World Development 179 (2024) 106598

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

Regular Research Article

Caste Differences in Child Growth: Disentangling Endowment and Investment Effects

Joseph Cummins [a], Jingyan Guo [a], Neha Agarwal [b], Anaka Aiyar [c,*], Vaishali Jain [a], Andrew Bergmann [d]

[a] Department of Economics, University of California, Riverside, 900 University Ave. Riverside, CA 92521, United States of America
[b] Department of Economics, University of Otago, PO Box 56, Dunedin 905, New Zealand
[c] Community Development and Applied Economics, University of Vermont, 146 Morrill Hall, Burlington, VT 05404, United States of America
[d] Department of Agricultural, Environmental, and Development Economics, Ohio State University, 250 Agricultural Administration Building, 2120 Fyffe Road, Columbus, OH 43210, United States of America

ARTICLE INFO

Dataset link: https://www.idhsdata.org/idhs/, https://dhsprogram.com/data/available-datasets.cfm, www.osf.io

JEL classification:
I14
I15
O10
O15
O53

Keywords:
Height for age Z scores
Child health
Health disparities
India
Caste
National Family Health Survey

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 undocumented 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 (e.g. maternal height) largely explain birth length gaps, and that variables related to health investments (e.g. household wealth, health care usage) 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.

1. Introduction

We document large disparities in child height for age z-scores (HAZ) across Indian caste groups using data from the fourth round of the 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 0.2 sd shorter. By age five, caste HAZ differentials have increased by an additional 0.1–0.3 sd in each group.

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 to studies analyzing data from previous rounds of the NFHS (Coffey, Deshpande, Hammer, & Spears, 2019; Van de Poel & Speybroeck, 2009). Our results are fully consistent with this 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 Coffey et al. (2019) and Van de Poel and Speybroeck (2009) 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. To accomplish this, we shift the analytic focus of our research from the static health measures used in previous work (HAZ, stunting), onto the dynamics of child growth as children age.

Fig. 1 summarizes our main empirical focus and findings, graphing mean HAZ across child age in months for four broad caste groups: UC, OBC, SC and ST. The overall pattern is similar across all groups

* Corresponding author.
E-mail addresses: jcummins@ucr.edu (J. Cummins), jguo092@ucr.edu (J. Guo), neha.agarwal@otago.ac.nz (N. Agarwal), Anaka.Aiyar@uvm.edu (A. Aiyar), vjain013@ucr.edu (V. Jain), bergmann.46@buckeyemail.osu.edu (A. Bergmann).

https://doi.org/10.1016/j.worlddev.2024.106598
Accepted 4 March 2024
Available online 1 April 2024

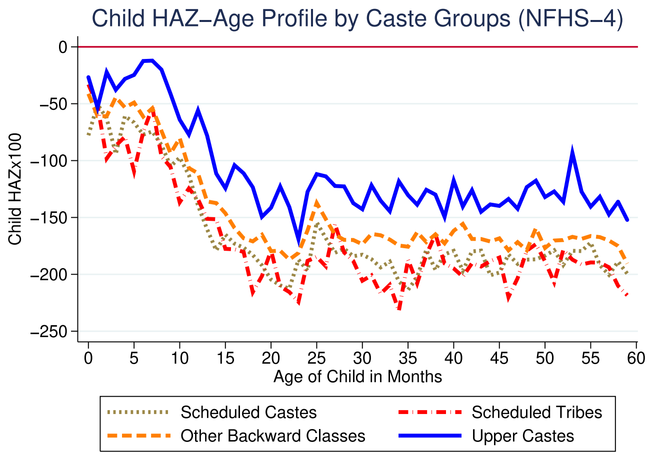

Fig. 1. Child HAZ by caste groups.
Notes: This figure graphs mean child HAZ score (x100) by caste groups for children from 0 to 5 years. The x-axis represents child age in months and the y-axis is mean weighted HAZ(x100). Median child HAZ for reference population children is 0. The mean child HAZ in 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.

and mirrors the HAZ-age profile shape common across the developing world (Shrimpton et al., 2001; Victora, de Onis, Hallal, Blössner, & Shrimpton, 2010). Indian children as a whole are, on average, born below the WHO reference population by between 0.2 sd and 0.5 sd. 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, in the standard deviation units of HAZ, then widen over the first two years of life, and subsequently remain relatively constant or slightly decrease between the ages of 2 and 5.

We structure our interpretation of these dynamics based on insights derived from a dynamic health capital accumulation model (Grossman, 1972). 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 birth length as a measure of child health endowments, and the rate of child growth as a measure of the interaction between this health endowment and the subsequent stream of health inputs consumed by the child.

Such a model predicts caste HAZ disparities that change as children age. Lower caste women are smaller, sicker and have less access to maternity care than higher caste women, potentially generating an "endowment" effect in the caste-HAZ deficit (Blunch & Gupta, 2020; Dommaraju, Agadjanian, & Yabiku, 2008; Hamal, Dieleman, De Brouwere, & de Cock Buning, 2020; Saikia, Bora, & Luy, 2019; Sanneving, Trygg, Saxena, Mavalankar, & Thomsen, 2013; Uddin, Acharya, Valles, Baker, & Keith, 2020). Similarly, lower caste households are significantly poorer than upper caste households, and many simply cannot afford health inputs like high-quality and sufficient food, or to live in areas with sanitation or clean air (Biswas, Pramanik, & Sonowal, 2023; LoPalo, Coffey, & Spears, 2019; Spears & Thorat, 2019; Thorat & Sadana, 2009; Vart, Jaglan, & Shafique, 2015). Such disparities in post-birth private and public investments may exacerbate differences at birth and generate divergent growth trajectories (Dewey & Begum, 2011; Dommaraju et al., 2008; Mertens et al., 2023).

Moreover, such a model predicts that the determinants of the HAZ gaps change as children age as well. Initial birth length deficits are likely to be the result of long-term maternal health and nutrition and/or pre-natal health care differentials, but should not be explainable by determinants of post-birth health inputs. On the other hand, household wealth disparities across caste groups are likely to generate differential health consumption patterns post-birth (e.g. in terms of nutritional intake or health care demand) that are likely to affect the rate at which children grow. Similarly, caste locational sorting may generate differential access to public goods that provide non-rival and non-excludable benefits to everyone in some area (e.g. public sanitation, pollution exposure), potentially generating differential growth patterns across caste groups as children age. The health capital theory prediction is clear: endowment-related health inputs should predominantly explain birth length disparities; post-birth consumption of (private and public) health inputs should become increasingly influential as children grow.

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 household health investment variables that have little explanatory power over HAZ gaps at birth, but which grow in explanatory power as children age. By the age of 5, we estimate that around half of the caste-HAZ deficit is due to disparities in average child health endowments, and around half due to deficits in post-birth child health inputs. Almost none of the caste gaps are ascribed by the model to disparities in public goods across caste groups. We conclude that initial endowment deficits across caste groups persist through early childhood, while private investment effects, driven by caste differentials in household wealth and expenditure, accumulate over the first few years of life.

While some of the dynamics we describe have been explored in the broader literature on growth faltering (Danaei et al., 2016; De Onis & Branca, 2016; Prendergast & Humphrey, 2014), we add new evidence documenting the age dynamics of childhood HAZ disparities across caste groups in India. We build on the economic demography literature examining the influence of health related variables in explaining caste HAZ disparities (Coffey et al., 2019; Van de Poel & Speybroeck, 2009) by studying the changing influence of household inputs over the life course of a child.

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 results cannot speak to questions of the economic and social discriminatory practices that have generated the disparities in child health endowments and investments that we document, nor to the effectiveness or efficiency of any particular policy aimed at remediating such disparities. We instead speak to the timing and subject of any policies or interventions aimed at reducing caste disparities in child health. An empirically equivalent definition of our "endowment" and "investment" related variables would be health inputs that could (in theory) be intervened upon either before or after a child is born. Our results thus reinforce an important overarching insight for both researchers and policy makers — understanding or ameliorating caste disparities in child HAZ will require addressing both contemporary disparities in early life child health inputs and nutritional intake, and disparities in the health and nutrition of the next generation of mothers.

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 (Blunch & Gupta, 2020; Borooah, 2005; Deshpande, 2000, 2001, 2007; Deshpande & Newman, 2007; Deshpande & Ramachandran, 2019; Deshpande & Sharma, 2016; Goraya, 2023; Kijima, 2006; Kumar, 2013; LoPalo et al., 2019; Maity, 2017; Munshi, 2019; Munshi & Rosenzweig, 2006; Perkins, Khan, Smith, & Subramanian, 2011; Subramanian et al., 2006; Van de Poel & Speybroeck, 2009; Vyas, Hathi, & Gupta, 2022; Zacharias & Vakulabharanam, 2011). Health outcome disparities across caste groups are well documented at every age: lower caste children are more likely to experience stunted growth; adult lower caste men and women are shorter and less healthy; and people from lower castes die younger (Kijima, 2006; Maity, 2017; Perkins et al., 2011; Raushan, Acharya, & Raushan, 2022; Saikia et al., 2019; Subramanian et al., 2006; Van de Poel & Speybroeck, 2009; Vyas et al., 2022). The literature on these caste disparities in child health outcomes has largely focused on either quantifying the extent of current discrimination against lower caste members, or 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 (Blunch & Gupta, 2020; Coffey et al., 2019; LoPalo et al., 2019).

Both sets of discriminatory practices – contemporaneous and historical – are potential explanations of the child HAZ disparities across castes that we investigate. 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; Munshi, 2019; Siddique, 2011), generating large income and wealth disparities across caste groups. These disparities in spending power then translate into disparities in health investment levels by the household, particularly disparities in food consumption and nutritional intake (Biswas et al., 2023; Choudhury, Shankar, Aleksandrowicz, Tak, & Dangour, 2021; Mahadevan & Suardi, 2013; Parappurathu, Kumar, Bantilan, & Joshi, 2015; Thorat & Sadana, 2009). Furthermore, direct discriminatory practices lower access to health care for SC, ST, and OBC children both directly (Coffey et al., 2019; LoPalo et al., 2019; Spears & Thorat, 2019) and indirectly (Blunch & Gupta, 2020; Debnath & Jain, 2020).

Additionally, long term economic inequality 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 & Currie, 2014; Chakrabarti, Scott, Alderman, Menon, & Gilligan, 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, independent of any contemporaneous discriminatory practices (Blunch & Gupta, 2020; Dommaraju et al., 2008; Hamal et al., 2020; Saikia et al., 2019; Sanneving et al., 2013; Uddin et al., 2020).

In the next section, we offer a biologically-focused economic perspective on the process through which child health disparities arise across 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 key temporal mechanisms through which disparities in child health come to exist and persist across castes in India.

3. Theoretical framework

We approach the question of child growth differentials across caste groups 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. Households can purchase consumption and health investments at a total 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 $f(H_0, I^*)$, which is realized as the health of a person of some age $H^*_A(H_0, I^*)$, and differences in health are attributed to differences in H_0 and/or 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 make it likely a priori that both of these causal channels operate on caste HAZ disparities: on average, UC women are healthier when they give birth and have more resources to provide for 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 to a child as they enter the world. We proxy for this theoretical construct at a population level 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 (Ahmed & Ray, 2018; Chari, Heath, Maertens, & Fatima, 2017; Currie, 2009; Maertens, 2013; Özaltin, Hill, & Subramanian, 2010; Swaminathan, Sharma, & Shah, 2019; Von Grafenstein, Klasen, & Hoddinott, 2023). Timely interventions on children born in the hospital can also play a role in ensuring that children are healthier at birth (Daysal, Trandafir, & Van Ewijk, 2015; Godlonton & Okeke, 2016). 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 population rate of growth of child HAZ as evidence of interactions between population-level initial health endowments and the subsequent stream of health inputs provided to the average child post-birth. 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) (Attanasio, Meghir, & Nix, 2020; Swaminathan et al., 2019). Alternatively, public goods like sanitation and health environment affect all children in an area (Geruso & Spears, 2018; Spears, 2020).

Health capital at any given age thus contains within it traces of an initial endowment component (H_0) and a component determined by the subsequent stream of health inputs consumed by the individual. While for adults we might expect the initial endowment effect to be wiped out by the subsequent health experiences of the person, for children the relative importance of the two is less clear. The health endowment is more recent, and the stream of privately and publicly provided health investments a child has received up until that point in their lives is relatively short.

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

A good measure of health capital for children, then, would be one that captures the cumulative health of a child from birth until they are measured. We argue that HAZ is precisely such a measure.

Children are born 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, child length or height contains within it information on the cumulative health experience of children from birth until they are measured, in the same manner as health capital.

Child HAZ is an age- and gender-normalized measure of child height. 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, India. These well-nourished children grow, on average, at the same rates across the sampled countries, implying that population level deviations from these standards are not determined by genetic origin, but instead by the circumstances of children's birth and growth. This normalization allows for comparisons of the relative magnitude of child health outcomes as children age.

Other anthropometric measures of child health, such as those based on weight or arm circumference that 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. Measures related to acute child morbidity are even shorter-term measures. Child mortality rates may contain information on average very early life health capital, but the infrequency of mortality at ages beyond the first year makes detecting differences in health levels of older children largely impossible.

3.4. HAZ-age profile

Shrimpton et al. (2001) 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 (Alderman & Headey, 2018; Rieger & Trommlerová, 2016; Roth et al., 2017). Focusing on the HAZ-age profile, as opposed to mean HAZ, centers our empirical perspective on the health capital accumulation process itself, as opposed to the stock of health capital at any particular moment in time.

Given this consistent and stable pattern in global child growth patterns, Aiyar and Cummins (2021) develop regression models to capture the effects of a key variable of interest on changes in the location and 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 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 HAZ-age profiles across socio-economic groups.

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. Table 1 provides summary statistics. Indian children are on average 1.49 sd 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)[2]. Overall, in our weighted sample, 23.7% of children 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 1.67 sd below, OBCs are 1.49 sd below, and ST children are around 1.72 sd below the reference height for their sex and age, while UC children are only 1.12 sd 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, variables for private health inputs, and public health inputs. A table of representative summary statistics is provided in Table 1, and the full table including all variables is available in the Online Appendix Table A.1.

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). Similarly, the total effects of private sanitation on health may nullify the effects of public sanitation on health. 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 two distinct channels of health disparities (health endowments and post-birth investment in health inputs) as determinants of the caste gaps in child HAZ.

An empirically equivalent definition of "endowment" and "investment" related variables would be variables that could (in theory) be acted upon either before (endowment) or after (input/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 important for 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 to define the contents of three families of explanatory variables.

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 2.03 sd shorter than the median height of healthy women across the world. UC mothers are on average healthier than

[1] for Population Sciences (IIPS) and ICF (2021).

[2] Our main estimates compare SC/ST/OBC children to UC children. In the Appendix we provide an analysis comparing Muslim Indian children with UC Hindu Indian children.

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. OBC mothers are slightly healthier than SC and ST groups but less healthy than US mothers. About half of SC and OBC mothers report having their deliveries in a hospital, whereas 71% of UC mothers deliver in hospitals.

4.2.2. Private investments

Private investment variables include household wealth index quintiles, 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 report access to clean drinking water and 45% of households own agricultural land. On average, UC households tend to be able to provide higher levels of private investments to their children relative to lower caste households. 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 in the poorest wealth quintile 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 women 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 mothers 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 the lowest at 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 about a third as likely to live in urban areas (11% for ST, 39% for UC).

These measures of public health inputs are limited relative to the richer and more predictive information we have on maternal and household characteristics. Sewer and health center access carry information about important channels, but they are rough, aggregate measures of key inputs such as the public health environment. We attempt to ameliorate this issue with a novel interpretation of a common regression adjustment. Regional or locational indicator variables by their nature capture the unobserved determinants of child health that are common to people in a specific place and time. In the context of child health, this could include access to health care, relative food prices, water and air quality, socio-cultural practices and any number of other unmeasured variables that affect every child in the community. We interpret these indicators as a kind of net public goods provision to children, determinants of HAZ that are non-rival and non-excludable for the local population, regardless of caste affiliation. If sorting across location (and thus public health inputs) statistically explains the caste HAZ gaps, regional variables should capture this effect.

Table 1
Summary statistics (NFHS-4).

	SC Mean/sd	ST Mean/sd	OBC Mean/sd	UC Mean/sd	All Mean/sd
HAZ (×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 (×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)
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)
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)
No toilet facility	0.60	0.74	0.51	0.25	0.51
	(0.49)	(0.44)	(0.50)	(0.43)	(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	34 913	21 441	64 206	26 218	146 778

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 differentials 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 towards 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:

$$Y_{irvg}^{A} = Caste_{g}^{\prime A}\delta_{g}^{A} + X_{1irvg}^{\prime A}\beta_{1}^{A} + X_{2v}^{\prime A}\beta_{2}^{A} + \lambda_{r}^{A} + \epsilon_{irvg}^{A} \quad (1)$$

where Y_{irvg}^{A} is the HAZ of child i, aged A, living in region r in PSU v and belonging to caste g. $Caste_{irvg}^{\prime A}$ 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_{r}^{A} 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_{g}^{A}, 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 Eq. (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 Eq. (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 HAZ-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 (the average 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. In contrast, when we bin regressions by age as above, 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 simultaneously 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 and estimated separately by caste-region cells. 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 (Shrimpton et al., 2001; Victora et al., 2010), and we include only cells with at least 20 observations to ensure reasonable within-cell estimates of the intercept and slope.

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

Y_i^{rg} is the HAZ of child i from region r and caste-group g. $\hat{\alpha}^{rg}$, the estimate of the implied (sub-group) average birth-length z-score, is interpreted as a measure of the average health endowment for caste group g in geographic region r, defined in our main specifications as district-by-urban groups. Similarly, $\hat{\beta}^{rg}$ provides an estimate of 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 interaction between the (subgroup average) initial health endowment and the (subgroup average) 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:

$$\hat{Z}_{rg} = Caste_g'\delta_g + Endowment_{rg}'\beta_1 + Private_{rg}'\beta_2 + Public_{rg}'\beta_3 + \epsilon_{rg} \tag{3}$$

Here, $\hat{Z}_{rg}$ is either the estimated health endowment, $\hat{\alpha}_{rg}$, or estimated rate of loss of HAZ, $\hat{\beta}_{rg}$, from Eq. (2). δ^g represents the caste gap in health endowments or rate of growth.

As above in the individual-level regressions, we estimate unconditional versions of Eq. (3) and conditional regressions that adjust the gap for observable covariates. For the conditional regressions, we include the district-urban-caste-group mean of our endowment and private investment variables. Public health variables are captured by state-urban cell fixed effects, sanitation access at the PSU level, and distance to the nearest health facility. These regressions are run at the district-urban-caste level and are weighted by cell-size of each district-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 an Oaxaca-Blinder (OB) decomposition to answer that question.

In the OB decomposition framework, the difference in mean HAZ between caste group g and the upper caste group, uc, for a specific age group A can be represented as:

$$\overline{HAZ}_g^A - \overline{HAZ}_{uc}^A = \overline{\beta}_g^A \bar{X}_g^A - \overline{\beta}_{uc}^A \bar{X}_{uc}^A \tag{4}$$

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

$$\overline{HAZ}_g^A - \overline{HAZ}_{uc}^A = \hat{\beta}_g^A(\overline{X}_g^A - \overline{X}_{uc}^A) + \overline{X}_{uc}^A(\hat{\beta}_g^A - \hat{\beta}_{uc}^A) \tag{5}$$

The first and the second terms on the right-hand side 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 (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, calculated as the explained variation (in HAZ units) divided by the total gap in the unconditional regressions (also in HAZ units) across all the age groups, and then multiplying by 100.

6. Results

6.1. Unconditional and conditional estimates

Fig. 2 (Top Panel) shows regression estimates from the unconditional model from Eq. (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 0.21 sd below UC children at birth. OBC children then grow less quickly than UC children after birth and average differences increase to around 0.38 sd by an OBC child's third birthday. These differences reduce to around 0.33 sd by the 4th year, providing some evidence of catch up in later years. SC children are born with HAZ scores around 0.37 sd below UC children. Over the first three years, they lose, on average, an additional 0.25 sd in HAZ, leading to a 0.62 sd difference by their third birthday.

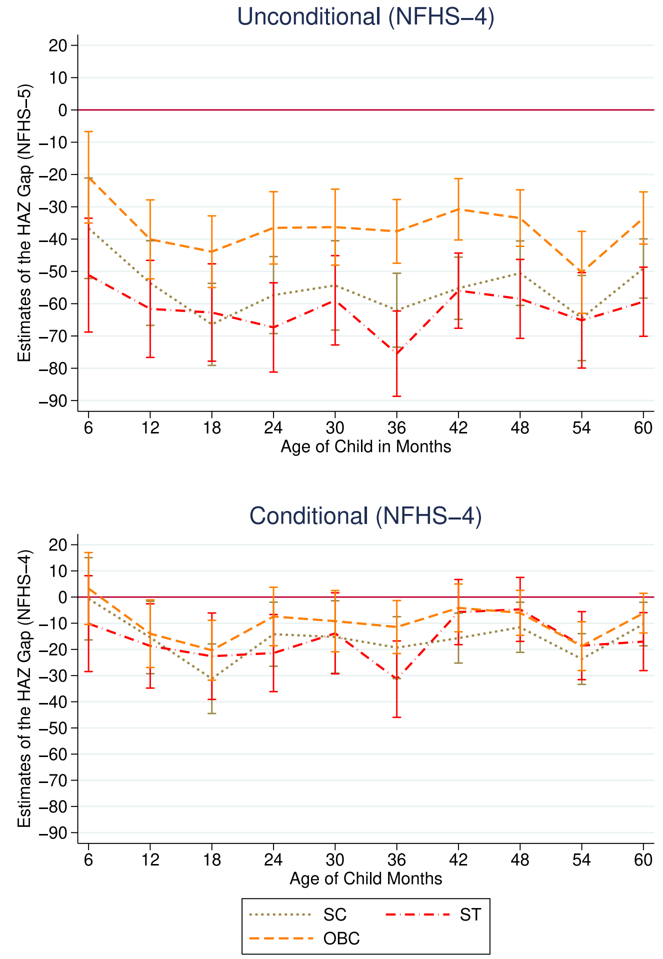

Fig. 2. Regression estimates of caste HAZ differentials by age (0–60 months).
Notes: This figure presents regression estimates for both unconditional and conditional regressions in Eq. (1). The y-axis provides the coefficient estimates (and confidence intervals) on the caste-group variables, and the 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.

After that, heights improve slightly and by the child's fifth birthday they are on average 0.5 sd below a UC child. ST children experience the largest deficit of around 0.51 sd at birth. They face a large dip in HAZ by their 3rd birthday where differences grow to around 0.75 sd. By the time they are done with early childhood, ST children are 0.6 sd below UC children.[3]

Fig. 2 (Bottom Panel) provides conditional (regression adjusted) caste gap estimates controlling for covariates related to the health endowment, private investment, and public health variables described above. Including covariates reduces differences in HAZ gaps across all caste groups at all ages. STs, who face the largest unconditional HAZ gap relative to UCs, see reductions in the gap at birth of around 0.41 sd (to about 0.10 sd) once covariates are included. Adjusted differences in birth length differences are indistinguishable from zero. 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 also statistically insignificantly shorter than UC children at birth. However, as children age, differences begin to appear. Coefficients for OBCs children hover between 0 and 0.2 sd across most ages, with estimates for only 4 of the 10 periods statistically distinguishable from 0 at a 95% confidence level. Adjusted estimates for SCs show similar patterns with somewhat larger point estimates that hover between 0.05 to 0.31 sd, with 9 of the 10 estimates being significantly different from 0. Even still, covariate adjustment reduces the height gaps between OBCs and SCs with UCs by well over half.[4]

One concern is that differential infant mortality across caste groups could be driving our estimates of the very early life caste gaps. Most children who die as children do so within the first six months of life. If lower caste children are more likely to die, and those that do die have differential potential HAZ across caste groups, this mortality effect would appear to us as part of a differential health endowment across caste groups. We provide two reasons this concern is not likely to prove fatal to our interpretation. First, in both theory and as documented around the world, child stunting is predictive of infant mortality (Olofin et al., 2013). In this case, our estimates of the caste HAZ deficits would likely be biased towards zero over the early period. Second, a recent paper by Panda (2020) argues that infant mortality risk is actually positively correlated with (potential) child HAZ in India. The study estimates that children who die would be approximately 17% taller (in units of HAZ) than the average child who lives, or approximately 0.34 sd at a mean HAZ of −2 sd. Meanwhile, Bora, Raushan, and Lutz (2019) estimates child mortality rates that are approximately 1 percentage point higher for SC/ST children than other children. Taken together, these would imply that absent differential mortality, mean HAZ for SC/ST children would be about 0.003 sd higher than we observe. This is approximately two orders of magnitude below our estimated caste gaps, and thus unlikely to greatly affect our inferences regarding the importance of health endowment deficits.

6.1.1. Regression estimates by gender and location

The results above provide average correlations across the entire population, estimated as though the associations are similar for all groups of children. There are numerous reasons to believe this may not be the case: child growth faltering is different in rural and urban settings (Rieger & Trommlerová, 2016), gender differences in the effects of caste discrimination are assured (Deshpande, 2007; Islam, Pakrashi, Sahoo, Wang, & Zenou, 2021), and different parts of India enforce and experience caste-based exclusionary social practices like untouchability in different ways (Coffey et al., 2019; LoPalo et al., 2019). Fig. 3 provides disaggregated estimates of unconditional and conditional caste gaps across three dimensions: child gender, rural/urban status, and across states of high and low density of UC Hindus.

The top row displays results when our models are estimated separately on girls and boys. Each graph shows unconditional (solid) and conditional (dashed) caste-gap estimates across child age for one gender, with girls on the left of the first row, and boys on the top right. There are large unconditional caste differentials for both boys and girls, on the order of 0.2–0.55 sd, and these gaps tend to grow in magnitude over the first two years of life. There is no noticeable difference in either the magnitude or the age-dynamics of the caste HAZ gaps across child gender. And as with the aggregate results, in both cases the unconditional estimates greatly attenuate and become, for the most part, statistically insignificant when controls are added.

The middle row divides the sample into rural (left graph) and urban (right graph) sub-samples according the NFHS definition. The results of the aggregate sample appear largely to represent attenuated dynamics in the rural areas. Rural children display the signature growth faltering dynamics seen in our aggregate results, but the growth faltering process is harder to characterize in the urban sample, in part due to noisier

[3] Regression estimates are provided in Table A.2 in the appendix.

[4] Detailed regression estimates are provided in Table A.3 in the appendix.

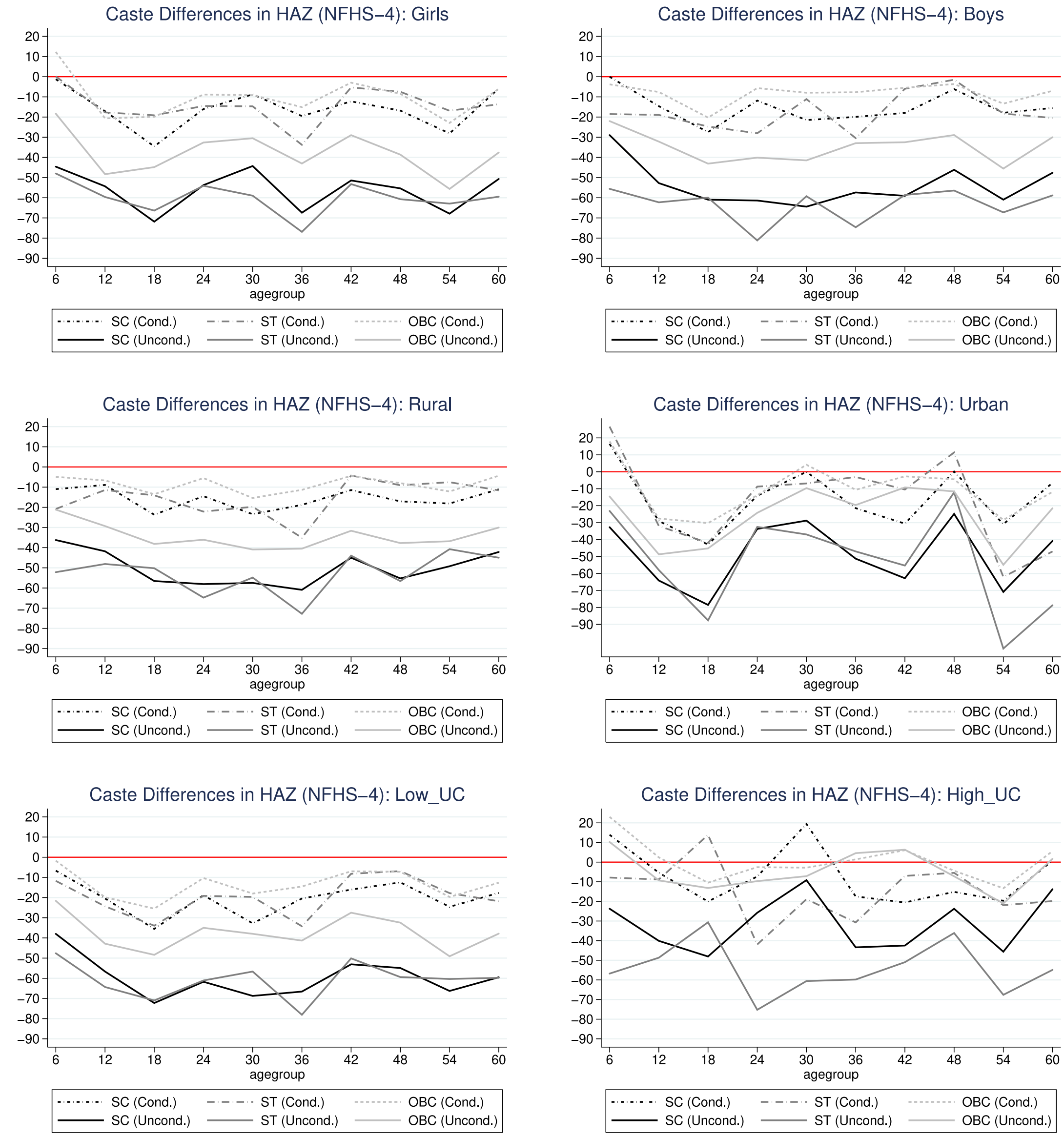

Fig. 3. Replication of individual-level results by gender and location (NFHS-4).
Notes: The graph on the top left panel plots the regression estimates for girls and the graph on the top right plots the regression estimates for boys. The graph in the middle panel on the left plots the regression estimates for rural children and the middle panel on the right plots the regression estimates for urban children. The bottom left panel has estimates from states with a below median share of UC children and the bottom right panel has estimates of the states with higher than median share. In each graph, the y-axis is the coefficient estimates on the caste-group variable, and the x-axis separates the estimates into 6-month age-bins. The solid lines presents regression estimates from the unconditional model and the dashed lines are estimates from the conditional model. 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.

estimates from a smaller sample size. Urban and rural OBC children are born approximately 0.2 sd below UC children. Compared with UC children, SC and ST children in urban areas are born about 0.3 sd and 0.2 sd shorter respectively. In rural areas, these differences at birth are much larger, 0.35 sd (SC) and 0.55 sd (ST). The loss of HAZ over the first two years of life is common across all groups of children in rural areas, leading to HAZ gaps of around 0.6 sd for SC and ST children by age 2, and around 0.4 sd for OBCs. The loss of HAZ in the urban sample is more rapid over the first 18 months, dropping approximately an additional 0.4 sd for each group, but HAZ gaps immediately narrow in the next six months of life, and it is difficult to distinguish the extent to which massive changes from 18 to 24 months constitute real improvements or simply mean reversion in the estimates. In both areas, however, adjustment for covariates greatly attenuates the gaps towards zero at all ages, and most coefficients become statistically insignificant.

The final row of Fig. 3 displays results across states with low (left graph) or high (right graph) share of UC children. This follows the intuition in Coffey et al. (2019), who divide their sample into regions with higher and lower concentrations of UC Hindus as a proxy for local discrimination, though we construct our sub-groups somewhat differently. We use the median of state-level shares of UC (~24%) to divide the states into those with a high and low share of UC. While this process generates an almost equal number of states in each group, it does not generate equivalent sample sizes. Low UC share states contain almost 80% of the total sample observations, and the relatively noisy estimates for the high UC states are likely driven by this much smaller sample size.

Overall, the results for low UC states are more in line with the aggregate results as compared to the high UC states. Children from lower castes in low UC states are born 0.2–0.5 sd shorter than UC children. The growth faltering in low UC states is characteristic of the aggregate results, and these initial gaps increase by 0.2–0.3 sd over the first two years. The estimates also attenuate at every age after adjusting for covariates. On the other hand, in the high UC states, OBC children are no shorter than UC children at any age, even in the unconditional regressions. However, SC and ST children are between 0.2 sd and 0.6 sd shorter than UC children in the unconditional estimates, though the growth faltering process in these states is not visually apparent and the age-dynamics are indistinguishable from noise.

We note a few commonalities across the results and in relation to the aggregate results. First, the unconditional estimates for most sub-groups are of similar magnitude and statistically significant. UC children are born with better health than all types of sub-groups across India. Second, these differences also largely attenuate towards 0 in conditional regressions for all sub-groups, following the pattern in the full sample. Moreover, the growth faltering dynamics we focus on in our interpretation hold in the larger subsamples, while in the smaller sample sizes (Urban and High UC) we are unable to distinguish these dynamics from noise in our estimates.

6.1.2. Regression estimates for 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 replicate our analyses by comparing Muslim children with UC Hindu children. We provide a brief overview here, but elaborate further in Online Appendix G.

In Figure A.2 we see that the HAZ-age profiles demonstrate that Muslim children display a similar growth faltering dynamic as seen for low caste Hindu children. We then replicate our regression results in Figure A.3. Unconditional estimates in the top panel shows that Muslim children are around 0.2 sd smaller than UC children at birth. These differences increase to around 0.45 sd by age 5. Conditional estimates are small and statistically insignificant for the first 6 months. By age 2, Muslim children are around 0.25 sd shorter than UC children, and this gap maintains a relatively constant magnitude (and largely maintaining statistical significance) through age 5.

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 & Cummins, 2021). Results from Eq. (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 (Eq. (3)) is estimated without covariates, and columns 2 and 4 provide estimates when district-caste-urban group level mean covariates and state-urban fixed-effects are included in the second stage.

Comparing results from columns 1 and 2 in Table 2 on birth length, we find that they are similar to those in the first age-bins of the individual-level regressions. SC children are on average born a statistically significant 0.47 sd shorter than the average UC child. Conditional on the inclusion of covariates, this difference reduces to 0.04 sd and is statistically insignificant. OBC children have a statistically significant average unconditional birth length deficit of 0.31 sd relative to UCs, a smaller gap compared to SC, but these differences reduce to 0.02 sd and are statistically insignificant once we condition on household covariates. Unconditionally, ST children have statistically significant

Table 2
Rate of HAZ loss and caste (NFHS-4).

	(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	−46.65***	−4.72	−1.46**	−0.47	−3.90***	−0.74**
	(9.88)	(13.45)	(0.60)	(0.90)	(0.41)	(0.35)
ST	−54.80***	−8.68	−0.92	−0.83	−3.79***	−1.32***
	(15.23)	(16.31)	(0.86)	(1.03)	(0.49)	(0.39)
OBC	−30.67***	−1.53	−0.72	−0.36	−2.32***	−0.45*
	(8.32)	(10.63)	(0.57)	(0.70)	(0.35)	(0.24)
Mean	−37.3	−37.3	−7.8	−7.8	−7.8	−7.8
R Square	0.04	0.25	0.01	0.18	0.68	0.81
Weighted N	37 959	37 959	37 959	37 959	37 959	37 959
Real N	920	920	920	920	920	920

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.

birth length deficit estimated at 0.55 sd. Conditionally, this difference reduces to 0.09 sd and is statistically insignificant.[5]

A similar picture emerges from the slope (β) estimates, where differences in growth rates are relatively large, and also largely attenuate when adjusted for observable group covariate means. Unconditional caste group point estimates in Column 3 are negative and statistically significant for SCs (−0.015 sd/month) but not statistically significant for OBCs and STs. The adjusted estimates in column 4 for SC and OBC estimates are insignificant and closer to 0.

Columns 5 and 6 in Table 2 provide 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 a birth length effect) would need to condition on the birth length. In columns 5 and 6 we include $\hat{\alpha}$ as a control variable for β. 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 including $\hat{\alpha}$ and our standard suite of covariates. Conditional on the intercept estimate, the slope estimates indicate large growth rate gaps between UC children and SC (−0.04 sd/month), OBC (−0.02 sd/month) and ST (−0.04 sd/month) children. These gaps would produce a cumulative effect of between −0.55 sd (for OBC) to −0.9 sd (for SC and ST) HAZ points by 24 months of age. Point estimates attenuate and remain statistically significant with the inclusion of group level covariates. Conditional on the intercept and covariates, the slope estimates are −0.007 sd/month for SC, −0.005 sd/month for OBC and −0.013 sd/ month for ST, a reduction on the order of 75% in magnitude relative to the unconditional estimates.

6.2.1. α and β estimates by gender and location

In Table A.5 to Table A.10 in the online appendix, we present the full sets of results from a disaggregation of the α and β estimates of caste

[5] 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. The top left panel in Figure A.1 provides estimates of caste gaps in α at alternative cutoffs and the results are relatively robust to selection of age cutoff after 20 m of age. The top right panel provides similar robustness checks for the unconditional estimates of β.

gaps by gender and location. A caveat is that by restricting our analyses to district-urban cells with more than 20 observations, we lose a large portion of the sample within each sub-group analysis. We discuss this in more detail in Section 8, but we note that these results should be interpreted with some caution.

In relation to gender differences, we estimate a higher unconditional birth length for girls than boys (−0.32 sd compared to −0.53 sd) though the two groups have comparable rate of loss estimates of approximately −0.08 sd (Table A.5 and Table A.6). However, the unconditional caste gaps for both birth length and rate of loss are of a similar magnitude across gender. The conditional HAZ-caste gaps for α among girls range between 0.16 sd–0.27 sd but are not statistically different from UC children heights. The estimate of the growth rate for ST girls, conditional on both our usual covariates and the estimated birth length, is the only statistically significant conditional estimate and indicates a rate of loss of 0.002 sd per month (less than .05 sd cumulative by age 2), which is also the largest estimated group difference in β. For boys, conditional estimates for birth-length or rate of loss are not statistically different from 0.

As shown in Table A.7 and Table A.8, the caste gaps for rural areas are largely in line with the overall results, while again those for the urban areas are less consistent. Rural children are generally born shorter (mean $\hat{\alpha}$ of 0.38) than urban children (mean $\hat{\alpha}$ of 0.28), and also lose HAZ at a faster rate (mean $\hat{\beta}$ of −.08 sd/month for rural, relative to −.06 sd/month for urban). Unconditional HAZ-caste differences exist at birth in rural areas (0.3–0.5 sd) and urban areas (0.2–0.7 sd) but the urban estimates are generally imprecisely estimated. The rural birth length estimates attenuate greatly towards 0 and become insignificant when conditioning on observable variables and the urban sample birth length estimates become increasingly imprecise and in some cases increase in magnitude. The slope estimates for both rural and urban samples are similar to the overall results, with the largest implied gaps in rate of loss of HAZ appearing for urban SC and ST children. Conditioning on birth length, rural ST children grow more slowly than UC children (0.01 sd), but there is no measurable caste gap in growth rates for SC and OBC children. In the urban sample, SC children show significant growth rate gaps with UC children and these estimates are relatively large at around 0.02 sd/month.

Consistent with the results in Section 6.1.1, the results for low UC states in Table A.9 are comparable with the aggregate results in Table 2, while those for the smaller sample in high UC states are not (Table A.10). A key difference is in the mean intercept results. Children in low UC states have a mean $\hat{\alpha}$ of −0.4 sd, but those in the high UC states have an estimated mean close to zero. This is likely a result of the relatively small sample size in high UC states, and not necessarily a feature of differential growth faltering. Estimates of the rate of loss of HAZ are similar across regions at around 0.07–0.08 sd/month for both groups. Unconditional HAZ caste gaps at birth vary from 0.14 sd to 0.25 sd in the low UC states, and are all statistically significant. In the high UC states, estimates of the birth length caste gaps are smaller and not statistically significant, but they are the opposite sign for ST children. Conditioning on covariates attenuates the estimates of birth length gaps by about half for the low UC regions, while in the high UC regions the estimates all become positive (but remain statistically indistinguishable from zero).

6.3. Decomposition results

Fig. 4 shows the OB decomposition results, presented as the explained percentage of the unconditional height gap as described in Eqs. (4) and (5). Child age binned into 6-month age groups is on the x-axis and the y-axis shows the share of the caste gap that can be explained by various groups of model covariates.

In general, well over half of the unadjusted caste HAZ differences can be explained by our observable covariate groups across all caste groups and all ages (top left panel). 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 with UC children is explainable by age 5 for every other caste group. 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 at any given age.

The remaining panels of Fig. 4 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 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.[6] 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.

In general, we find surprisingly small associations between public health variables and child HAZ, despite the well-known importance to child health of public goods like clean water and sanitation services. While public health environment is clearly an important determinant of child HAZ, it is not, according to our estimates, important to explaining caste HAZ gaps. There are two important caveats to this interpretation. The first is one of measurement: there is a general lack of high quality measures of public health inputs in our data, and we rely on regional indicator variables to capture much of any such effect. These regional variables are relatively large to allow for variation in our observable public health inputs at the community level, and thus may not capture imbalances in public health environment that are real in the world. The second is that even our direct local measures of public health – access to a sewer system and distance to health care facility – capture only the separate effect of having a community sewer system or health facility access, while the actual presence of a toilet in the house or health care visits for the child are ascribed to the private investment channel. Our work should not be interpreted as suggesting that public health is not an important determinant of child HAZ, but if it is important for explaining caste HAZ gaps, then it must be operating at either finer geographic levels and/or via inputs that are not observable in the data.

6.3.1. Decomposition results by gender and location

We also provide decomposition results for the previously analyzed subgroups in the Online Appendix. The loss of sample size (about half in each sub-group analysis) leads to noisier estimates, but they largely mimic the aggregate results described above.

Figures A.5 and A.6 show that overall results differ slightly by gender. For both the groups, birth length is almost fully explained, after which the explanatory power of covariates falls up to the age of 18 months. By age 5, almost all of the caste HAZ gaps for girls and 60% of the caste HAZ gaps for boys are explained by included set of variables. Endowment variables explain the birth length entirely for girls, while accounting for 40%–65% of the caste HAZ gap in birth length for boys. Thereafter, the share explained remains stable at around 40% for girls

[6] Information on the shares from the OB decomposition can be found in Table A.4.

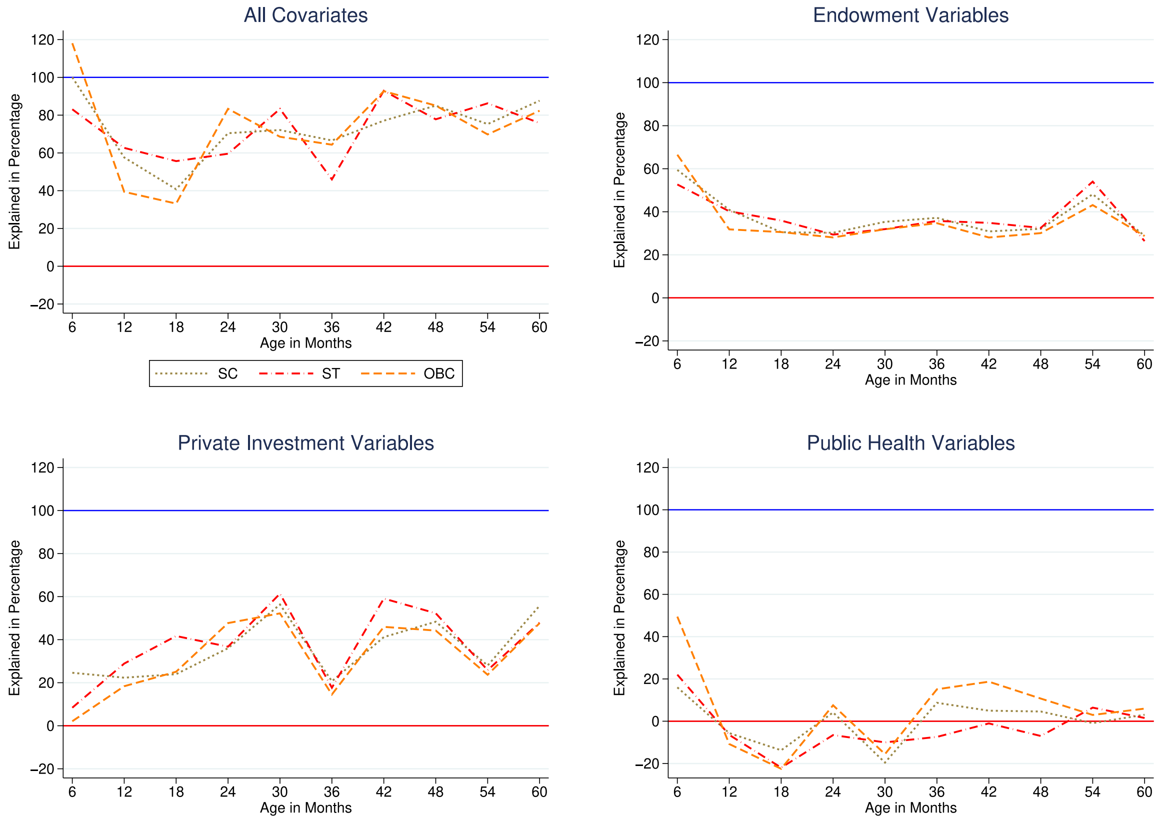

Fig. 4. OB decomposition results: percent explained.
Notes: This figure presents Oaxaca–Blinder decomposition results. Child age measured in 6-month age groups is on the x-axis and the y-axis is the percent of the caste gap that can be explained by model covariates. The top left panel presents the percent 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.

and 20%–40% for boys at any age. The share explained by private investment variables starts low at birth and increases overtime for both genders. As before, the share of public health variables is null overall, with the exception that it explains a high share of the differences in the caste-height gap for OBC boys at their birth. These effects attenuate within the first 12 months.

Figures A.7 and A.8 present the results by rural and urban areas. Consistent with earlier results, the estimates from rural areas are similar to the aggregate results. The share of variation explained by endowment variables starts high and remains stable. The share of variation explained by private investment variables increases as children age. Public health variables continue to explain little of the HAZ caste gaps. Similar to the urban estimates of caste gaps themselves, our decomposition results do not reveal consistent patterns for urban children. This is likely attributable to both a decrease in sample size, and to the general differences in growth faltering patterns for urban children noted above.

In figures A.9 and A.10, we separate our results by the share of UCs in a state. Once again, consistent with findings in Section 6.1.1, the patterns for children in states with a low share of UCs (which accounts for 80% of the sample) are broadly consistent with the aggregate results. Birth length caste gaps are almost entirely explained by the included set of covariates up to 18 months. Endowment covariates have the highest explanatory share at birth and maintain a stable share at around 30% afterwards. The share explained for private investment begins at 20%–30% at birth and grows as children age. Public health variables, on the other hand, do not contribute in explaining the caste HAZ gaps. States with a high share of UC provide noisy estimates, driven by the relatively small sample size for these states.[7]

6.3.2. *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, private investment variables 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 Fig. 5.

The top row of Fig. 5 graphs the outcome-age profile for two representative 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 take-up of various vaccines, a representative age-varying private investment. If age-varying measures like vaccinations are driving the explanatory power of private investments, then the apparent increase in the explanatory power of private investment variables could, in theory,

[7] 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 do not seem to explain much of HAZ differences between UC and Muslim children (Figure A.4).

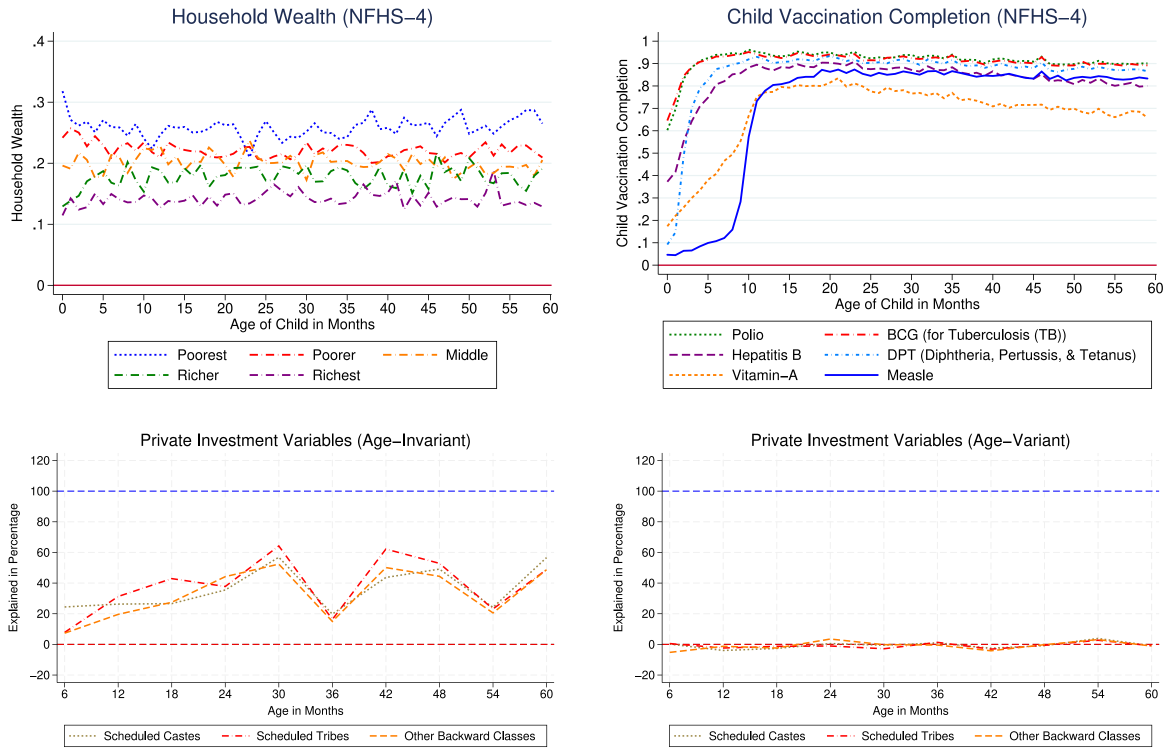

Fig. 5. Age-specific and non-age-specific private investments.
Notes: The top two panels of the figure shows the frequency of two particular private investment variables for children from 0 to 5 years. The top left panel shows the age-invariant nature of the household wealth measure of private investments across child age, while the top right panel shows the age-specific frequency of vaccination status by child age. The bottom row displays Oaxaca–Blinder 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.

be simply the result of the increasing take-up of vaccines as children age. In that case, the estimates of increasing explanatory power of private investments would be driven by some of them simply turning on.

The bottom row of Fig. 5 shows the Oaxaca-Blinder decomposition for the suites of age-invariant and age-varying private investment variables (estimated from the full model, but with their contributions accounted for separately). 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 caste HAZ gaps.

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. Our initial choices over the sample selection rules and regression model specifications 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. Subsequently, data from the next round of the NFHS-5 was made available to researchers. We then replicated the results from the NFHS-4 on data from the newly available NFHS-5, using the same methods and (to the extent possible) the same regression specifications as before. 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 analyzing 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. The sample that answered the vaccination (Polio, DPT, Hepatitis B, Measles, BCG) question also changed in NFHS-5. Nevertheless, using this newer dataset, we can replicate a majority of the analyses conducted on the NFHS-4.

Fig. 6 presents a comparison between the results from the NFHS-4 (left column) and NFHS-5 (right column). The patterns in the HAZ-profiles (first row) are quite similar between the two waves. Children from lower castes 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. In the second row, we present the unconditional regression estimates of caste group differences in HAZ. Compared to NFHS-4, there appears to be a small secular decrease in unconditional caste gaps overall, with all lower caste groups seeing smaller unconditional gaps relative to UC children. In the NFHS-5 sample, differences in height at birth between OBC children and UC children have become statistically insignificant, while the gaps for SC and ST children remain significant but are relatively smaller than in the previous round, decreasing from 0.4–0.5 sd in the NFHS-4 to 0.2–0.3 sd in the NFHS-5. This upward level shift of the HAZ-age profile across rounds persists through the growth faltering process, and HAZ caste gaps from age 2 onwards are all about 0.1 sd smaller in the newer round of data. Conditional estimates, presented in the bottom panel, are more similar across rounds, and 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 (Coffey et al., 2019; Van de Poel & Speybroeck, 2009).

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

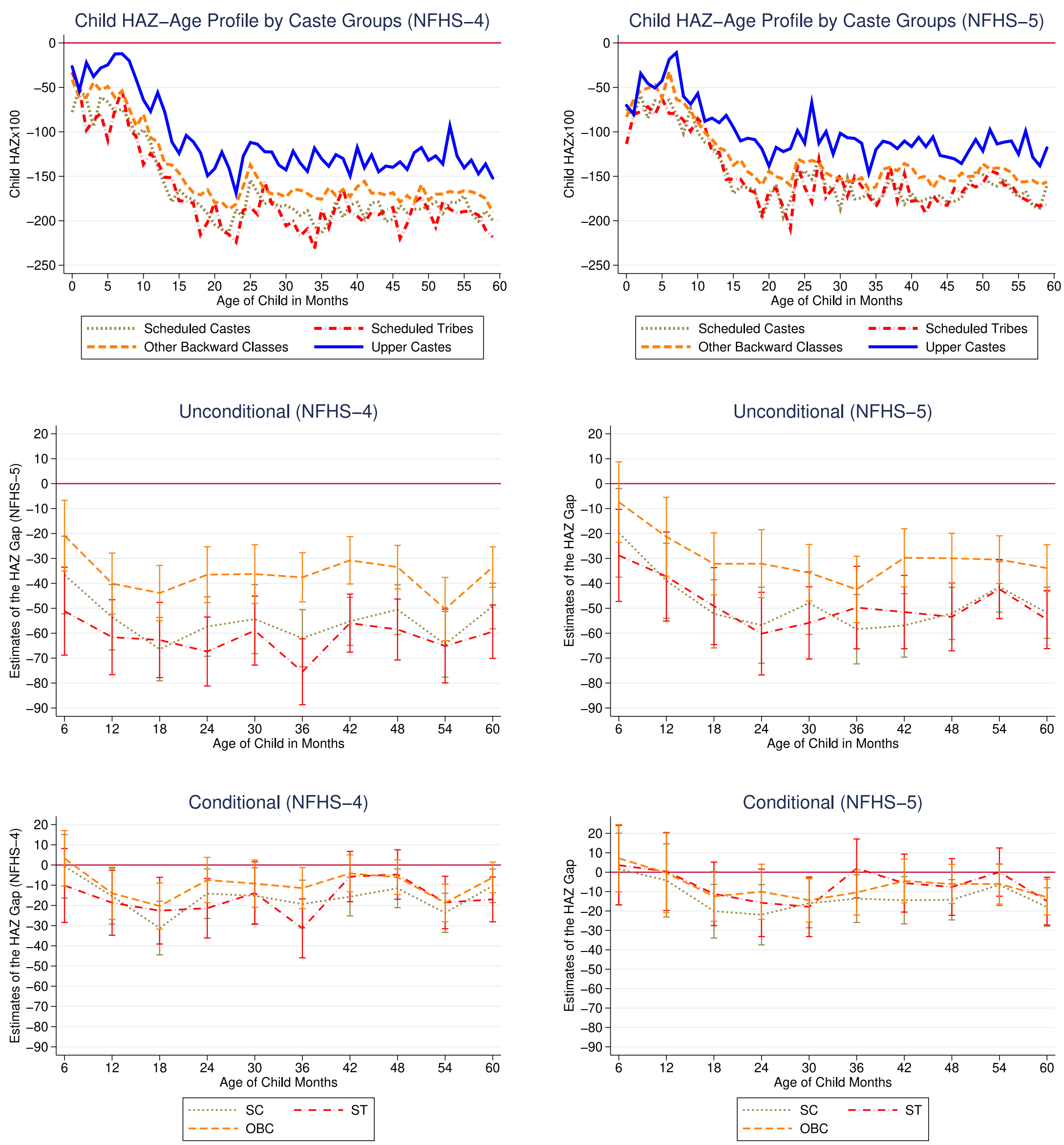

Fig. 6. Replication of individual-level results: NFHS-4 versus NFHS-5.

Notes: 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. The gray dotted line is the difference of SC and UC children, the red dashed-dotted line represents differences between ST and UC children, and the orange dashed line represents differences between OBC and UC children. 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. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

data. By and large, the patterns are similar across the two rounds in sign and magnitude but some differences emerge. Mean implied birth length (α) is actually lower in NFHS-5 (−.45 sd compared to −.37), but mean rates of loss (β) are lower as well (−0.06 sd/month compared to −0.08 sd in NFHS-4). On the other hand, the unconditional caste gaps in HAZ at birth decline over time (−0.31 sd to −0.55 sd in NFHS-4 as compared to −0.2 sd to −0.4 sd in NFHS-5). Conditional on covariates, the results from NFHS-5 follow those from NFHS-4 and all differences in birth-lengths become less than 0.1 sd and statistically insignificant. The estimated gaps for rate of loss of HAZ across caste groups are relatively stable across rounds. For ST children, the gaps in unconditional rate of loss increases from 0.009 sd in NFHS-4 to 0.012 sd in NFHS-5, while the unconditional slope gap estimates are small and statistically insignificant for SC and OBC children. Comparing columns 5 and 6, the slope estimates conditional on $\hat{\alpha}$, there is no meaningful change in the estimated gaps in β across rounds.

The striking similarity of results over the two waves indicates an amount of temporal stability in the child growth dynamics we originally observed in the NFHS-4, even when applying a de facto pre-defined regression model to the NFHS-5 data. This eases concerns related to cherry picking results or 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 across caste that we identified in the 2015

Table 3
Rate of HAZ loss and caste (NFHS-5).

	(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	−37.09***	−12.54	−1.35**	−0.56	−3.44***	−1.27**
	(8.02)	(12.75)	(0.60)	(0.83)	(0.55)	(0.50)
ST	−40.46***	−11.91	−1.20*	−0.69	−3.47***	−1.36**
	(11.40)	(14.26)	(0.69)	(0.92)	(0.51)	(0.60)
OBC	−20.38**	−4.03	−0.71	−0.37	−1.86***	−0.60
	(8.65)	(12.15)	(0.58)	(0.77)	(0.45)	(0.43)
Mean	−45.5	−45.5	−6.2	−6.2	−6.2	−6.2
R Square	0.02	0.13	0.00	0.16	0.69	0.77
Weighted N	33133	33133	33133	33133	33133	33133
Real N	836	836	836	836	836	836

The results are weighted by numbers of individuals in each district-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.

data are meaningful beyond just a particular sample at a particular moment in time.[8]

8. Limitations

Our empirical analysis faces limitations, including our interpretation of results and our ability to estimate our models with sufficient precision.

First, to interpret our decomposition results as reflecting exact, true underlying contributions of each variable group, we would need a fully specified structural model of child health and measures of all the relevant inputs. We provide only a stylized growth model with three sets of key determinants, along with the relevant measures for such determinants that are observable in our data. In cases where our variable groups have rich information on key determinants, as for maternal health or household wealth, our estimates may be more reasonable. However, our measurement and empirical definition of public health inputs are both unorthodox and imperfect. We acknowledge that a large number of public health determinants of child HAZ may remain unmeasured and thus not accounted for in our decomposition estimates.

Second, data requirements for α/β slope and intercept models are high for both the number of geographic cells and the number of observations required per cell. This creates two important limitations for our work. The first is that sub-group estimates of slope and intercept coefficients we provide are imprecise and provide less insight than we would hope. The second is that we are unable to statistically decompose the changes in the intercept and slope coefficients as we can with the fixed-effects estimates. This is due to two factors. There is the large number of covariates, including state-urban fixed-effects, relative to the number of district-urban cells. But more fundamentally, these estimates rely almost exclusively on within-cell variation; they are designed to compare two HAZ-age profiles from the same region but from different caste groups. This puts tremendous importance on the locational fixed-effects, and makes it difficult to pin down the effects of other important covariates, including and in particular those associated with public health.

9. Conclusion

Child HAZ gaps across caste groups in India are significant. In line with previous literature, we confirm that caste HAZ gaps remain large and are largely explainable by observable differences in households across caste groups (Coffey et al., 2019; LoPalo et al., 2019; Van de Poel & Speybroeck, 2009). Our framework, though, reveals new features of these caste gaps. We document that these caste gaps are present at birth and grow in size, particularly over the first two years of life. Our decomposition results suggest that observable characteristics of households related to both health endowments and (private) health investments matter significantly in explaining the differential child growth patterns across caste groups, and do so 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 in explaining those gaps as children age. By the age of 5, the two variable groups each explain around half of the HAZ disparities across caste groups. These patterns are remarkably similar in both the fourth and fifth waves of the NFHS.

Our findings imply two things. First, 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 (both economic and social). Second, 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 do not change after birth, and remedying the health endowment disparities across castes is likely to be a multi-generational process.

CRediT authorship contribution statement

Joseph Cummins: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Jingyan Guo:** Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Neha Agarwal:** Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Anaka Aiyar:** Formal analysis, Investigation, Project administration, Supervision, Methodology, Visualization, Validation, Writing – original draft, Writing – review & editing, . **Vaishali Jain:** Investigation, Writing – original draft, Writing – review & editing. **Andrew Bergmann:** Writing – original draft, Writing – review & editing.

Declaration of competing interest

We declare we have no conflicts of interest

Data availability

Data for this paper is available through the IPUMS-DHS (https://www.idhsdata.org/idhs/) and Demographic Health Survey (https://dhsprogram.com/data/available-datasets.cfm) websites for download. Registered users can download the National Family Health Survey of India for 2015-16 (NFHS-4) and 2019-20 (NFHS-5) for free. Stata do-files that can replicate all figures and tables produced in the paper are hosted on the open science framework website (www.osf.io).

[8] We also present the sub-group specific replication results in the Online Supplement for completeness on www.osf.io. In summary, we see that the patterns are broadly similar across subgroups, although relative to results from the NFHS-4, the results for girls more closely resemble the results for boys. The results for the rural sample are again similar to the full sample results, but urban sample continues to generate divergent patterns, again likely due to the generally less severe growth faltering experienced by children in urban areas. The results across states with high or low concentrations of UC children both remain consistent with the aggregate results in both survey rounds, with the high UC states again generating noisy results due to the much smaller sample size.

Acknowledgments

We would like to thank the editor, Nidhiya Menon and two anonymous reviewers whose comments helped us clarify ideas and contributions of the paper. We would also like to thank participant comments received at the University of Nevada, Reno (2021), Applied Economics Seminar at University of California, Riverside (2022), Pacific Conference for Development Economics (PacDev) (2022), Western Economic Association International (WEAI) Annual Conference (2022), Economics of Discrimination and Disparities Seminar (2022), Global Health Economics Workshop (2022), Southern Economics Association (SEA) Annual Conference (2022), and Annual Conference on Economic Growth and Development (2023). Any errors are the authors' own. This research did not receive any specific grant or financial support from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.worlddev.2024.106598.

Supplementary data including stata codes to replicate this analysis and analysis from the NFHS 5 can be found at - https://osf.io/5dge7/.

References

Addo, O. Y., Stein, A. D., Fall, C. H., Gigante, D. P., Guntupalli, A. M., Horta, B. L., et al. (2013). Maternal height and child growth patterns. *The Journal of Pediatrics*, *163*(2), 549–554.

Ahmed, S., & 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., & 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*, Article 102569.

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

Alderman, H., & 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), Article e0195904.

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

Biswas, S., Pramanik, K. R., & Sonowal, C. (2023). Marginalised social groups differentials in nutritional status (BMI) among reproductive-aged women in West Bengal. *BMC Public Health*, *23*(1), 1–18.

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

Bora, J. K., Raushan, R., & Lutz, W. (2019). The persistent influence of caste on under-five mortality: Factors that explain the caste-based gap in high focus Indian states. *PLoS One*, *14*(8), Article e0211086.

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., & 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., & Fatima, F. (2017). The causal effect of maternal age at marriage on child wellbeing: Evidence from India. *Journal of Development Economics*, *127*, 42–55.

Choudhury, S., Shankar, B., Aleksandrowicz, L., Tak, M., & Dangour, A. (2021). Caste-based inequality in fruit and vegetable consumption in India. *Food and Nutrition Bulletin*, *42*(3), 451–459.

Coffey, D., Deshpande, A., Hammer, J., & 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.

Danaei, G., Andrews, K. G., Sudfeld, C. R., Fink, G., McCoy, D. C., Peet, E., et al. (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), Article e1002164.

Daysal, N. M., Trandafir, M., & 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., & Branca, F. (2016). Childhood stunting: a global perspective. *Maternal & Child Nutrition*, *12*, 12–26.

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

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

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

Dewey, K. G., & Begum, K. (2011). Long-term consequences of stunting in early life. *Maternal & Child Nutrition*, *7*, 5–18.

Dommaraju, P., Agadjanian, V., & Yabiku, S. (2008). The pervasive and persistent influence of caste on child mortality in India. *Population Research and Policy Review*, *27*, 477–495.

for Population Sciences (IIPS), I. I. I., & ICF (2021). [Dataset:] national family health survey (NFHS-4) 2015-16 [dataset]. data extract from IAIR72.sav, IAKR72.sav, IABR72.sav, IAMR74.sav, and IAPR74.sav. IPUMS demographic and health surveys (IPUMS DHS), version 7, IPUMS and ICF [distributors]. Accessed from http://idhsdata.org on June 2021.

Geruso, M., & Spears, D. (2018). Neighborhood sanitation and infant mortality. *American Economic Journal: Applied Economics*, *10*(2), 125–162.

Godlonton, S., & 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(2). *Journal of Political Economy*, *223*(10.2307), Article 1830580223.

Hamal, M., Dieleman, M., De Brouwere, V., & de Cock Buning, T. (2020). Social determinants of maternal health: a scoping review of factors influencing maternal mortality and maternal health service use in India. *Public Health Reviews*, *41*(1), 1–24.

Islam, A., Pakrashi, D., Sahoo, S., Wang, L. C., & Zenou, Y. (2021). Gender inequality and caste: Field experimental evidence from India. *Journal of Economic Behaviour and Organization*, *190*, 111–124.

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., & 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*, 239–254.

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

Mahadevan, R., & Suardi, S. (2013). Is there a role for caste and religion in food security policy? A look at rural India. *Economic Modelling*, *31*, 58–69.

Maity, B. (2017). Comparing health outcomes across Scheduled Tribes and Castes in India. *World Development*, *96*, 163–181.

Mertens, A., Benjamin-Chung, J., Colford Jr, J. M., Coyle, J., van der Laan, M. J., Hubbard, A. E., et al. (2023). Causes and consequences of child growth faltering in low-resource settings. *Nature*, *621*(7979), 568–576.

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

Munshi, K., & 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.

Olofin, I., McDonald, C. M., Ezzati, M., Flaxman, S., Black, R. E., Fawzi, W. W., et al. (2013). Associations of suboptimal growth with all-cause and cause-specific mortality in children under five years: a pooled analysis of ten prospective studies. *PLoS One*, *8*(5), Article e64636.

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

Panda, P. (2020). Selective mortality and malnutrition in India. *Journal of Quantitative Economics*, *18*(4), 861–890.

Parappurathu, S., Kumar, A., Bantilan, M., & Joshi, P. (2015). Food consumption patterns and dietary diversity in eastern India: evidence from village level studies (VLS). *Food Security*, *7*, 1031–1042.

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

Prendergast, A. J., & Humphrey, J. H. (2014). The stunting syndrome in developing countries. *Paediatrics and International Child Health*, *34*(4), 250–265.

Raushan, R., Acharya, S. S., & Raushan, M. R. (2022). Caste and socioeconomic inequality in child health and nutrition in India. *CASTE: A Global Journal on Social Exclusion*, *3*(2), 345–364.

Rieger, M., & 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., & 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.

Saikia, N., Bora, J. K., & Luy, M. (2019). Socioeconomic disparity in adult mortality in India: estimations using the orphanhood method. *Genus*, *75*(1), 1–14.

Sanneving, L., Trygg, N., Saxena, D., Mavalankar, D., & Thomsen, S. (2013). Inequity in India: the case of maternal and reproductive health. *Global Health Action*, *6*(1), 19145.

Shrimpton, R., Victora, C. G., de Onis, M., Lima, R. C., Blossner, M., & Clugston, G. (2001). Worldwide timing of growth faltering: implications for nutritional interventions. *Pediatrics*, *107*(5), e75.

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*, Article 102277.

Spears, D., & 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., & 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., & Shah, N. G. (2019). Does the relationship between income and child health differ across income groups? Evidence from India. *Economic Modelling*, *79*, 57–73.

Thorat, S., & Sadana, N. (2009). Discrimination and children's nutritional status in India. *IDS Bulletin*, *40*(4), 25–29.

Uddin, J., Acharya, S., Valles, J., Baker, E. H., & Keith, V. M. (2020). Caste differences in hypertension among women in India: diminishing health returns to socioeconomic status for lower caste groups. *Journal of Racial and Ethnic Health Disparities*, *7*, 987–995.

Van de Poel, E., & Speybroeck, N. (2009). Decomposing malnutrition inequalities between Scheduled Castes and Tribes and the remaining Indian population. *Ethnicity & Health*, *14*(3), 271–287.

Vart, P., Jaglan, A., & Shafique, K. (2015). Caste-based social inequalities and childhood anemia in India: results from the National Family Health Survey (NFHS) 2005–2006. *BMC Public Health*, *15*, 1–8.

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

Von Grafenstein, L., Klasen, S., & Hoddinott, J. (2023). The Indian Enigma revisited. *Economics & Human Biology*, *49*, Article 101237.

Vyas, S., Hathi, P., & Gupta, A. (2022). Social disadvantage, economic inequality, and life expectancy in nine Indian states. *Proceedings of the National Academy of Sciences*, *119*(10), Article 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., & Vakulabharanam, V. (2011). Caste stratification and wealth inequality in India. *World Development*, *39*(10), 1820–1833.