agricultural Land	0.33	0.55	0.48	0.49	0.45
	(0.47)	(0.50)	(0.50)	(0.50)	(0.50)
Treat Drinking Water	0.22	0.38	0.31	0.44	0.32
	(0.41)	(0.49)	(0.46)	(0.50)	(0.47)
No Toilet Facility	0.60	0.74	0.51	0.25	0.51
	(0.49)	(0.44)	(0.50)	(0.43)	(0.50)
Flush Toilet	0.34	0.20	0.44	0.66	0.43
	(0.47)	(0.40)	(0.50)	(0.47)	(0.49)
Pit Toilet	0.06	0.06	0.05	0.08	0.06
	(0.24)	(0.23)	(0.21)	(0.27)	(0.24)

Postnatal Care-Health Worker	0.03	0.03	0.03	0.02	0.03
	(0.16)	(0.18)	(0.16)	(0.15)	(0.16)
Postnatal Care-Health Facility	0.01	0.01	0.01	0.01	0.01
	(0.11)	(0.12)	(0.11)	(0.11)	(0.11)
Completed Polio Vaccination	0.92	0.88	0.92	0.93	0.92
	(0.27)	(0.32)	(0.27)	(0.25)	(0.27)
Completed DPT Vaccinations	0.87	0.83	0.88	0.89	0.87
	(0.34)	(0.38)	(0.33)	(0.31)	(0.33)
Completed Hepatitis B Vaccination	0.83	0.80	0.84	0.87	0.84
	(0.37)	(0.40)	(0.37)	(0.34)	(0.37)
Completed Measles Vaccinations	0.72	0.69	0.74	0.76	0.73
	(0.45)	(0.46)	(0.44)	(0.42)	(0.44)
Received Vitamin A	0.68	0.67	0.69	0.72	0.69
	(0.47)	(0.47)	(0.46)	(0.45)	(0.46)
Completed BCG Vaccination	0.91	0.88	0.92	0.93	0.91
	(0.28)	(0.33)	(0.28)	(0.26)	(0.28)
Public Health Variables					
PSU has sewer system access	0.10	0.07	0.12	0.23	0.13
	(0.31)	(0.26)	(0.33)	(0.42)	(0.34)
Distance to Health Facility as Barrier	0.36	0.44	0.33	0.25	0.33
	(0.48)	(0.50)	(0.47)	(0.43)	(0.47)
Urban = 1	0.23	0.11	0.25	0.39	0.26
	(0.42)	(0.32)	(0.43)	(0.49)	(0.44)
N	34913	21441	64206	26218	146778

Unconditional Effects

Table A.2: Unconditional Caste Differences in HAZ

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	0-6	6-12	12-18	18-24	24-30	30-36	36-42	42-48	48-54	54-59
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
SC	-36.65***	-53.61***	-66.41***	-57.34***	-54.34***	-62.02***	-55.22***	-50.55***	-64.45***	-49.10***
	(7.94)	(6.68)	(6.48)	(6.08)	(7.06)	(5.84)	(4.92)	(5.09)	(6.73)	(4.68)
ST	-51.16***	-61.61***	-62.73***	-67.36***	-58.96***	-75.47***	-55.96***	-58.51***	-65.15***	-59.40***
	(9.00)	(7.67)	(7.69)	(7.05)	(7.05)	(6.74)	(5.94)	(6.24)	(7.54)	(5.46)
OBC	-20.86***	-40.08***	-43.90***	-36.52***	-36.30***	-37.60***	-30.77***	-33.49***	-50.29***	-33.46***
	(7.23)	(6.24)	(5.66)	(5.71)	(6.00)	(5.03)	(4.85)	(4.46)	(6.47)	(4.13)
Observations	11811	15053	14470	14933	14493	14917	15546	15641	14705	15209

p-values: * 0.10, ** 0.05, *** 0.01

Age is measured in months and one age group is 6 months.

Conditional Effects

Table A.3: Conditional Caste Differences in HAZ

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	0-6	6-12	12-18	18-24	24-30	30-36	36-42	42-48	48-54	54-60
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
SC	-0.65	-15.44**	-31.20***	-14.19**	-15.24**	-19.32***	-15.68***	-11.55**	-23.67***	-10.31**
	(8.02)	(7.06)	(6.78)	(6.24)	(7.06)	(6.03)	(4.86)	(4.88)	(4.95)	(4.24)
ST	-10.15	-18.67**	-22.60***	-21.39***	-13.86*	-31.36***	-5.73	-4.71	-18.55***	-17.02***
	(9.34)	(8.21)	(8.42)	(7.52)	(7.89)	(7.44)	(6.35)	(6.24)	(6.63)	(5.65)
OBC	3.28	-13.99**	-20.32***	-7.43	-9.20	-11.44**	-4.14	-6.03	-18.75***	-6.15
	(7.01)	(6.59)	(5.83)	(5.70)	(5.97)	(5.16)	(4.66)	(4.38)	(4.76)	(3.87)
Percent BO = 1	0.00	0.00	0.00	56.53***	0.00	0.00	42.77***	0.00	55.51***	48.48***
	(.)	(.)	(.)	(15.88)	(.)	(.)	(13.94)	(.)	(12.08)	(10.59)
Percent BO = 2	12.07**	-14.35***	-17.83***	46.23***	-17.97***	-12.08***	23.75*	-15.12***	38.14***	32.96***
	(5.88)	(5.33)	(4.72)	(15.58)	(4.80)	(3.89)	(13.75)	(3.77)	(12.14)	(10.43)
Percent BO (2-6)	7.13	-24.73***	-18.23***	30.82**	-30.05***	-24.66***	13.34	-21.42***	23.99**	23.71**
	(7.16)	(6.70)	(5.82)	(14.90)	(5.98)	(5.04)	(13.33)	(4.78)	(11.17)	(9.94)
Percent BO > 6	0.95	-29.02*	-52.55***	0.00	-57.73***	-53.48***	0.00	-30.08**	0.00	0.00
	(21.05)	(17.57)	(15.49)	(.)	(14.79)	(16.42)	(.)	(12.97)	(.)	(.)
Mother's Age at Child's Birth	8.46**	17.72***	2.00	6.38	7.42**	10.18***	10.10***	8.37***	5.42**	8.89***
	(4.31)	(3.56)	(3.19)	(4.18)	(3.11)	(2.58)	(2.37)	(2.63)	(2.29)	(2.21)
Mother's Age Squared	-0.12	-0.30***	-0.00	-0.08	-0.10*	-0.14***	-0.15***	-0.12**	-0.06	-0.13***
	(0.08)	(0.06)	(0.06)	(0.08)	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)
Mother's HAZ (x 100)	0.23***	0.31***	0.30***	0.30***	0.34***	0.31***	0.32***	0.33***	0.28***	0.28***
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Mother's WAZ (x 100)	0.16***	0.12***	0.13***	0.13***	0.08***	0.10***	0.10***	0.12***	0.12***	0.12***
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Delivery Care	2.42	5.20	12.45***	6.51*	5.12	6.32*	2.31	3.47	3.10	-2.68
	(4.76)	(4.40)	(4.10)	(3.65)	(4.76)	(3.52)	(3.24)	(3.24)	(3.18)	(3.07)
Poorest	0.00	-30.84**	0.00	-31.64**	0.00	-25.80***	-39.12***	0.00	0.00	0.00

	(.)	(13.28)	(.)	(16.03)	(.)	(9.97)	(9.19)	(.)	(.)	(.)
Poorer	1.98	-26.76**	7.14	-26.73**	3.41	-20.47**	-32.02***	13.44***	2.22	5.93
	(6.23)	(11.94)	(5.66)	(12.63)	(5.74)	(8.96)	(8.20)	(4.09)	(4.03)	(3.83)
Middle	15.98**	-15.88	9.52	-18.96*	24.27***	-12.07	-18.85***	17.86***	4.47	18.88***
	(8.14)	(10.16)	(7.19)	(9.72)	(7.35)	(8.05)	(7.21)	(5.08)	(5.25)	(4.88)
Richer	25.48**	-12.13	4.28	-22.77**	35.96***	-5.98	-17.53**	29.99***	9.69	25.22***
	(11.79)	(9.30)	(9.42)	(9.16)	(9.05)	(6.91)	(6.96)	(6.42)	(6.97)	(6.61)
Richest	39.44***	0.00	5.21	0.00	54.70***	0.00	0.00	42.21***	30.19***	36.62***
	(14.02)	(.)	(11.56)	(.)	(11.88)	(.)	(.)	(9.03)	(9.12)	(8.10)
Percent No Educ (mother)	-15.89	-5.12	-2.57	-10.22*	-27.40***	-11.35**	-33.37***	-38.23***	-21.42***	-8.38**
	(9.88)	(6.47)	(5.69)	(5.61)	(8.61)	(4.97)	(8.03)	(6.98)	(8.12)	(3.96)
Percent Primary Educ (mother)	-7.03	0.00	0.00	0.00	-20.16**	0.00	-29.91***	-24.46***	-22.89***	0.00
	(10.65)	(.)	(.)	(.)	(8.71)	(.)	(8.39)	(7.10)	(8.33)	(.)
Percent Secondary Educ (mother)	-9.76	-1.64	24.36***	14.54***	-14.24*	9.30*	-18.71**	-16.97***	-11.90	8.67**
	(8.28)	(6.31)	(5.72)	(5.17)	(7.74)	(4.75)	(7.57)	(6.15)	(7.41)	(3.91)
Percent Higher Educ (mother)	0.00	3.03	33.77***	28.46***	0.00	24.95***	0.00	0.00	0.00	14.27**
	(.)	(9.65)	(8.74)	(8.62)	(.)	(7.90)	(.)	(.)	(.)	(6.04)
Owns Motorcycle	-11.14*	1.50	9.70**	10.32**	4.12	9.81**	0.67	7.27*	2.70	-1.60
	(5.77)	(5.75)	(4.92)	(4.92)	(5.02)	(4.27)	(3.97)	(3.84)	(4.26)	(3.40)
Owns Agricultural Land	14.40***	3.88	-0.30	-3.63	-0.25	-1.57	-2.51	1.08	1.72	2.76
	(4.72)	(4.31)	(3.93)	(3.81)	(4.15)	(3.49)	(3.23)	(3.18)	(3.51)	(2.71)
Treat Drinking Water	-2.39	-0.58	-8.50	-8.65*	9.89**	1.25	-4.48	7.85*	4.26	8.11**
	(6.01)	(5.56)	(5.37)	(4.81)	(4.93)	(4.45)	(4.07)	(4.11)	(3.87)	(3.43)
No Toilet Facility	5.18	-1.53	-15.28*	-23.18***	-0.01	-5.79	0.00	-9.23**	-9.20	7.45
	(7.14)	(8.67)	(8.97)	(8.65)	(8.86)	(4.97)	(.)	(4.15)	(6.41)	(6.15)
Flush Toilet	0.00	9.60	1.02	-14.83*	-2.96	0.00	4.63	0.00	-7.11	16.52***
	(.)	(8.62)	(8.82)	(8.07)	(9.26)	(.)	(4.47)	(.)	(6.17)	(6.13)
Pit Toilet	-10.49	0.00	0.00	0.00	0.00	1.25	2.70	4.92	0.00	0.00
	(9.85)	(.)	(.)	(.)	(.)	(6.94)	(7.16)	(6.88)	(.)	(.)
Postnatal Care-Health Worker	-1.53	2.76	13.82	-13.30	-9.13	-12.96	7.62	-8.70	-15.27	-12.90
	(8.88)	(12.27)	(11.54)	(10.81)	(11.10)	(10.81)	(11.36)	(9.62)	(11.32)	(9.63)
Postnatal Care-Health Facility	-13.71	10.06	6.98	-32.69*	14.38	33.32	15.47	8.41	-15.11	-3.30
	(11.83)	(21.60)	(21.91)	(18.60)	(15.04)	(25.58)	(21.96)	(21.36)	(19.71)	(14.56)

Completed Polio Vaccination	9.24	14.38	10.93	-26.15*	-22.48**	-14.70	-9.33	-12.11	-3.12	3.35
	(9.76)	(13.81)	(12.58)	(14.39)	(11.14)	(10.33)	(9.16)	(9.28)	(10.08)	(8.04)
Completed DPT Vaccinations	2.37	-7.68	-0.21	-3.25	-3.38	-6.95	-8.15	-0.22	-4.50	-2.37
	(5.43)	(7.73)	(10.60)	(15.22)	(11.50)	(9.77)	(9.97)	(8.47)	(9.29)	(7.29)
Completed Hepatitis B Vaccination	-2.50	-5.08	-8.93	-9.37	-7.61	11.81*	9.28	0.32	10.19**	-4.23
	(5.81)	(6.45)	(7.56)	(9.47)	(9.27)	(7.01)	(6.14)	(5.47)	(5.17)	(4.08)
Completed Measles Vaccinations	26.28**	-8.54*	3.36	13.76**	-2.44	10.21	-3.18	1.84	4.90	4.20
	(11.14)	(5.04)	(5.55)	(6.76)	(10.61)	(6.37)	(5.92)	(5.35)	(5.90)	(5.12)
Received Vitamin A	-2.76	10.11**	-4.69	-2.43	-12.08**	-14.77***	-7.67*	-5.83	-9.70***	-2.43
	(5.40)	(4.57)	(5.03)	(5.55)	(4.74)	(3.86)	(3.92)	(3.61)	(3.31)	(2.96)
Completed BCG Vaccination	-3.17	-3.23	-15.34	8.80	27.56**	2.59	3.18	11.45	-1.20	-3.05
	(9.83)	(12.12)	(11.67)	(11.72)	(10.74)	(12.00)	(8.99)	(9.39)	(10.53)	(8.76)
PSU has sewer system access	11.11	-4.38	-9.08	-3.39	3.50	-0.88	-6.92	-1.41	-8.45	0.78
	(9.03)	(8.49)	(7.93)	(7.07)	(9.62)	(6.68)	(6.60)	(6.51)	(6.98)	(5.10)
Distance to Health Facility as Barrier	5.20	2.28	3.21	0.64	1.05	1.89	0.02	0.05	-3.89	-5.15*
	(4.63)	(4.46)	(4.19)	(4.53)	(4.06)	(3.57)	(3.29)	(3.12)	(3.24)	(2.74)
Observations	11811	15053	14470	14933	14493	14917	15546	15641	14705	15209

p-values: * 0.10, ** 0.05, *** 0.01

Alpha / Beta Estimates

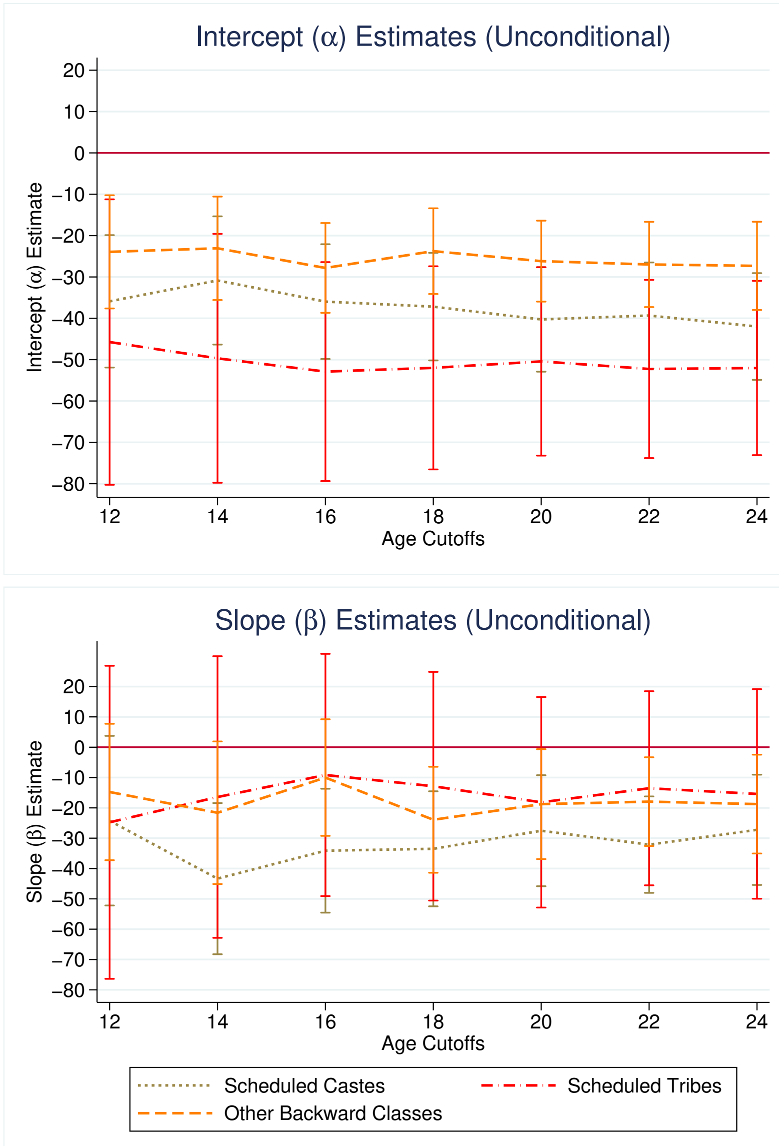

Figure A.1: α/β Estimates by Age Cutoffs

This figure shows the robustness of estimated coefficients on caste group dummies for the α and β regressions across in Equation 3 by age cutoff, along with their 95% confidence intervals. The x axis represents the age cutoff in months. The top panel presents the unconditional estimates for birth length ($\hat{\alpha}$). The bottom panel presents the unconditional estimates for the rate of loss ($\hat{\beta}$). The results are weighted by numbers of individuals in each state-caste-urban cell and clustered at state-urban level.

Oaxaca Blinder Decompositions

Table A.4: Explained in Percentage for Blinder-Oaxaca Decomposition

	All			Endowment			Private			Public		
	(1) SC mean	(2) ST mean	(3) OBC mean	(4) SC mean	(5) ST mean	(6) OBC mean	(7) SC mean	(8) ST mean	(9) OBC mean	(10) SC mean	(11) ST mean	(12) OBC mean
Child Age = 6m	1.00	0.83	1.18	59.45	52.69	66.51	24.62	8.39	2.10	16.03	22.10	49.58
Child Age = 12m	0.58	0.63	0.40	40.89	40.23	31.83	22.33	28.90	18.37	-5.55	-6.40	-10.62
Child Age = 18m	0.41	0.56	0.33	30.52	35.92	30.57	23.98	41.69	25.11	-13.68	-21.75	-22.34
Child Age = 24m	0.71	0.60	0.84	30.25	29.39	28.09	36.03	36.75	47.71	4.50	-6.54	8.25
Child Age = 30m	0.72	0.84	0.69	35.34	31.93	31.89	56.34	61.37	52.22	-19.41	-9.26	-15.32
Child Age = 36m	0.67	0.46	0.64	37.18	35.74	34.71	20.69	17.67	14.55	8.72	-7.41	15.17
Child Age = 42m	0.77	0.93	0.93	30.89	34.85	28.06	41.20	59.11	45.95	5.07	-0.83	18.90
Child Age = 48m	0.85	0.78	0.85	32.09	32.50	30.10	48.35	52.28	44.24	4.50	-7.13	10.66
Child Age = 54m	0.69	0.85	0.58	48.16	54.01	43.12	28.01	25.84	23.70	-7.37	5.51	-8.40
Child Age = 60m	0.88	0.76	0.83	28.84	26.38	28.73	55.75	47.89	47.67	3.12	1.82	6.33
Total	0.73	0.72	0.73	37.36	37.37	35.36	35.73	37.99	32.16	-0.41	-2.99	5.22

This table shows the results of Blinder-Oaxaca Decomposition in terms of explained in percentage.
Child age is measures in months.

Gelbach Decomposition

Another approach to disentangling the contribution of covariates in explaining the caste differences in HAZ is to examine variation in the caste coefficient from regressions that project HAZ on caste, with and without the covariates, following Gelbach (2016). Consider the following regressions:

Base/ Unconditional model:

$$Y_{irvg}^{a} = \delta^{ag} * Caste_{irvg}^{a} + \lambda_{r}^{a} + \epsilon_{irvg}^{a} \tag{6}$$

Full/Conditional model:

$$Y_{irvg}^{a} = \delta_{g}^{a} * Caste_{g}^{a} + X_{1irvg}^{\prime a} * \beta_{1}^{a} + X_{2v}^{\prime a} * \beta_{2}^{a} + \lambda_{r}^{a} + \epsilon_{irvg}^{a} \tag{7}$$

where the subscripts i, r, v, and g represent child, region, village, and caste group, respectively. The coefficient on the caste variable $\hat{\delta}^{ag}$ from the benchmark model is the simple difference in mean HAZ of group g with respect to UC (baseline). The decomposition suggested by Gelbach (2016), based on a formulation of the omitted variable bias formula, exploits this difference in coefficients on the $Caste_{irvg}^{a}$ dummy in the base model to decompose the fraction of the change in $\hat{\beta}$ attributable to each element of $X_{1irvg}^{\prime a}$. The contribution of a specific variable to caste gap in HAZ is the mean difference in that covariate between g and UC group scaled by the effect of that covariate on HAZ. As above, our interest centers on understanding the individual contributions of the three covariate groups related to health endowments, private investments, and public investments in child health.

Similar to the OB approach above, we estimate the difference in $\hat{\delta}^{ag}$ for each age group a and trace out the contribution of different set of covariates by age. We calculate the explained portion for a covariate group as the percent (net) contribution of the variable group to the coefficient differences between the base (Equation 6) and full model (Equation 7).

Figure A.2 presents the results from this decomposition exercise. The top left panel shows that the contribution of endowment variables ranges between 40% to 60% for all caste groups. Among STs and SCs, the share explained by endowments starts at about 45% at birth and reaches to 60% for one-year old children. For both groups this decreases over child's age— for STs, this share decreases over the child's age to around 55% by age 3 and 50% by age 5; for SCs, after the first year, the share remains constant at around 50% after the first two years. For OBC children, health endowments explain around 40% of the differences at birth and this increases slightly to around 56% over the first 18 months and then hovers between 40-50%. In the top right panel of figure 4, we see how private investments matter. 30% of the differences in coefficients can be explained by private investments at childbirth for ST and SC children. This effect accumulates to around 60% by age 3 for STs and 50% for SC. The contribution for OBC children starts off lower at birth but follows the trajectory of SCs closely at later ages.

As shown in the bottom left panel, public health variables do not seem to contribute in explaining the change in the $Caste^{a}_{irvg}$ coefficient, except for the initial bump at around 25-35% immediately after the birth. The broad patterns are consistent with the main findings in Figure 3 using the OB decomposition — endowment related variables matter early on and remain important in explaining the HAZ gap; private investment is less important for newborns but becomes important as children age; and public health variables do not seem to account for the HAZ gap among caste groups.

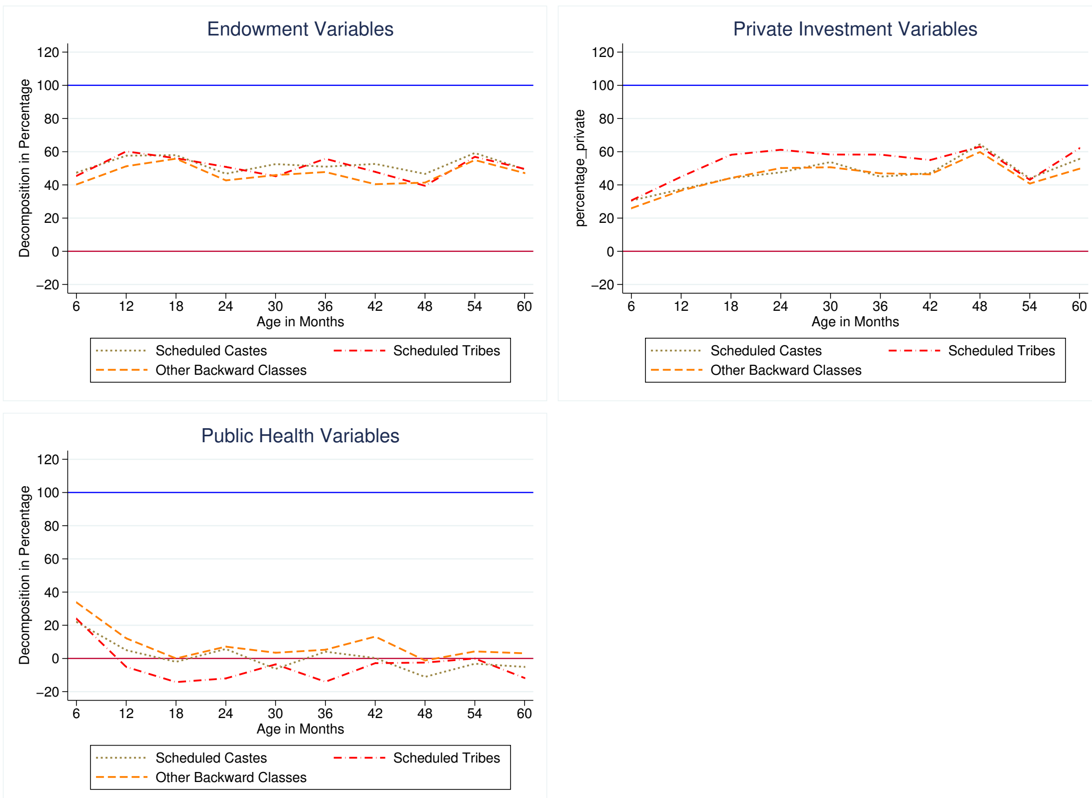

Figure A.2: Gelbach Decomposition Results: Coefficient Differences in Percentage

This is the graphical representation of the Gelbach decomposition results. The y axis is the percent explained of the HAZ difference after including all controls after dividing the coefficient by the coefficient differences in the unconditional model. Child age is measured in months on the x-axis. The top left panel shows the decomposition of coefficient difference for endowment variables in percentages. The top right panel shows the decomposition of coefficient difference for private investment variables. The bottom left panel shows the decomposition of coefficient difference for public health variables.

Table A.5: Explained in Percentage for Gelbach Decomposition

	Endowment			Private			Public		
	(1) SC mean	(2) ST mean	(3) OBC mean	(4) SC mean	(5) ST mean	(6) OBC mean	(7) SC mean	(8) ST mean	(9) OBC mean
Child Age = 6m	47.38	45.39	40.25	30.55	30.64	25.93	22.07	23.97	33.82
Child Age = 12m	57.61	60.18	51.24	37.38	44.91	36.61	5.01	-5.09	12.15
Child Age = 18m	57.93	56.08	55.80	44.09	58.15	44.17	-2.01	-14.24	0.03
Child Age = 24m	46.73	50.85	42.73	47.56	61.17	50.16	5.71	-12.02	7.12
Child Age = 30m	52.54	45.24	45.91	53.76	58.29	50.64	-6.30	-3.53	3.45
Child Age = 36m	50.97	55.76	47.77	44.94	58.27	46.98	4.09	-14.03	5.25
Child Age = 42m	52.64	47.86	40.42	47.09	54.98	46.37	0.27	-2.84	13.20
Child Age = 48m	46.62	39.46	41.38	64.47	63.02	59.82	-11.09	-2.48	-1.19
Child Age = 54m	59.23	56.93	54.97	43.94	43.09	40.78	-3.17	-0.01	4.25
Child Age = 60m	49.39	49.63	47.13	55.68	62.14	49.75	-5.07	-11.77	3.11
Total	52.10	50.74	46.76	46.95	53.46	45.12	0.95	-4.20	8.12

This table shows the results of Gelbach Decomposition in terms of explained in percentage.
Child age is measures in months.

HAZ differences between Muslim & UC children

Figure A.3 summarizes the mean HAZ across child age in months for Muslim children and upper caste children from the NFHS-4 data. Muslim children are, on average, born smaller than upper caste children and grow at a slower rate. In the first 6 months, UC children seem to recover rapidly.

In Figure A.4, the top and bottom panel shows regression estimates from the unconditional and conditional regression models. Each point estimate in Figure A.4 is the coefficient of Muslim children relative to upper caste children (UC). In the unconditional model, Muslim children are around 0.2 sd smaller than UC children at birth. These differences increase to around 0.5 sd by age 5. In the bottom panel, religion gap estimates lower after controlling for covariates related to the health endowment, private investment, and public investment. There are no differences in the first 6 months and these differences increase to 0.3 sd in the first 12 months. These changes support the issue of selective mortality that UC children face as discussed in Bhalotra et al. (2020). By age 2, Muslim children are around 0.2 sd shorter and this gap remains a constant over the rest of their age.

Figure A.5 shows the Blinder-Oaxaca decomposition results explained in percentage as described in Equation 4 and Equation 5. In the top left panel, we see that more than 50% of the religion gaps in birth length can be explained by covariates. One third of the height differences between Muslim children and UC children can be explained by their endowment covariates. These shares are decreasing over the child's age. In the bottom left panel, around 80% of the differences at birth can be explained by differences in private investment variables. These differences decrease by 12 months and then increase by age. By age 5, around two thirds of the differences can be explained by differences in private investment variables across groups. Public health variables explain almost no share of the differences.

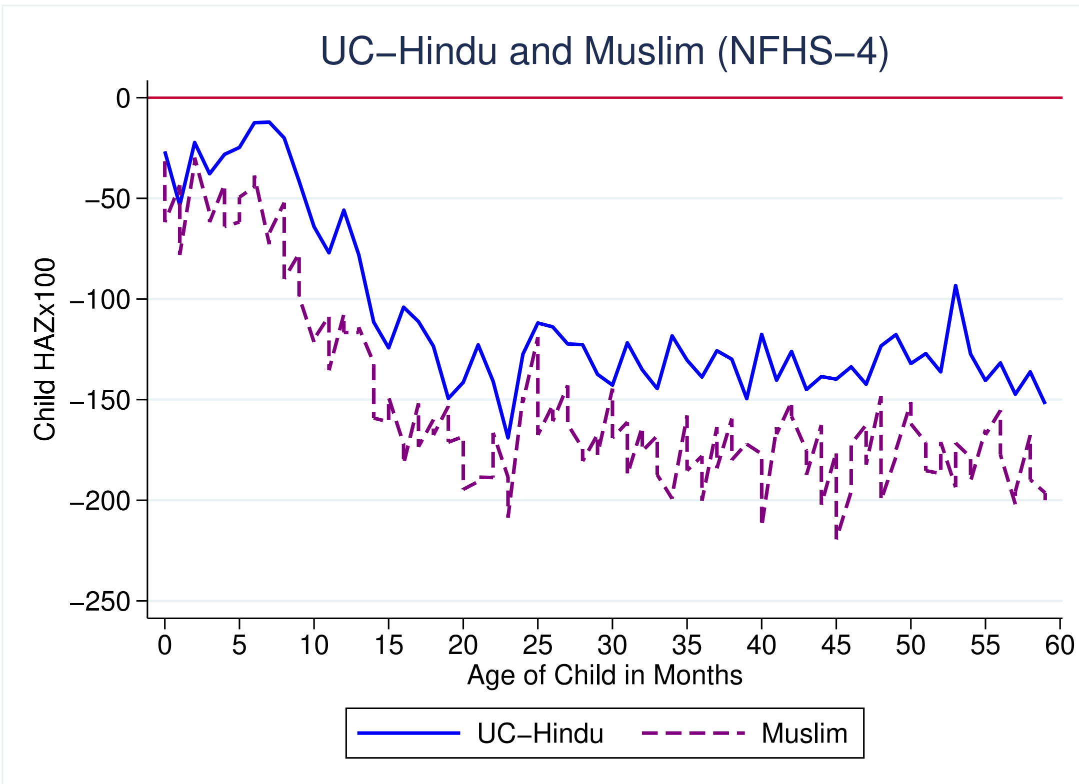

Figure A.3: Child HAZ for UC Hindu and Muslim

This figure presents child HAZ score by religion (UC-Hindus and Muslims). The x-axis represents age in months and the y-axis is the weighted mean HAZ. The results are weighted by survey weights.

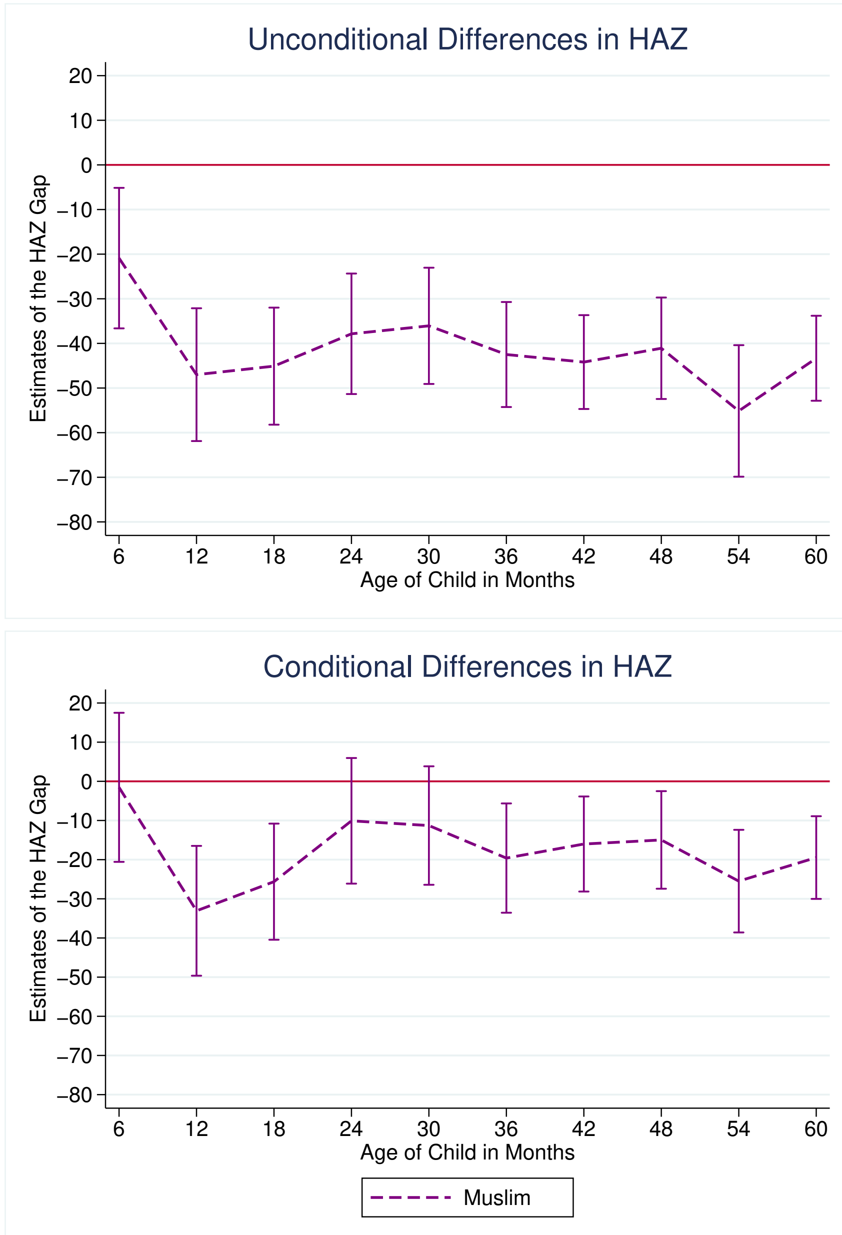

Figure A.4: Regression Estimates (UC Hindu vs. Muslims)

This figure presents regression estimates for both unconditional and conditional regressions in Equation 1, with non-UC Hindus exlcuded and Muslim children comprising the group of interest. The y-axis is the coefficient estimates (and95% confidence intervals) on the caste-group variable, and the x-axis separates the estimates into 6-month age-bins. The top panel presents regression estimates from the unconditional model. In the bottom panel, the regression (conditional) estimates include all controls related to endowment, private investment, and public health. The results are weighted by survey weights and clustered at PSU level.

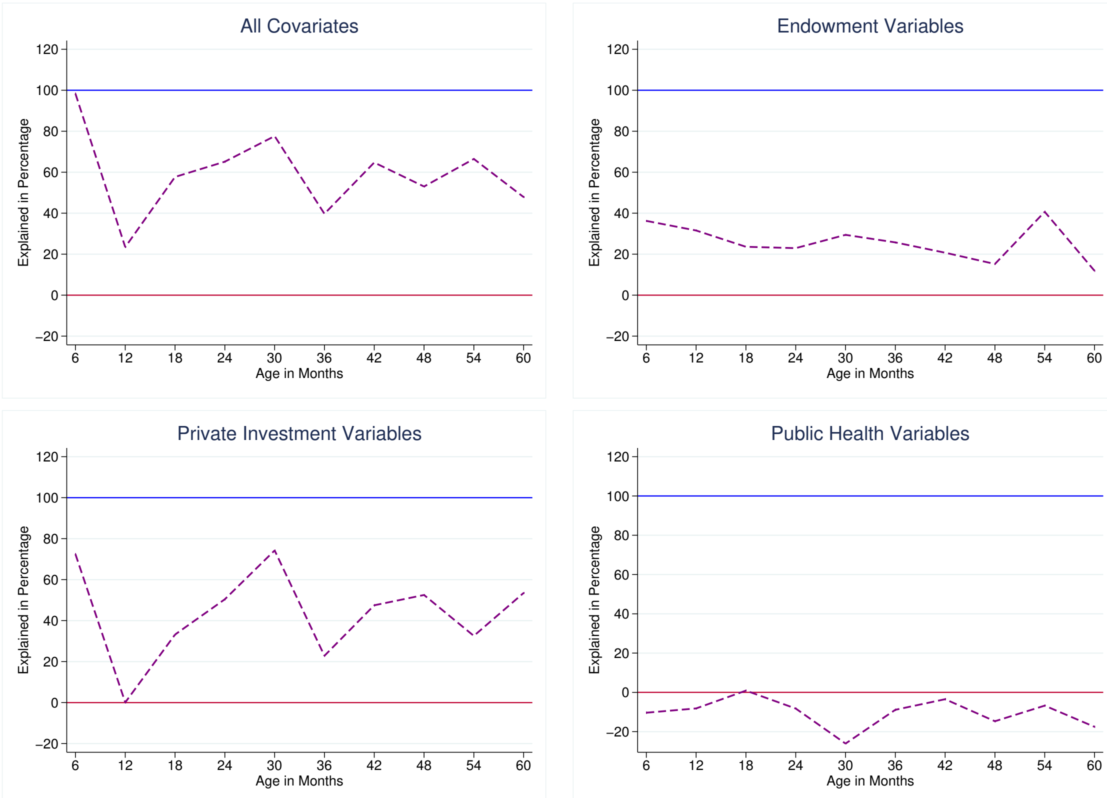

Figure A.5: Blinder-Oaxaca Decomposition Results: Explained in Percentage (UC Hindu vs. Muslims)

This figure presents Blinder-Oaxaca decomposition results for Muslim children relative to UC-Hindu children. Child age measured in 6-month age groups is on the x-axis and the y-axis is the percent of the religion gap that can be explained by model covariates. The top left panel presents the percent of the HAZ gap that is explained when all control variables are included. The top right panel is the the percent of the HAZ gap explained by endowment variables. The bottom left panel depicts the percent of the HAZ gap explained by the private investment variables. The bottom right panel presents the percent HAZ gap that is explained in percentage by public health variables.

HAZ differences between states with low and high share UC groups

Figure A.6 is the histogram of the share of upper caste households across states in India as interviewed in the NFHS-4. The median share of UC households across the country is around 15%. Using this figure, we divide the sample into states with a high share of upper caste ($\geq$15%) and low share of upper caste (<15%).

Figure A.7 shows the Blinder-Oaxaca decomposition results explained in percentage for children living in regions with high shares of UC households. In the top left panel, we see that more than 50% of the share gaps among regions with high share of upper caste can be explained by covariates. Nearly 80% of the height differences at birth can be explained by endowment covariates when low caste children live in areas with high shares of upper caste (top-right panel). These shares are decreasing over the child's age. Endowment differences explain around 40% of differences after age 2 among SC and ST children. Among OBC children, endowment differences explain less than 20% after age 2. In the bottom left panel, less than 40% of the differences at birth can be explained by differences in private investment variables. These differences increase by 12 months and then increase, inconsistently, by age.By age 3, close to 50% of the differences can be explained by private investment variables. This reduces to around 20% among ST and ST children by age 5 but remains high for OBC children (around 70%). In the bottom right panel, we see that public health variables explain almost no share of the differences among SC and ST children in regions with high shares of UC groups. Only among OBC children do public health variables matter. Around 40% of the differences in health can be explained by these variables.

Figure A.8 shows the Blinder-Oaxaca decomposition results for children living in areas with low shares of UC households. In the top left panel, we see that more than 80% of the share gaps among lower and upper caste children can be explained by covariates. This briefly decreases till age 3 after which all the differences can be explained by covariates in the model. Around 40% of the height differences at birth can be explained by endowment covariates when low caste children live in areas with lower shares of upper caste (top-right panel). These shares are slightly decreasing over the child's age. Endowment differences explain around 30% of differences by age 5. In the bottom left panel, less than 20% of the differences at birth can be explained by differences in private investment variables for ST children. Among SC and OBC children close to 60% of the differences can be explained by private investments. These differences increase and by age 5 explain around around 70% of the differences. In areas with low shares of UC households, public health variables explain almost no share of the differences after birth (bottom-right panel).

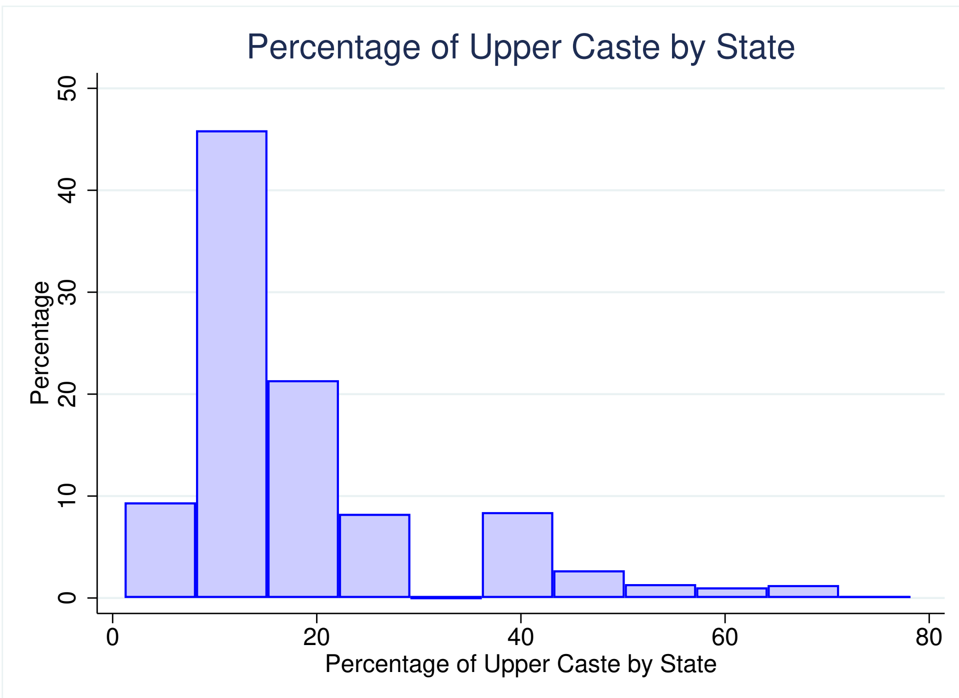

Figure A.6: Percentage of Upper Caste by State

This figure presents a histogram of the share of Upper Caste by State. Observations are weighted by sample weights.

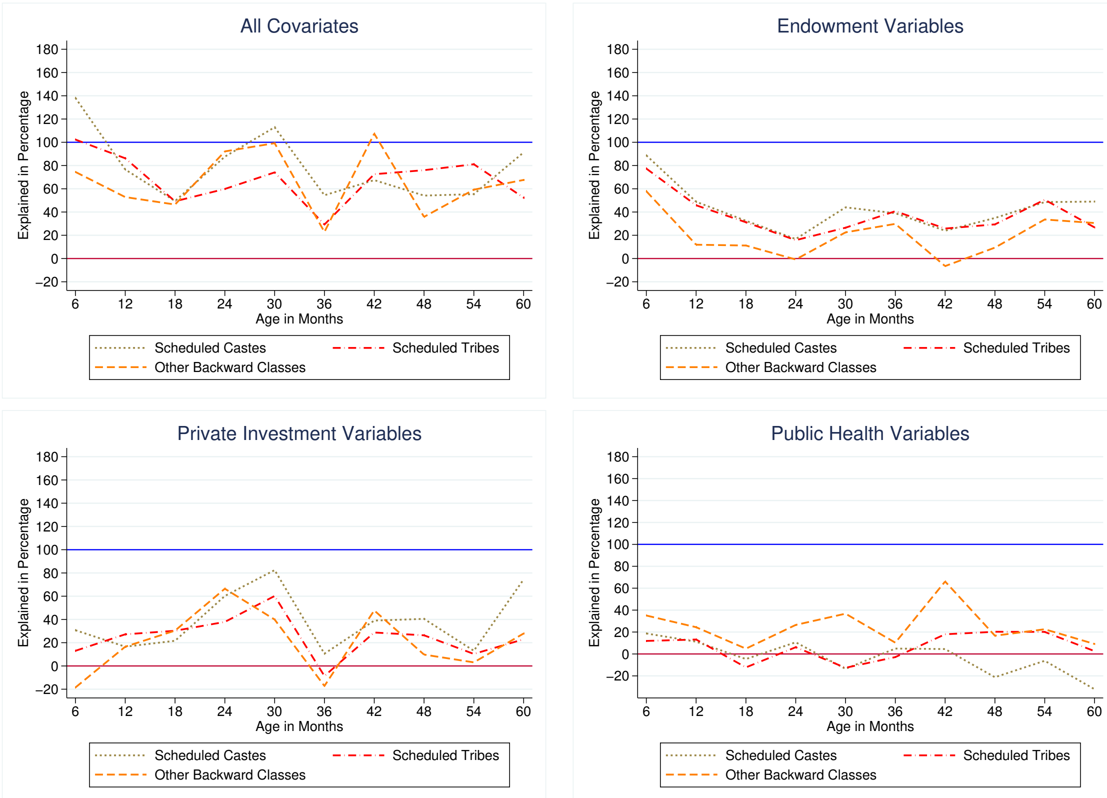

Figure A.7: Blinder-Oaxaca Decomposition Results: Explained in Percentage (High Share of Upper Caste by State)

This figure presents Blinder-Oaxaca decomposition results. Results are expressed in percentage of the unconditional caste gap among children who live in regions (states) with a *high* share of UC children. Child age measured in 6-month age groups is on the x-axis. The top left panel presents the percent of the HAZ gap that is explained when all control variables are included. The top right panel is the the percent of the HAZ gap explained by endowment variables. The bottom left panel depicts the percent of the HAZ gap explained by the private investment variables. The bottom right panel presents the percent HAZ gap that is explained in percentage by public health variables.

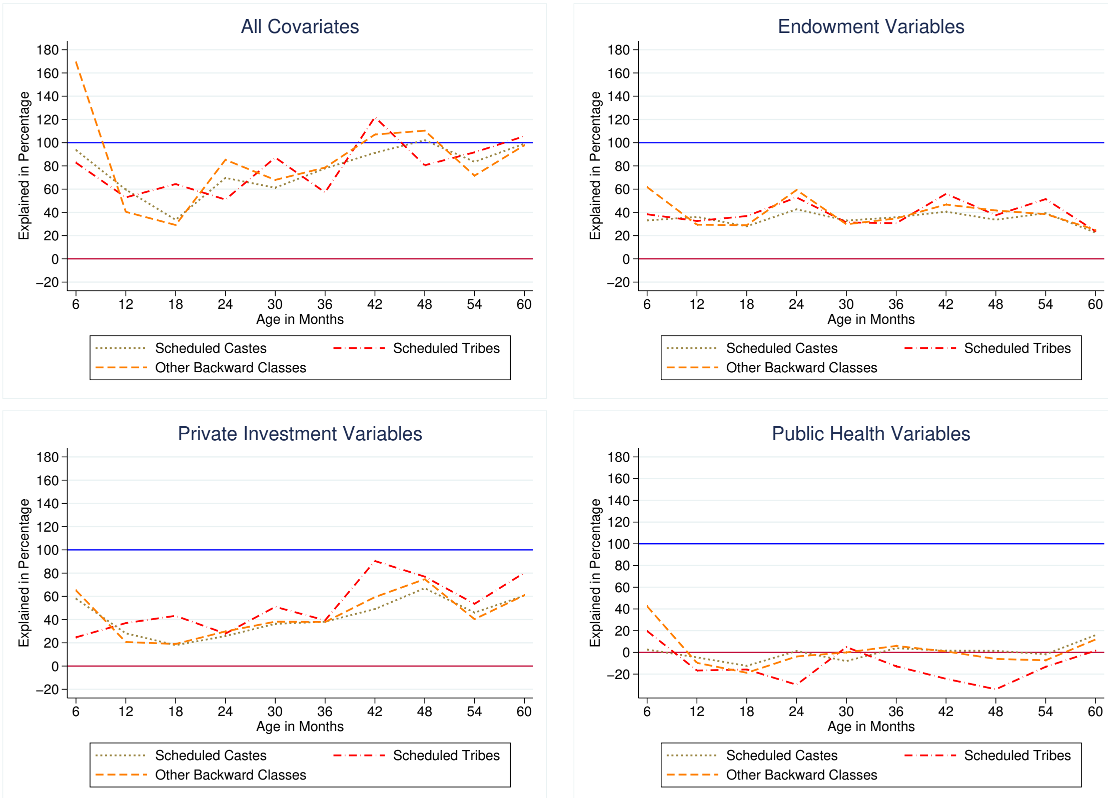

Figure A.8: Blinder-Oaxaca Decomposition Results: Explained in Percentage (Low Share of Upper Caste by State)

This figure presents Blinder-Oaxaca decomposition results. Results are expressed in percentage of the unconditional caste gap among children who live in regions (states) with a *low* share of UC children. Child age is measured in 6 month bins. The top left panel presents the percent of the HAZ gap that is explained when all control variables are included. The top right panel is the the percent of the HAZ gap explained by endowment variables. The bottom left panel depicts the percent of the HAZ gap explained by the private investment variables. The bottom right panel presents the percent HAZ gap that is explained in percentage by public health variables.

Data Availability Across NFHS-4 and NFHS-5

The National Family Health Survey 5 (NFHS-5) contains information on population, health, and nutrition for India from 2019-2021. Similar to NFHS-4, NFHS-5 includes district-level data for important variables. The total sample size of NFHS-5 is around 610,000 households. Variables collected in NFHS-5 are similar to the variables in NFHS-4, but with slight differences.

For endowment variables group, we use the same variables of birth order, maternal age, squared maternal age, maternal HAZ, maternal WAZ. However, 30% of delivery care data is missing. For the private investment variables, we include comparable variables such as wealth status, maternal education, motorcycle, own land, toilet type. To overcome the lack of data on water treatment status we utilize a water source variable. We divide this up into private water access and public water access at the home level.

The major differences is the access to vaccination variables. This information was not available for children: 1. Who were born before 2016; 2. Who were not the "last birth, next to last birth, or second to last birth" child; 3. Who died during any stage of this. In other words, the NFHS-5 questionnaire only collect information on vaccine for last 3 births but has more restrictions than the NFHS-4. Thus vaccination information from many older children (33 months and above) was not available and hence comparable across surveys.

For the public investment variables, we used sewer system access at PSU level, distance to health facilities, state-urban fixed effects. We use the water source variable to create variables that capture access to private and public water sources at the PSU level.