Department of Economics

# DISCUSSION PAPER SERIES IN ECONOMICS

DP No. 27

# Which Indian Children are Short and Why? Social Identity, Childhood Malnutrition and Cognitive Outcomes

March 2020

Ashwini Deshpande
Rajesh Ramachandran

https://www.ashoka.edu.in/ecodp

Plot No. 2, Rajiv Gandhi Education City,
National Capital Region
P.O.Rai,Sonepat Haryana-131029

# Which Indian Children are Short and Why? Social Identity, Childhood Malnutrition and Cognitive Outcomes*

Ashwini Deshpande† Rajesh Ramachandran‡

March 31, 2020

**Abstract**

This paper addresses four previously under-explored facets related to early childhood malnutrition, as manifested in stunting. First, it provides a new explanation for the "Indian Enigma": why Indian children are shorter than those from poorer countries. Our analysis, based on data for over 213,000 children, finds that the Indian height deficit is driven by the shorter height-for-age of children from the lower-ranked and stigmatised castes. Upper-caste children are taller than the average for low middle income countries, but stunting in the lower-ranked caste groups is 10-15 percentage points higher. This is true even after controlling for variables used in earlier explanations in the literature, viz., birth order and disease environment. Second, using repeated cross sections of national level data for India, we show that between 1998 and 2016 caste gaps in childhood stunting did not decline. Third, using longitudinal data from the Young Lives Project, we estimate the probabilities of transition in and out of stunting at various ages, conditional on being stunted at age 1. We show that children stunted at age 1 are 22% points more likely to be more likely to be stunted at age 15. Finally, we show that stunting at age 1 year significantly lowers cognitive and learning outcomes at ages 5, 8, 12 and 15 years. We explore the role of the enduring stigmatisation of lower-castes as the reason for the persistence of caste gaps. Our results underscore the importance of focusing on social identity and discrimination as key markers of malnutrition, and indicate that the origins of adult life disparities between caste groups lie in early childhood.

Keywords: caste, stunting, early childhood development, cognition, discrimination
JEL: I12, J13, O15, Z12, Z13

*We would like to thank Gaurav Datt, Andreas Grunewald, Karen Macours, Devesh Rustagi, Anisha Sharma and Alessandro Tarozzi for helpful comments and suggestions. We are also grateful to conference participants at Yale University; Indian Statistical Institute, New Delhi; Goethe University; International School of Business, Hyderabad; Summer School in Development Economics, Prato; and German Development Association Meeting, Gottingen, where earlier versions of this paper were presented, for their suggestions and useful comments. We are responsible for all remaining errors and omissions.

†Ashoka University, Department of Economics, Rajiv Gandhi Education City, Sonipat, Haryana 131029, India. ashwini.deshpande@ashoka.edu.in

‡Heidelberg University, Faculty of Economics, Bergheimer Straße 58, Heidelberg, Germany. rajesh.ramachandran@awi.uni-heidelberg.de

# 1 Introduction

There is growing and compelling evidence on the impact of early childhood conditions on a variety of later life outcomes, including health, nutrition, cognition and mortality. (Victora et al., 2008; Case and Paxson, 2010; McGovern et al., 2017). Of particular interest has been the phenomenon of stunting, which is defined as "impaired growth and development that children experience from poor nutrition, repeated infection, and inadequate psychosocial stimulation."[1] Stunting among children is widespread, with one in four children worldwide classified as stunted (UNICEF, 2018). Thus, it remains a global health challenge, because of its adverse consequences for later life morbidity and mortality risk, non-communicable diseases, learning capacity and productivity.

India exhibits high prevalence of stunting, being home to nearly a third of all stunted children. This makes the country an outlier among countries at comparable levels of development. The "Asian enigma" (Ramalingaswami et al., 1996) of high rates of stunting in Asia compared to poorer countries has been explored by a number of studies in the literature (Deaton, 2007; Tarozzi, 2008; Jensen, 2012). Recent work suggests that this is more of an Indian than an Asian enigma (Headey et al., 2015).

In this context, our paper addresses four aspects of early childhood stunting in India. First, it provides a new explanation for the "Indian Enigma": why Indian children are shorter than those from poorer countries, which highlights the salience of social identity, specifically caste in India. Second, using repeated cross sections of national-level data, we show that between 1998 and 2016 caste gaps in childhood stunting did not decline. Third, using longitudinal data from the Young Lives Project, we estimate the probabilities of transition in and out of stunting at various ages, conditional on being stunted at age 1. Finally, we estimate the impact of stunting at age 1 on cognitive and learning outcomes at ages 5, 8, 12 and 15 years. We explore the role of the enduring stigmatisation of lower-castes as the reason for the persistence of caste gaps.

To the best of our knowledge, this investigation of multiple facets of stunting, highlighting the salience of traditional social identities, is a major contribution to the literature. Our analysis underscores the continued relevance of caste in understanding contemporary inequalities in Indian society, and sheds light on how traditional social hierarchies not only tend to persist, but can in fact become more salient, despite globalization, marketization and access to political power by lower-ranked groups. The early childhood onset of disparities, and the continuing presence of deeply stigmatizing, discriminatory practices and their negative effect on health and learning outcomes, suggests that multi-faceted policy interventions that combine affirmative action with effective early childhood intervention might be required for the elimination of deep-rooted social hierarchies.

[1] https://www.who.int/nutrition/healthygrowthproj_stunted_videos/en/

### 1.1 The Role of Social Identity in India's Child Height Deficit

There are three broad explanations for the unusually high levels of stunting in India. One explanation stresses the role of birth order gradient driven by son preference (Jayachandran and Pande, 2013, 2017). This argues that India's steeper birth order gradient in child height can account for over one-half of the India-Africa gap in child height. The authors show that first-born Indian children are not significantly shorter than children in sub-Saharan Africa (SSA). The relative height disadvantage materializes for second-born children in India, and increases for third and higher order births. The second major alternative explanation focuses on the disease environment due to high rates of open defecation combined with high population density (Hammer and Spears, 2016; Coffey and Spears, 2017; Spears, 2018). This strand of the literature argues that higher rates of open defecation can account entirely for why Indian children are shorter than African children from poorer countries. The evidence of height disadvantage materializing only for higher order births also raises a serious question mark against a third argument, such as that advanced by Panagariya (2013), that the Indian children are shorter than African children because of genetic reasons.

We advance a fourth explanation, one that has been surprisingly overlooked in the literature, *viz.,* the role of social identity manifested through the caste system in India. The literature on childhood stunting discusses the role of caste primarily in the context of how caste-based notions of purity and pollution affect toilet usage and sanitation behavior. These beliefs account for high rates of open defecation, which in turn leads to higher rates of stunting. However, we argue that caste hierarchy defines stunting more fundamentally, in that lower-ranked groups are more likely to be stunted. We demonstrate that the caste gap in childhood stunting needs to be unpacked to fully understand the puzzle posed by the "Asian Enigma". To the best of our knowledge, our paper is the first to demonstrate this.

Figure 1 provides important motivating evidence for the importance of accounting for caste in the analysis of childhood stunting in India. The figure plots the percentage of under-five children stunted for 115 countries clubbed by income level – high, upper middle, lower middle and low. It shows that 38.5 percent children in India, a lower middle-income country, are stunted. This is more than 10 percentage points, or 33 percent higher, than the average for the lower middle-income group countries.

Figure 1 also shows the incidence of childhood stunting for three broad caste groups in the country – Hindu upper castes, Other Backward Classes (OBCs) and the Scheduled Castes and Scheduled Tribes (SC-ST).[2] The percentage of upper caste (UC) children who are stunted is 28.2, which

[2] India's caste system consists of several thousand *jatis* or castes. The formerly untouchable castes and several marginalized tribal communities are the most socioeconomically disadvantaged groups in the country. Thesse groups receive preferential affirmative action, for the purpose of which these are listed in a government schedule. Thus, this collection of castes and tribes are referred to as the Scheduled Castes and Scheduled Tribes (SC and ST). While these are official, administrative categories, SCs often self-identify themselves as *Dalit*, meaning "oppressed", as a term of

is slightly below the average for the lower middle-income group of countries. However, stunting proportions for OBC and SC-ST children are 37 and 43.1 percent, respectively. Thus, the level of stunting among Hindu UC children is identical to that in lower middle income countries, but the incidence among SC-ST and OBCs are around 10 and 15 percentage points higher, respectively.

Figure 2 provides further evidence on the importance of focusing on caste differences in order to understand the dynamics of stunting in India. We plot the height-for-age z-score (HAZ) curves in Figure 2 for 30 countries in sub-Saharan Africa (SSA), for all-India, and separately for SC-ST and UC children in India. We see that the mean child's HAZ is higher at all ages for SSA as compared to India. However, the lower HAZ in India driven by lower height of SC-ST children: whereas the HAZ curve for UCs is higher at all ages relative to children in SSA, for SC-ST children, the opposite is true. Thus, Figure 1 and 2 suggest that the Indian anomaly is largely due to the extraordinarily high rates of stunting among the lower caste groups. We argue that in order to understand the problem of stunting in India, we need to explicitly recognise the importance of social identity on life outcomes.

## 1.2 Main Results

### 1.2.1 Evolution of Caste Gaps in Stunting

We trace the evolution of caste gaps in stunting over time based on three rounds of nationally representative data from the National Family and Health Survey (NFHS) for 1998-99, 2005-06 and 2015-16, for children aged 0-59 months. Pooled data over the three rounds reveal that on average, 32 percent of upper caste Hindu children are stunted; SC-ST and OBC children are 13 and 7.2 percentage points (or 40 and 23 percent), respectively more likely to be stunted. The evolution of gaps in stunting over the 18-year period reveals that the caste gaps have either remained static, or increased (comparing UCs and OBCs). These gaps are not fully accounted for by the conventional socioeconomic status (SES) factors: accounting for a wide range of observable SES characteristics – mother's education, health inputs at birth, mother's health status, mother's birth history, household wealth and sanitation – reduces the raw gaps only by half. These gaps are clearly present even when we restrict the sample to mothers at either end of the SES spectrum: whose education or wealth is above the 95th percentile or below the 5th percentile. This further confirms that conventional SES factors cannot explain the entirety of caste disparities in stunting.

pride. *Adivasi* means "indigenous people" and is often used to describe tribal communities. A group of intermediate to low-ranked castes and communities, which also receive affirmative action, are called the "Other Backward Classes" (OBCs). All those not classified as SC, ST or OBC are the residual "Others" category. The caste system can be seen in all major religions in India, but we restrict our sample to Hindu individuals, as Hinduism has the most formalised caste system. Thus, in our case, the "Others" category captures Hindu Upper Castes (non-SC-ST-OBC Hindus).

### 1.2.2 Transition in and out of stunting

As the literature recognises, stunting has far-reaching consequences, since it represents a state of chronic malnourishment with long-term implications for cognition and life outcomes. NFHS data do not allow us to track the later life consequences of early childhood stunting. In particular, there are two key questions of interest that arise. One, are the documented caste gaps transient, i.e. are they likely to become smaller as children grow older, or are they persistent over time? Two, how does early childhood stunting affect later life cognition and educational outcomes? To answer these questions, we use data from the Young Lives Project. This is an international study of childhood poverty; in India, it follows the lives of 3000 children in the states of Andhra Pradesh and Telangana between 2002 to 2016. This gives us longitudinal data on children at ages 1, 5, 8, 12 and 15 years, enabling us to explore the longer-term patterns in, as well as consequences of, stunting.

Our analysis with panel data shows that caste gaps in stunting are durable and remain as large at age 15 years as they were at 1 year. The percentages of UC children stunted at ages 1, 5, 8, 12 and 15 years are 16.6, 23.4, 14.8, 16.6 and 15.4, respectively. The corresponding percentages of SC-ST children are 41.2, 40.2, 35, 33.6 and 33.4, respectively. Thus, caste gaps persist into teenage years, and if anything, are larger at 15 years than at 5 years. As we are able to track individuals over time, our analysis sheds light on the dynamics of caste gaps in stunting. We find that, one, caste gaps in stunting at 1 year are persistent and responsible for a large share of caste gaps at older ages. Two, conditional on being stunted at age 1, SC-ST children are more likely to remain stunted at later ages compared to UC children. Three, conditional on *not* being stunted at age 1, OBC children are more likely to transit into being stunted at later ages, relative to UC children.

### 1.2.3 Consequences of Stunting on Cognitive and Learning Outcomes

We then investigate the implications of being stunted at age 1 on cognitive outcomes at ages 5, 8, 12 and 15. We find that stunting at age 1, after accounting for an extensive set of controls, reduces the standardized score on the cognitive development assessment (CDA) test at age 5 by 0.18 standard deviation units, and increases the probability of being in the bottom quartile of the distribution by around 33 percent. For ages 8 and above, we estimate value-added regressions which not only account for a rich set of socioeconomic controls, but also for past achievement. The value-added model at age 8 shows that stunting at age 1 reduces the probability of being literate by about 20 percent and the number of correct words read per minute by 13 percent. It also reduces the standardized math test score by 0.28 standard deviation units and the probability that the child correctly can subtract 4 from 20 by 8.4 percent. The effects of being stunted at age 1 persist at ages 12 and 15. At age 12, even after accounting for learning outcomes at age 8, the effect of stunting at age 1 remains statistically significant and economically meaningful, reducing test scores in reading, Mathematics, English and Peabody Picture Vocabulary Tests (PPVT) by 0.20, 0.17, 0.19 and 0.12

units of a standard deviation. Finally, at age 15, the cumulative effect of stunting at age 1 is to reduce language and math scores by 0.28 and 0.25 units of a standard deviation.

### 1.3 Main Contributions

Our paper extends the existing literature in several ways. First, it highlights the importance of accounting for social identity, specifically caste in the Indian context, to understand and address the challenge of childhood stunting. The small literature on differences in stunting by caste has typically only explored the differences between two broad groups of SC-ST and "Others" (everyone else), whereas we have a more nuanced three-way classification, including intermediate castes. Existing studies also do not distinguish between religion, thus "Others" includes all religious groups, and is not a good proxy for Hindu UCs (the group of castes conventionally regarded as the top of the socioeconomic pyramid). Hence, the existing literature understates the extent and importance of caste gaps (for instance, see Subramanyam et al. 2010). A recent notable exception is Coffey et al. (2019), which focuses on caste gaps in stunting in rural India. However, the purpose of this paper is to understand the role of local social hierarchies in affecting differences in child height across caste groups. The paper uses data from one round of NFHS (2005-06) and is unable to comment on how gaps change over time. In this paper, we present a nuanced and detailed picture of caste gaps in stunting over a period of 18 years based on nationally representative data.

Second, using longitudinal data on 1497 children, our paper also sheds light on the factors shaping the transition in and out of stunting. We estimate the probabilities of entry and exit into stunting, conditional on being stunted at age 1 year, and find that children stunted at age 1 year are 22% points more likely to be stunted at age 15. Thus, caste gaps in stunting at age 1 persist all the way into teenage years.

Third, we contribute to the literature on the long-term effects of childhood malnutrition in general, and stunting in particular, on cognitive and educational outcomes. The bulk of the literature linking early childhood development (ECD) to later life outcomes is based on data from developed countries. Ours is one of the few studies from a large emerging economy[3]. Regardless of the estimation method and data source, this literature indicates that poor health in childhood translates into poor cognitive and non-cognitive abilities in later life, and affects educational and occupational attainment, as well as wages.

Our study is restricted to measuring effects of childhood stunting on human capital measures[4] until

[3] Victora et al. (2008) present evidence from longitudinal studies in Brazil, Guatemala, India and South Africa. Studies for developing countries are few, possibly because of the paucity of good quality and consistent panel data. However, researchers have used panel data where available, such as in Zimbabwe (Alderman et al., 2006) or Guatemala (Maluccio et al., 2009). Elsewhere, we find the use of innovative estimation techniques in the absence of panel data, such as evidence from natural experiments, e.g. Yemenite children in Israel (Gould et al., 2011), or quasi-natural experiments, e.g. effect of a malaria eradication program in India (Cutler et al., 2010).

[4] It is important to recognise that in India nationwide, almost 30 percent of children in Grade 8 cannot read at

the age of 15, and thus cannot shed light on labor market outcomes or fertility decisions. This is due to the nature of data availability. We would like to emphasise that the evidence presented in our paper is the best possible and most comprehensive evidence yet for India. We are able to use a range of outcomes such as literacy, math, language, reading scores etc. to better measure the extent of gaps in learning arising due to chronic malnutrition in early childhood. This is based on a subnational dataset as the Young Lives Project covers two states of India, Andhra Pradesh and Telangana. However, the sample is not small, as these two states have a combined population greater than that of Germany.

Our results suggest that if India is to reap its potential demographic dividend, the challenge of chronic malnutrition needs to be addressed urgently. The salience of caste differences over and above conventional SES factors indicates that any serious solution of the stunting problem cannot sidestep the enduring power of caste hierarchies. In order to understand the mechanisms underlying this result, we explore the potential role of enduring societal discrimination and the stigmatizing practice of untouchability as potential factors that need to be investigated rigorously in the future (Barbour et al., 2007; George, 2015; Human Rights Watch, 2014; Coffey et al., 2019). The data from the Indian Human Development Survey (IHDS) show that the illegal practice of untouchability continues to be widely practiced, with more than 40% respondents in several states openly admitting practicing it. We show that the caste gaps in stunting between UC and SCs are sharply increasing in the self-reported incidence of untouchability. The fact that stunting outcomes of SC children, but not UC children, are correlated with the practice of untouchability indicates that the stigmatisation that accompanies this highly discriminatory practice contributes to SC children being more stunted than UC.

The remainder of the paper is organized as follows: Section 2 provides evidence on the evolution of caste gaps in stunting for the period 1998-2016, and explores the importance of socioeconomic factors in explaining these gaps. Section 3 uses panel data to understand the persistence of height deficits between caste groups until the age of 15, and the dynamics of entry and exit into stunting status. Section 4 provides evidence about the effect of stunting at age 1 on cognitive and learning outcomes at ages of 5, 8, 12 and 15. Section 5 discusses the association between untouchability and stunting, and Section 6 discusses the findings and concludes.

## 2 Evolution of stunting and caste gaps

To explore the extent and evolution of caste gaps in child stunting, we employ data from three rounds of the National Family and Health Survey (NFHS): NFHS-II-1998-99, NFHS-III-2005-06

the level of Grade 2, and thus years of schooling is not a reliable way of measuring human capital (ASER, 2018).See http://img.asercentre.org/docs/ASER%202018/Release%20Material/aser2018nationalfindingsppt.pdf for further details.

and NFHS-IV-2015-16. To gauge the full extent of caste gaps, we restrict our analysis to the religious denomination of Hindus for whom caste is a strong and codified marker of social stratification. We first calculate anthropometric z-scores using the 2006 WHO child growth standards for height-for-age (stunting).[5] Following international convention, we classify children with z-scores less than -2, i.e. two standard deviations below the reference median of height-for-age, as stunted. The proportion of stunted children is 51, 47 and 38 percent in 1999-99, 2005-06 and 2015-16 respectively.

Panel A of Figure 3 shows how the aggregate figure masks very sharp differences in stunting between caste groups. In 1998-99, 58 percent of the SC-ST children were stunted compared to 45 percent of UC children. The corresponding figures for the next two rounds are 53 and 37 percent in 2005-06 and 43 and 28 percent in 2015-16. In 2015-16, stunting among UC children is 10 percentage points lower than the national average, whereas for SC-ST children, the incidence is 5 percentage points above the national average, resulting in a gap of 15 percentage points. In other words, SC-ST children are 50 percent more likely to be stunted than UC children.[6]

Panel B of Figure 3 shows what stunting gaps would look like if we clubbed OBCs with UC. The gap in stunting between the two groups, based on a binary division of SC-ST and non SC-ST, reduces to around 7.5 percentage points. Thus, aggregating caste categories halves the extent of relative disadvantage faced by the SC-ST. This highlights the importance of accounting for the OBCs, who comprise around 45 percent of the Hindu population, as a separate group in analyzing caste gaps. Aggregating OBCs and UCs into one category would lead to an overestimation of convergence among the caste groups in India. For instance, the results of convergence on a range of indicators from the Hnatkovska et al. (2012) study, which compares SC-ST to non-SC-ST, are not robust when OBCs are accounted for separately, as in the Deshpande and Ramachandran (2019) study.

To formally estimate the incidence and evolution of caste gaps between the three social groups over the three rounds, we estimate an Ordinary Least Squares (OLS) regression given by:

$$\begin{aligned} O_{ikr} = \psi_0 + + \delta_1 SCST_i + \delta_2 OBCs_i + \beta_1 SCST_i * NFHS3 + \beta_2 SCST_i * NFHS4 \\ + \beta_3 OBCs_i * NFHS3 + \beta_4 OBCs_i * NFHS4 + \gamma_1 Rural_i + \gamma_2 State_k + \gamma_3 NFHS_r + \epsilon_{ikr}, \quad (1) \end{aligned}$$

[5] Refer to http://siteresources.worldbank.org/INTPAH/Resources/Publications/459843-1195594469249/HealthEquityCh4.pdf for further details. We use the zscore06 package in Stata to calculate the height-for-age z-score (Leroy, 2011).

[6] Figure 13 in the appendix shows the evolution of stunting by caste and six different regions, It shows that there are large regional differences with the group classified as BIMARU - Bihar, Chhattisgarh, Jharkhand, Madhya Pradesh, Rajasthan, Uttar Pradesh and Uttaranchal - showing the highest rates of stunting and the region classified as North East - Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim and Tripura - showing the lowest stunting levels.

where $O_{ikr}$ is either the height-for-age-Z (HAZ) score or a dummy that takes value one if child $i$ resident in state $k$ and from NFHS round $r$ is stunted. $\delta_1$ and $\delta_2$, two of the principal coefficients of interest, capture the average pooled difference between UC and SC-STs and UC and OBCs, respectively. $\beta_1$ and $\beta_2$ capture how the gaps between the upper castes and SC-ST changes from 1998-99 to 2005-06 and 2015-16, respectively, whereas, $\beta_3$ and $\beta_4$ do the same for the comparison between the upper castes and OBCs. We additionally account for a set of exogenous controls that might have a bearing on child height, namely, age in months and sex of the child. Additionally, we include dummies for rural residence, the NFHS round and 36 dummies for the 29 states and 7 union territories in India to account for time trends, as well as state level differences and cluster standard errors at the state level.

Table 1 shows the results of the estimation based on a sample of 213,665 children. Columns (1) and (2) show the results for the regression with the dependent variable HAZ, whereas columns (3) and (4) show the results where the dependent variable is a dummy for stunting. The coefficients on the SC-ST and OBCs dummy confirm the raw gaps shown in Panel A of Figure 3. Column (3) shows that the SC-ST and the OBCs are 13 and 7.2 percentage points more likely to be stunted as compared to the upper castes, in other words, the incidence of stunting is 40 and 23 percent higher among the SC-ST and OBCs respectively.

Column (4) of Table 1 also shows the interaction of the caste dummies with the round dummies to see how gaps have evolved over the 18 year period. $\beta_1$ and $\beta_2$ are positive, though statistically insignificant, suggesting that gaps in stunting between the SC-ST and UC have remained unchanged; if anything, have widened. The coefficients $\beta_3$ and $\beta_4$ are positive and significant, showing that the relative incidence of stunting among the OBCs compared to the UCs increased by 4.3 and 4.5 percentage points respectively over the two rounds. The macro evidence shows that not only are caste gaps in stunting large in magnitude, but these remain either persistent or increase over the 18 year period between 1998-99 and 2015-16.

## 2.1 Socioeconomic factors and caste gaps in stunting

To understand whether observable socioeconomic characteristics are associated with caste gaps in stunting, we make use of the rich information on the socioeconomic characteristics available in the NFHS. Empirically, we re-estimate Equation 1 but sequentially introduce controls. We introduce the following set of controls:

1. Estimates labeled W2 account for the years of education of the mother.
2. Estimates labeled W3 account for the health inputs at birth - (i) place of delivery; (ii) prenatal access to doctor; (iii) breastfed immediately after birth
3. Estimates labeled W4 account for the mother's health status - (i) weight; (ii) height; (iii)

height-for-age percentile; (iv) body mass index; (v) Rohrer index[7]; (vi) age; (viii) age at first marriage.

4. Estimates labeled W5 account for the mother's birth history - (i) number of children who have died; (ii) total children born; (iii) age at first birth; (iv) birth order of child.
5. Estimates labeled W6 accounts for the factor score of the wealth index of the mother's household.[8]
6. Estimates labeled W7 account for sanitation proxied by - (i) source of drinking water; (ii) type of toilet facility.
7. Estimates labeled W8 include all the above specified controls (W2-W7) in a single regression.

The results of this exercise are shown in Figure 4. In order to enable comparison of the coefficients across different specifications, we restrict the sample to children for whom information on all the above covariates is available. This results in a total sample of 157,467 children. The estimates labeled W1 are the caste gaps resulting from re-estimating equation 1 on the sample of 157,467 children, and provides the reference point to evaluate the importance of various characteristics in muting these gaps. Panel A of Figure 4 shows the coefficient on the SC-ST dummy, and Panel B shows the coefficients on the OBC dummy.

Estimates W1, similar to Table 1, show that SC-ST are 13 percentage points more likely to be stunted. Accounting for mother's years of education (W2) reduces the gap to 8 percentage points. Estimates labeled W3, which account for the three health inputs at birth, reduces the gap to 11 percentage points, though they are not statistically different from the baseline estimates. Accounting for mother's health status (W4) reduces the stunting gap to 7.8 percentage points, whereas accounting for the birth history (W5) does not result in a statistically different gap compared to the baseline estimate. Accounting for the factor score of the wealth index of the mother's household (W6) reduces stunting gaps to 7.7 percentage points, and this is statistically different from the baseline estimate. Finally, accounting for the source of drinking water and type of toilet facility reduces gaps to around 10 percentage points (W7), though these are not statistically different from the baseline estimates. Finally, estimates labeled W8 include all the specified controls (W2-W7) in a single regression; this reduces the coefficient on the SC-ST dummy to 0.036 or by about three-fourths compared to the baseline estimates.

[7]The Rohrer index is used to describe the relationship between weight (body mass M) and height (L) in order to allow categorisation of subjects according to obesity. It is similar to the body mass index, but the mass is normalized with the third power of body height rather than the second power.

[8]The wealth index is a composite measure of a household's cumulative living standard. The wealth index is calculated using easy-to-collect data on a household's ownership of selected assets, such as televisions and bicycles; materials used for housing construction; and types of water access and sanitation facilities. See https://www.dhsprogram.com/topics/wealth-index/Wealth-Index-Construction.cfm for further detials.

Panel B of Figure 4 shows that accounting for the characteristics outlined above leads to a reduction in the OBC-UC gap by about half compared to the baseline estimate of 6.6 percentage points. However, these differences are not statistically significant. The estimate labeled W8, which include all the controls shows that the gaps between the OBCs and the upper castes reduces to 1.5 percentage points, and this difference is statistically significantly different from the baseline estimate.

## 2.2 The Role of Mother's Wealth and Education

Figure 4 shows that accounting for mother's wealth or education reduces the gaps by around 40 percent for SC-ST children. To better understand the role of these two factors, we restrict our sample in two distinct ways: (i) mothers whose household wealth is above the 95th or below the 5th percentile; and (ii) mothers whose years of schooling is above the 95th or below the 5th percentile. We then re-estimate Equation 1 and the results are shown in Table 2.

Column (1) reproduces the baseline regression for the purpose of comparison. Column (2) shows that even for households whose wealth is above the 95th percentile, the proportion of stunted UC children is 19 percent. In this group, SC-ST children are 5.8 percentage points more likely, or 30 percent more likely to be stunted. Considering households whose wealth is below the 5th percentile in Column (3), the proportion of stunted UC children is 52 percent, with SC-ST children 4.9 percentage points, or 10 percent more likely to be stunted.

Turning to education, within the subsample of mothers whose years of schooling is greater than the 95th percentile (which amounts to having 15 or more years of schooling), Column (4) shows that SC-ST children are 4.4 percentage points or about 23 percent more likely to be stunted than UC children. In column (5), when we consider mothers with years of schooling below the 5th percentile (no education), the 51 percent of UC children are stunted, with SC-ST 5.6 percentage points, or about 10 percent more likely to be stunted. The associations in Table 2 suggest that caste gaps remain important in magnitude, even at the two extremes of the education and wealth distribution. Interestingly and somewhat counter-intuitively, relative gaps are larger at the top end of the distribution than at the bottom.

## 2.3 Summary of Findings

The results in this section show that the burden of child malnutrition, as manifested in stunting, is unequally distributed in India, with children in the socially disadvantaged groups of SC-STs and OBCs being more than 41 and 22.5 percent respectively more likely to be stunted compared to UC children. The evidence from Figure 1 shows that the 'Indian enigma' in stunting is largely an artefact of higher rates of stunting among the socially disadvantaged groups. The incidence of stunting among the upper caste children is very similar to the national average for the lower middle income group of countries, whereas stunting among SC-ST and OBC children is around 10 and 15 percentage points higher than the national average for lower middle income countries. The HAZ curves in Figure 2 show that whereas the mean child's height-for-age is higher at all ages for SSA than India, this does not hold when we compare children in SSA to only UC children. This gap either remains persistent, or increases over the 18 year period of our data. Finally, we see that even after accounting for mother's education, wealth and other household and mother level characteristics the caste gaps in stunting remain substantively important.

# 3 Caste Gaps in Stunting in Later Life

The evidence in Section 2 opens up other important questions. The emphasis on stunting stems from its long-term adverse consequences. To understand the contours of caste gaps in stunting in later-life, it is important to investigate whether, one, do caste gaps in stunting become smaller as children become older? Two, what are the longer term consequences of stunting on cognition and educational outcomes? We now turn to these two questions.

## 3.1 Data

In order to examine the durability of caste gaps, we use data from the Young Lives Project (YLP). The YLP is an international study of childhood poverty following the lives of 12,000 children in Ethiopia, India, Peru and Vietnam over a period of 15 years. In India, it follows 3000 children in the states of Andhra Pradesh and Telangana over 2002-2016. This is a panel data with rounds in 2002, 2006, 2009, 2013 and 2016. The children are divided in two cohorts, a younger and an older cohort. Information on the younger cohort was collected at ages of 1, 5, 8, 12 and 15. For the older cohort, the data was collected at the ages of 8, 12, 15, 18 and 21.

The surveys collects information about a host of child and household characteristics, and most importantly for our purpose, information on stunting status and cognitive and/or learning outcomes. Even though these data are available for only two states in India, it is important to note that the combined population of these two states is around 85 million, which is larger than the second most populous state in Europe, Germany, with a population of around 82 million.

How comparable are these two states to the rest of India? The data shows that the proportion of stunted children under age 6 is 33 percent, thus slightly below the Indian average, but broadly comparable. Second, in these states, 40, 35 and 20 percent of SC-ST, OBC and UC children under age of 6 are stunted, again broadly comparable to the rest of India. Thus, these states are not outliers among Indian states.

### 3.2 The evolution of stunting by caste and age

To evaluate how caste gaps evolve as children get older, we concentrate on the younger cohort for whom the anthropometric measures are available in all rounds. This results in a sample of 1479 children of which 33 percent are SC-ST, 51 percent OBCs and 16 percent are upper castes. We calculate the proportion of children stunted by caste and age, and the results are shown in Figure 5. The figure shows that the percentage of UC children stunted at ages of 1, 5, 8, 11, 12 and 15 years are 16.6, 23.4, 14.8, 16.6 and 15.4, respectively. The comparable percentages of SC-ST children are 41.2, 40.2, 35, 33.6 and 33.4 respectively. The comparable figures for the OBCs are 29.8, 38, 29.6, 29.6 and 28.6 respectively. The data show that caste gaps remain persistent into teenage years, and if anything, are larger at age 15 than at age 5. Thus, the raw gaps show that at the age of 15, the SC-ST are 2.16 times more likely to be stunted as compared to the upper castes, whereas the comparable figure for the OBCs is 1.75.

To evaluate the extent of caste gaps in stunting at each of the age profiles, we estimate the following OLS regression:

$$S_{ikr} = \psi_0 + \delta_1 SCST_i + \delta_2 OBC_i + XBasic_{ikr} + XExtd_{ikr} + \epsilon_{ikr}, \tag{2}$$

where $S_{ikr}$ refers to the outcome of child $i$ resident in cluster $k$ and of age $r$; The coefficients of interest are $\delta_1$ and $\delta_2$, which help us the evaluate the position of SC-ST and OBCs in relation to the upper castes. $XBasic_{ikr}$ refers to the set of basic controls that have a bearing on child height. These include the child's age in month, as well as dummies for gender, rural residence and the sampling cluster to which the child belongs. We also estimate regressions, where we have an extended set of controls denoted by $XExtd_{ikr}$, namely, household size, mother's years of education and a wealth index.[9] Additionally, for children aged 1, we also account in the extended regression for the number of antenatal visits of mother, whether mother was attended by skilled health personnel, whether the mother received at least two injections for tetanus and whether the child received vaccination for BCG, measles and polio. Finally, we cluster standard errors at the level of the 21 sentinel sites from which the samples were draw.

The results of the exercise are shown in Figure 6, where the coefficient on the SC-ST and OBC dummy are plotted in left hand and right hand panel, respectively.[10] For every age, there are two

[9]The wealth index is constructed from three indices: housing quality, access to services, and ownership of consumer durables. It assumes that its three indicators are of equal importance, the wealth index is computed as a simple average of the three indices. The average produces a value between 0 and 1, where a higher wealth index indicates a higher socioeconomic status. The indicators included in housing quality are material of wall, roof, floor and household density. The access to services includes whether the child's household has access to electricity, safe drinking water source, sanitation facility and the type of fuel for cooking. The list of consumer durables considered are Radio, Television, Bicycle, Motorbike, Automobile, Landline phone, Mobile phone, Refrigerator, Fan.

[10]The corresponding regression table is shown in Table 8 in the appendix.

plots: one, the coefficients for the regression with only the basic controls, and two, the coefficient from the regression with the extended set of controls.

Examining the coefficient on the SC-ST dummy from the regression with the extended set of controls shows that the SC-ST are 16, 9.7, 10, 8.1 and 11 percentage points more likely to be stunted than the upper castes at the ages of 1, 5, 8, 12 and 15 y, respectively. The corresponding figures for OBCs are 7.7, 6.9, 8.8, 7.8 and 7, respectively. The results from this exercise suggest that (i) caste gaps are large and sizable at all age profiles; for instance, at 1 and 15 y, SC-ST are more than 94 and 73 percent respectively, more likely to be stunted than the UCs. OBCs at 1 and 15 y are 45 and 46 percent respectively, more likely to be stunted than the upper castes; (ii) the gaps in stunting are persistent well into the teenage years; and (iii) these gaps cannot be explained by education, wealth, levels of urbanization or other socioeconomic characteristics included in $XBasic_{ikr}$ and $XExtd_{ikr}$.

### 3.3 The importance of stunting at age 1 on future stunting

The evidence from Figure 6 shows that caste gaps in stunting remain persistent. However, it does not allow us to distinguish between three important explanations, namely: (1) do we observe persistence because of the fact that gaps already present at age 1 determine the entire trajectory of gaps? (2) Is it the case that from the group of children who were stunted at age 1, the children from the socially disadvantaged groups are less likely to transit out of stunting? (3) Is it the case that from the group of children who were *not* stunted at age 1, children from the socially disadvantaged groups are more likely to transit into stunting?

In order to gain traction on these questions, Panel A of Figure 7 shows the probability of being stunted at age 15, conditional on being stunted at age 1, for the three caste groups. It shows that 50 and 44 percent of SC-ST and OBC children respectively, are likely to continue to be stunted at age 15, conditional on being stunted at age 1. For the upper castes this figure is only 33 percent. This suggests that children from the socially disadvantaged groups are less likely to transit out of stunting by age 15, conditional on being stunted at age 1.

Panel B of Figure 7 shows the probability of being stunted status by age 15, conditional on not being stunted at age 1, for the three caste groups. For this group, we see that 21 and 20 percent of SC-ST and OBC children, who were not stunted at age 1, are likely to be stunted by age 15, compared to only 12 percent for UCs. This shows that children from socially disadvantaged groups are more likely to transit into stunting compared to upper castes, conditional on not being stunted at age 1.

To more formally evaluate the importance of stunting at age 1, and the differential entry and exit based on the caste identity, we estimate the following OLS regression

$$S_{ikr} = \psi_0 + \gamma \text{Stunted Age } 1_i + \delta_1 \text{Stunted Age } 1_{ik} * SCST_i + \delta_2 \text{Stunted Age } 1_i * OBCs_i + \psi_1 SCST_i + \psi_2 OBCs_i + XBasic_{ikr} + XExtd_{ikr} + \epsilon_{ikr}, \quad (3)$$

where $S_{ikr}$ refers to the outcome of child $i$ resident in cluster $k$ and of age $r$, where $r = 5, 8, 12, 15$. The coefficient $\gamma$ captures the effect of stunting at age 1 on stunting at future ages. The coefficients $\delta_1$ and $\delta_2$ tell us whether the impact of stunting at age 1 on future stunting outcome varies by caste group. The coefficients $\psi_1$ and $\psi_2$ show whether the probability of being stunted at age greater than 1 varies by caste status, conditional on not being stunted at age 1. $XBasic_{ikr}$ refers to the set of basic controls, which include the child's age in month, as well as dummies for gender, rural residence and the sampling cluster to which the child belongs. We also estimate regressions where we have an extended set of controls denoted by $XExtd_{ikr}$, namely, household size, mother's years of education and a wealth index. We again cluster standard errors at the level of the 21 sentinel sites from which the samples were drawn.

The results in Table 3 show that the being stunted at age 1 has a sizable impact on the probability of being stunted at ages 5, 8, 12 and 15. Children who were stunted at age 1 are 43, 38, 30, and 22 percentage points more likely to be stunted at the age of 5, 8, 12 and 15 compared to children who were not stunted at age 1. This confirms why early childhood chronic malnutrition as captured by stunting status at age 1 is likely to affect longer run health outcomes.

Examining the interaction of the SC-ST dummy with the dummy for being stunted at age 1 ($\delta_1$) shows that at ages 8, 12 and 15, these coefficients are positive and economically meaningful, though not statistically significant at conventional levels. For instance, the point estimates suggest that at ages 12 and 15, SC-ST children (who were stunted at age 1) are 11 and 7.2 percentage points more likely to be stunted. The interaction terms for the OBC dummy with the dummy for being stunted at age 1 ($\delta_2$) is always insignificant, though smaller and often also negative. This suggests that SC-ST are less likely to transit out stunting, conditional on being stunted at age 1. However, it is important to reiterate that the coefficients are not significant.

The coefficients on the SC-ST dummy ($\psi_1$) are always positive but never significant, though again the point estimates suggest that children from SC-ST who were not stunted are around 5 percentage points more likely to become stunted at age 15 as compared to the upper castes. Finally, $\psi_2$, the coefficient on the OBC dummy is positive and significant at ages 8 and 12. OBC children, conditional on not being stunted at age 1, are 8 and 6.5 percentage points more likely to be stunted at ages 8 and 12, respectively.

Summarizing the evidence from Table 3 shows that the differences in caste gaps in stunting emerge from: (1) differential levels of stunting at age 1 between the caste groups that are very persistent;

(2) a portion of the caste gaps between the upper castes and SC-ST arises from the fact that being stunted at age 1 has a larger impact on stunting at later ages for the SC-ST, though this evidence is weaker; and (3) a portion of the caste gaps arises from the fact that OBC children are more likely to become stunted at later ages, relative to the upper castes.

# 4 Stunting, cognition and learning outcomes: Evidence from panel estimates

Early childhood stunting is significantly negatively associated with cognitive performance of children. This channel can also explain the lasting effects of stunting on adult life outcomes. In the Indian context, the disproportionate burden of stunting faced by the SC-ST and OBCs has important implications directly beyond health, as it potentially effects future human capital, and hence future living standards.

The Young Lives datasets have been used to estimate the impact of stunting on cognition. For instance, Woldehanna et al. (2017), using data for Ethiopia, show that stunted children scored 16.1% less in the Peabody Picture Vocabulary Test and 48.8% less in the Quantitative Assessment test at the age of eight compared to non-stunted children[11]. Fink and Rockers (2014), using longitudinal data from 3327 children aged 8-15 years collected in Ethiopia, India, Peru, and Vietnam as part of the Young Lives project, find that first, 36% of children stunted at age 8 managed to catch up with their peers by age 15, and those who caught up had smaller deficits in cognitive scores than did children who remained stunted. Second, physical growth faltering was not restricted to early childhood but rather affected a substantial share of children in the 8-15 year age range, with large negative consequences for cognition and schooling outcomes.

We explore the impact of being stunted at age 1 on a range of cognitive and learning outcomes at ages 5, 8, 10 and 12 years.

## 4.1 Stunting and cognition at age 5

At age 5, we examine the following outcomes: (1) standardized score on the Peabody Picture and Vocabulary Test (PPVT) and whether the child falls below the 25th percentile on the PPVT Score; (2) The standardized score on the cognitive development assessment (CDA) test and whether the child falls below the 25th percentile on the CDA test[12]. These measures help us understand not

[11]Controlling for confounding variables such as length of breastfeeding, relative size of the child at birth, health problems of early childhood such as acute respiratory illness and malaria, baseline household wealth, child gender, household size and parental education

[12]The PPVT intends to measure vocabulary acquisition in persons from the age of two and half into adulthood. The main task involves the test taker selecting the picture that best represents the meaning of a stimulus word presented orally by the examiner. The CDA was only conducted on the quantitative sub-scale of the CDA and "Notions such as a few, most, half, many, equal, a pair, etc. are assessed with statements such as: 'Point to the plate that has a few cup-

only the average effect of stunting at age 1 on cognitive outcomes at age 5, but also shed light on how stunting might affect the weight in the lower part of the distribution.

Figure 8 shows the kernel density for the standardized CDA and PPVT score by stunting status at age 1 in Panel A and B, respectively. The distribution shows that the cognitive abilities of children who are stunted at age 1 is left shifted with a higher mass in the bottom part of the distribution. The corresponding regression estimates are shown in Panel A of Table 4. The odd numbered columns, (1), (3), (5) and (7) account for the basic controls, namely, sex of child, age of child in months, rural urban residence, type of school, whether child is enrolled in pre-school and language of the PPVT test. The even numbered columns, (2), (4), (6) and (8), with extended controls, additionally accounting for the wealth index, household size, monthly expenditure on food and non-food items measured in real terms and whether the child has been enrolled in caste-based welfare program at age 5. The results show that being stunted at age 1 increases the probability of being in the bottom quartile of the PPVT or the CDA score by 10 and 6.4 percentage points. Compared to children who are not stunted, those stunted at age 1 are 62 and 32 percent more likely to be in the bottom quartile of the PPVT or CDA scores respectively. Stunting at age 1 also reduces the standardized score on the PPVT and CDA by 0.096 and 0.18 standard deviation units, respectively, though only the latter is significant at conventional levels. This suggests that stunting at age 1 has important implications for cognitive development at age 5.

### 4.2 Stunting and cognition at age 8

At age 8 the four outcomes we consider are: (1) a dummy for whether the child can read; (2) the number of correct words read per minute; (3) the standardized score on the Mathematics test; and (4) whether the child can subtract 4 from 20. Figure 9 shows the four outcomes by stunting status at age 1. Panel A shows that whereas 40 percent of stunted children at age 1 can read, this goes up to 60 percent for children who are not stunted, or by 50 percent. Similarly, for all the other three outcomes, we see large differences contingent on stunting status at age 1. Panel B shows that relative to children stunted at age 1, children who are not stunted at age 1 can read 45 percent more words per minute; Panel D shows that the density of the standardized test scores in Math is shifted significantly to the right for children who are not stunted at age 1; and finally, Panel C shows that children who were not stunted at age 1 are 18 percent more likely to be able to subtract 4 from 20.

The regression estimates on the impact of stunting at age 1 on cognitive outcomes at age 8 is shown in Panel b of Table 4. The odd numbered columns again account only for basic controls, namely, sex of child, age of child in months, rural urban residence, language of instruction in schools, the

cakes'." For futher details on the PPVT and the CDA refer to https://www.younglives.org.uk/content/psychometric-characteristics-cognitive-development-and-achievement-instruments-round-2-young.

distance to school in minutes, whether the child has ever dropped out of school, the wealth index, household size, monthly expenditure on food and non-food items measured in real terms. The even numbered columns labeled extensive controls additionally accounts for the standardized CDA score at age 5. These are akin to value-added models in the literature, and show the impact of deficits in child height over and above their impact through affecting skills and cognition in the past (Chetty et al., 2014a,b).

Column (2), which accounts for the basic controls plus the CDA score at age 5, shows that stunting at age 1 reduces the likelihood of being able to read at age 8 by 12 percentage points, or about 20 percent. It also reduces the number of correct words read per minute by 13 percent (column 4), the standardized math test score by 0.28 standard deviation units (column 6), and finally, the probability that the child correctly can subtract 4 from 20 by 8.4 percent (column 8). These results are robust to controlling for important observable socioeconomic characteristics, as well as the standardized CDA score at age 5. As Panel A of Table 4 shows, stunting at age 1 also affects CDA scores at age 5, the estimates reported in Panel B of of Table 4 can be considered the lower bound estimates of stunting at age 1 on cognition at age 8.

### 4.3 Stunting and cognition at age 12

We consider the following outcomes: the percentage and standardized scores obtained in (1) language; (2) mathematics; (3) English; and (4) PPVT. Figure 10 shows the raw outcomes by stunting status at age 1. Relative to children who were stunted at age 1, the children who were not stunted at age 1 score 18, 37, 17 and 9 percent more in reading, mathematics, English and PPVT, respectively. Thus, the raw gaps remain sizeable at age 12, especially for mathematics.

The regression estimates, where now the test scores have been standarized to have mean 0 and standard deviation 1 are shown in Table 5; the basic controls included in the odd numbered columns are identical to those of Panel B of Table 4, but measured at age 12. The extended controls specifications of columns (2), (4), (6) and (8) additionally includes a dummy for whether the child could read at age 8 and the standardized score from the math test conducted at age 8. Without accounting for learning outcomes at age 8, stunting at age 1 reduces language, math, English and the PPVT score by 0.33, 0.33, 0.29 and 0.22 units of a standard deviation, and all the estimates are significant at least at the 5 percent significance level. Once we consider the value added model, as expected, the effect of stunting on age 1 on learning outcomes at age 12 becomes smaller, but still remains statistically significant and economically meaningful, reducing test scores in reading, math, English and PPVT by 0.20, 0.17, 0.19 and 0.12 units of a standard deviation.

### 4.4 Stunting and cognition at age 15

Figure 11 shows the raw outcomes by stunting status at age 1 for two learning outcomes at age 15, the percentage scores obtained in: (1) language; and (2) mathematics. The figure shows that the raw gaps are relatively small, though the difference in language outcomes are still statistically significant. Moreover, both the groups do poorly in Mathematics, with the average being less than a third of the questions answered correctly.

The regression estimates, where the test scores have been standardized to have mean 0 and standard deviation 1 are shown in Table 5. The basic controls included in the odd numbered columns are identical to those of Panel B of Table 4, but measured at age 15. The extended controls specifications of columns (2), (4), (6) and (8) additionally includes the standardized score from the reading and math test conducted at age 12.

Without accounting for learning outcomes at age 12, stunting at age 1 reduces reading and math scores by 0.32 and 0.15 units of a standard deviation, respectively, though only the former is statistically significant at conventional levels. Once we include the learning outcomes from age 12, the effect of stunting at age 1 reduces further; stunting now reduces reading and math outcomes by 0.19 and 0.032 units of a standard deviation, where only the former is significant. However, it is important to note that both the standardized score in reading and mathematics at age 12 are economically meaningful correlates of reading and math outcomes at age 15, and at ages 5, 8 and 12. A back of the envelope calculation would suggest that the cumulative effect of stunting at age 1 is to reduce language and math scores by 0.28 and 0.25 units of a standard deviation.

## 5 Stigmatizing (illegal) social norms and child health

The evidence so far has highlighted the persistence of caste gaps. To what extent might this be a reflection of societal discrimination and stigma towards lower ranked groups? This section contains a preliminary investigation of the role of discriminatory practices in stunting outcomes.

Despite decades of legal or formal equality between caste groups, discriminatory practices against Dalits such as residential segregation, violent hate and sexual crimes especially against Dalit women, denial of entry into temples, prohibitions on inter-caste marriages, forms of bonded labor, segregation in classrooms and discrimination by teachers, discriminatory access to water and irrigation facilities, unequal treatment under the justice system and discrimination in public streets and market places among others remain widespread and rampant (Barbour et al., 2007).

More directly related to field of health, the study by Shah et al. (2006, 104) on *Untouchability in Rural India* reports that Dalits were denied entry into private health centers or clinics in 74 out of 348 villages surveyed. Moreover, the study found that in 30-40 percent of the villages surveyed, public health workers refused to visit Dalit villages. In 15-20 percent of villages, Dalits were denied admission to public health clinics; if admitted, in 10-15 percent of the villages they received discriminatory treatment. Similarly, Acharya (2010, Table 7.3, 218), based on interviews with around 200 Dalit children, reports widespread discrimination in rural public health care services in the states of Gujarat and Rajasthan. These patterns of discrimination are reported in all spheres ranging from home visits of health care professionals, practice of untouchability, information provision, dispensation of medicine, diagnosis and laboratory testing (Acharya, 2010, 215). For instance, 91 percent of Dalit children report experiencing caste-based discrimination in receiving medicines, and 87 percent in the conduct of pathological tests. These practices are especially prevalent among grassroots workers with 94 percent of respondents reporting that Auxiliary Nurse Midwife (ANM) refuse to enter Dalit homes; 93 percent reporting that public health workers refuse to touch children while dispensing medicine, and 98 percent report they serve food last to SC children. These discriminatory practices are driven by prejudiced notions harbored by service providers about Dalit children. As Acharya (2010, 225) writes, "Conventionally, improper drainage, flies and garbage, and consumption of stale food mark their understanding of the Dalits. However, during the group discussions, 'children with running nose, which they keep licking', ill-clad or naked children playing in the dirty streets also emerged as the markers" (see also Acharya 2010, Table 7.8, 225).

A recent study by Coffey et al. (2019, 1429) finds that "Although the average gap in height between higher caste and lower caste (SC and OBC) children in rural India cannot fully be explained by household-level SES variables, these variables can fully explain the caste height gap in those localities where SC and OBC children do not live with higher caste neighbors." They interpret the remaining gaps, after accounting for the socioeconomic determinants, as presence of 'local' discrimination. They posit the hypothesis that in areas where the lower caste people inhabit with the higher castes, the higher caste "might enforce the social rank of lower caste households, especially SCs, in ways that could create stress and limit access to common resources, such as clean water, which would matter for child health but would not show up in household economic status" (Coffey et al., 2019, 1432).

These studies suggest that stigmatizing and extralegal practices against Dalits might be responsible for the enduring gaps in stunting that remain impervious to a whole range of socioeconomic controls. In this section, we present some additional descriptive evidence in this regard that supports the link between the practice of untouchability and caste gaps in stunting.

To study this link, we leverage a unique question in the second round of the IHDS conducted in 2011-12, which asks, "In your household, do some members practice untouchability", which is further followed up by "Would there be a problem if someone who is scheduled caste were to enter your kitchen or share utensils." We code households who answer in the affirmative to either of these questions as households who practice untouchability. We aggregate all households at the level of the state, and then create a state level indicator that captures the proportion of households that practice untouchability. The data shows that on average 24 percent of the households in a state report practicing untouchability, with the state at the $25^{th}$ (Nagaland) and $75^{th}$ percentile (Odisha) having 7 and 43 percent, respectively.

Next, using the latest round of the NFHS, we calculate the state level gaps in stunting between the UC and SCs, the latter group subject to the stigmatizing and dehumanizing practice of untouchability. We then merge the data on the prevalence of untouchability with the extent of caste gaps in stunting. Figure 12 shows the scatter plot and the fitted line between the caste gaps in stunting and the prevalence of untouchability. It shows that caste gaps between the SC and UC are increasing in the proportion of households who report practicing discrimination.

Table 6 looks at the association between the prevalence of untouchability at the state level and stunting probability of SC children. Column (1), which accounts for the baseline controls (age in months, sex and rural or urban residence) shows that moving from a state like Kerala to Uttar Pradesh increases the probability of a SC child being stunted by 13 percentage points. Column (2) includes mother years of education; Column (3) includes health inputs at birth, mother's health status and birth history (these are controls labeled as W3, W4 and W5 in Section 2.1); Column (4) includes the factor score on the wealth index; Column (5) includes controls for water and toilet source (these are controls labeled as W6 in Section 2.1); and finally, Column (6) includes all the above specified controls. Columns (1)-(6) show that the prevalence of the practice of untouchability is a robust and statistically significant correlate of the probability of stunting for SC children. The coefficient in column (6) suggests that even after accounting for socioeconomic controls, moving from a state like Kerala to Uttar Pradesh increases the probability of a SC child being stunted by 6.5 percentage points. These results however are only meant to be suggestive and point to the potential importance of discriminatory practices in explaining differences in early childhood outcomes between caste groups.[13] Future work is needed to explore whether there is indeed a causal relationship and the pathways underlying the relationship between caste discrimination and childhood stunting.

## 6 Conclusion

In the literature on adult life discrimination, reference is made to "pre-market" discrimination that differently shapes the acquisition of income-earning characteristics, such as education and health, across social groups, such as caste. These gaps are an important aspect of contemporary caste realities, but they need to be further unpacked. To the best of our knowledge, ours is the

[13]Reassuringly, Column (7) of Table 6 shows that this documented correlation is absent when exploring the probability of stunting for upper caste children and the prevalence of untouchability at the state level.

first comprehensive examination that analyses the trends, dynamics and trajectories of gaps in chronic malnutrition in early childhood up until the age of 15. We show that the gaps in chronic malnourishment between caste groups that result in stunting are large and remain persistent. Relative to upper caste children, the incidence of stunting is 40 and 23 percent higher among the SC-ST and OBCs, respectively. Gaps in stunting between caste groups at age one are very striking; moreover, children from socially disadvantaged groups are both more likely to become stunted, and less likely to transit out of stunting, at later ages. We also demonstrate that childhood stunting has substantively important effect on learning and cognitive outcomes in later years. Specifically, we show the effects of stunting at age 1 on learning outcomes at ages 5, 8, 12 and 15. For instance, stunting at age 1 year reduces the probability of being able to read at age 8 by 20 percent, and reading scores at age 12 by 0.20 units of a standard deviation. Similarly, it reduces math scores at ages 8 and 12 by 0.28 and 0.17 units of a standard deviation, respectively. Our evidences provides an important reason as to why adult life caste disparities persist, and in fact, in most instances have widened in India (Deshpande and Ramachandran, 2019).

The caste gaps in stunting only reduce by half after accounting for observable socio economic characteristics – mother's education, health inputs at birth, mother's health status, mother's birth history, household wealth and sanitation – suggests that caste operates through other mechanisms besides affecting material standard of living to have a bearing on anthropometric outcomes. Dalits have been historically associated with ritually polluting occupations, and hence used to be considered "untouchable". For this reason, they have been deeply stigmatized. Most of the existing research has focused on how stigma can affect the stigmatized even in the absence of discriminatory behavior on the part of others, as dominant cultural stereotypes are widely known, and this knowledge may affect behavior through ideomotor process (Major and O'brien, 2005). However, the role of discriminatory practices in mediating learning and health outcomes of lower caste groups remains an important future task.

# References


Acharya, S. (2010). Caste and patterns of discrimination in rural public health care services. *Blocked by Caste: Economic Discrimination in Modern India (New Delhi: Oxford University Press, 2010)*, 208–229.

Alderman, H., J. Hoddinott, and B. Kinsey (2006). Long term consequences of early childhood malnutrition. *Oxford Economic Papers 58*(3), 450–474.

Barbour, S., T. Palikovic, J. Shah, and S. Narula (2007). Hidden apartheid: Caste discrimination against India's "untouchables". Shadow report to the UN committee on the elimination of racial discrimination. *New York: The Center for Human Rights and Global Justice and the New York. University School of Law*.

Case, A. and C. Paxson (2010). Causes and consequences of early-life health. *Demography 47*(1), S65–S85.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9), 2593–2632.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014b). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review 104*(9), 2633–79.

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

Coffey, D. and D. Spears (2017). *Where India goes: Abandoned toilets, stunted Development and the costs of caste*. Harper Collins.

Cutler, D., W. Fung, M. Kremer, M. Singhal, and T. Vogl (2010). Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in India. *American Economic Journal: Applied Economics 2*(2), 72–94.

Deaton, A. (2007). Height, health, and development. *Proceedings of the National Academy of Sciences 104*(33), 13232–13237.

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

Fink, G. and P. C. Rockers (2014). Childhood growth, schooling, and cognitive development: further evidence from the young lives study. *The American journal of clinical nutrition 100*(1), 182–188.

George, S. (2015). Caste and care: Is Indian healthcare delivery system favourable for Dalits? Institute for social and economic change. WP 350.

Gould, E. D., V. Lavy, and M. D. Paserman (2011). Sixty years after the magic carpet ride: The long-run effect of the early childhood environment on social and economic outcomes. *The Review of Economic Studies 78*(3), 938–973.

Hammer, J. and D. Spears (2016). Village sanitation and child health: Effects and external validity in a randomized field experiment in rural India. *Journal of Health Economics 48*, 135–148.

Headey, D., J. Hoddinott, D. Ali, R. Tesfaye, and M. Dereje (2015). The other Asian enigma: explaining the rapid reduction of undernutrition in Bangladesh. *World Development 66*, 749–761.

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

Human Rights Watch (2014). "They say we're dirty": Denying an education to India's marginalized'. Available at: http://www.refworld.org/docid/5368a3db4.html. Technical report.

Jayachandran, S. and R. Pande (2013). Choice not genes: Probable cause for the India-Africa child height gap. *Economic and Political Weekly 48*(34), 77–79.

Jayachandran, S. and R. Pande (2017). Why are indian children so short? The role of birth order and son preference. *American Economic Review 107*(9), 2600–2629.

Jensen, R. (2012). Another mouth to feed? The effects of (in) fertility on malnutrition. *CESifo Economic Studies 58*(2), 322–347.

Leroy, J. L. (2011). zscore06: Stata command for the calculation of anthropometric z-scores using the 2006 WHO child growth standards. *Statistical Software Components S 457279*.

Major, B. and L. T. O'brien (2005). The social psychology of stigma. *Annual Review of Psychology 56*, 393–421.

Maluccio, J. A., J. Hoddinott, J. R. Behrman, R. Martorell, A. R. Quisumbing, and A. D. Stein (2009). The impact of improving nutrition during early childhood on education among Guatemalan adults. *The Economic Journal 119*(537), 734–763.

McGovern, M. E., A. Krishna, V. M. Aguayo, and S. Subramanian (2017). A review of the evidence linking child stunting to economic outcomes. *International Journal of Epidemiology 46*(4), 1171–1191.

Panagariya, A. (2013). Does India really suffer from worse child malnutrition than sub-Saharan Africa? *Economic and Political Weekly*, 98–111.

Ramalingaswami, V., U. Jonsson, and J. Rohde (1996). The Asian enigma.

Shah, G., H. Mander, S. Thorat, A. Baviskar, and S. Deshpande (2006). *Untouchability in rural India*. Sage.

Spears, D. (2018). Exposure to open defecation can account for the Indian enigma of child height. *Journal of Development Economics*, 1–17.

Subramanyam, M. A., I. Kawachi, L. F. Berkman, and S. Subramanian (2010). Socioeconomic inequalities in childhood undernutrition in India: analyzing trends between 1992 and 2005. *PLoS One 5*(6), e11392.

Tarozzi, A. (2008). Growth reference charts and the nutritional status of Indian children. *Economics & Human Biology 6*(3), 455–468.

UNICEF (2018). World bank group, levels and trends in child malnutrition, 2016.

Victora, C. G., L. Adair, C. Fall, P. C. Hallal, R. Martorell, L. Richter, H. S. Sachdev, Maternal, C. U. S. Group, et al. (2008). Maternal and child undernutrition: Consequences for adult health and human capital. *The Lancet 371*(9609), 340–357.

Woldehanna, T., J. R. Behrman, and M. W. Araya (2017). The effect of early childhood stunting on children's cognitive achievements: Evidence from young lives Ethiopia. *Ethiopian Journal of Health Development 31*(2), 75–84.

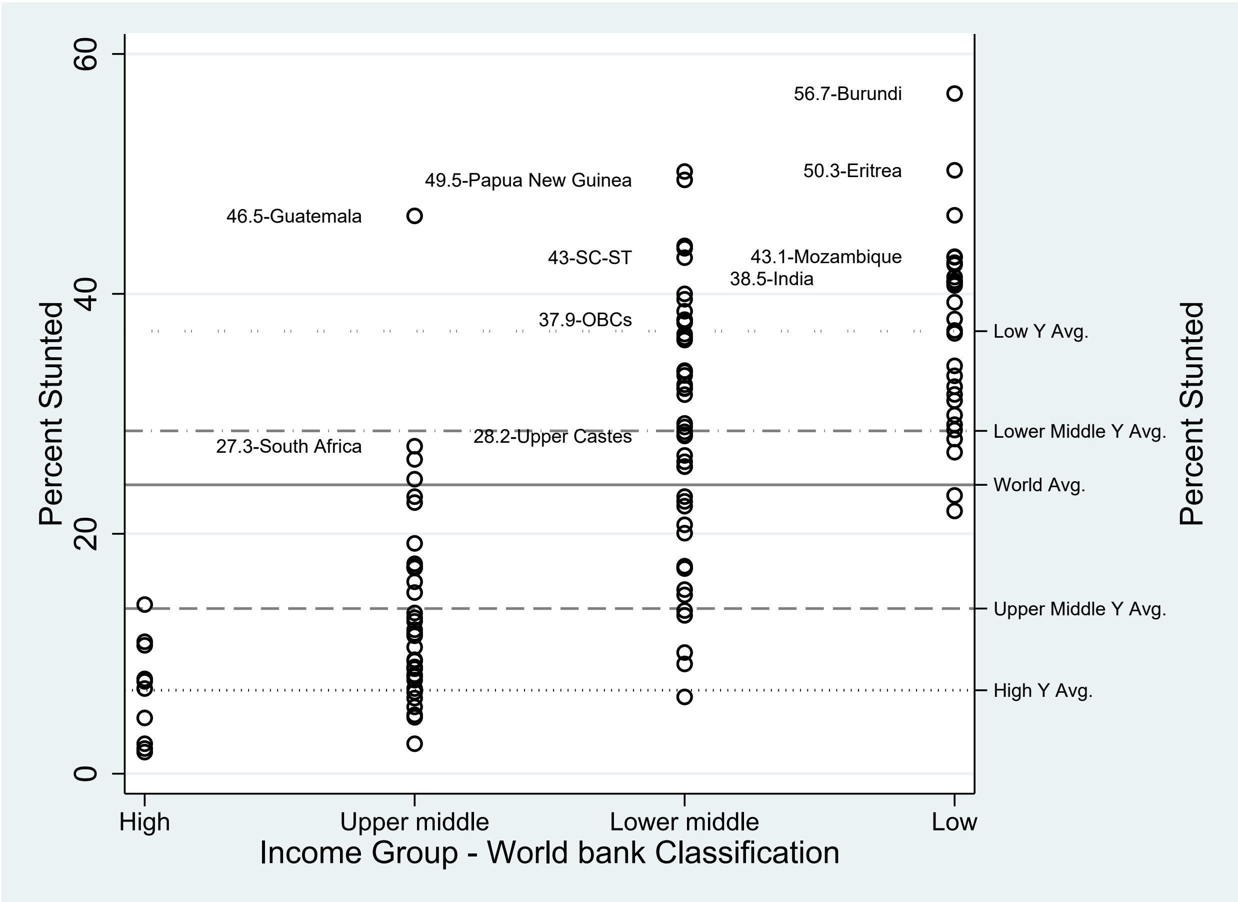


Figure 1: Prevalence of stunting by country and income group

The data on stunting comes is an average over the years 2010-19 for 115 countries and is drawn from the UNICEF, WHO, World Bank: Joint child malnutrition estimates. The data on the stunting levels for the SC-ST, OBCs and upper caste Hindus is from the NFHS-4 2015-16. Each dot represents a country. The average stunting levels for the world and the four income groups are depicted on the right hand side y-axis. For the high and upper middle income group, only countries which have a higher level of stunting that the upper caste Hindus in India is depicted.

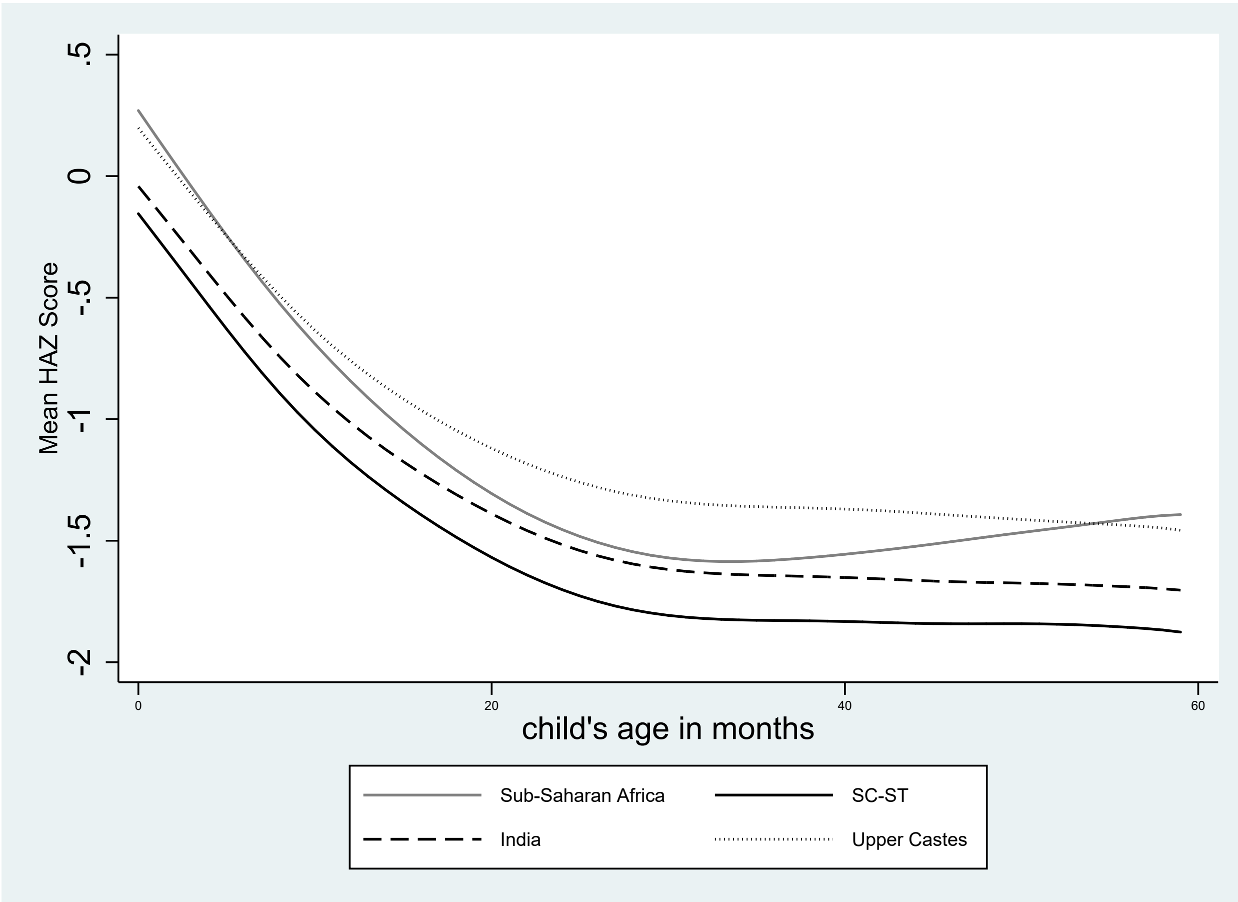


Figure 2: Height-for-age-Z-score comparing Sub-Saharan Africa with India and the upper castes and Schedule Castes and Tribes

The data on the HAZ scores is from the latest available round of the Demographic and Health, conditional on being later than 2010. The mean for sub-Saharan Africa is based on the average of 30 countries for whom data is collected post 2010, namely, Angola (2015-16), Benin (2017-18), Burkina Faso (2010), Burundi (2016-17), Cameroon (2011), Chad (2014-15), Comoros (2012), Republic of Congo (2011-12), Democratic Republic of Congo (2013-14), Ethiopia (2016), Gabon (2012), Ghana (2014), Guinea (2018), Ivory Coast (2010-11), Kenya (2014), Lesotho (2014), Liberia (2013), Malawi (2015-16), Mali (2018), Namibia (2013), Niger (2017), Nigeria (2018), Rwanda (2014-15), Senegal (2017), Sierra Leone (2013), South Africa (2016), Tanzania (2015-16), Togo (2013-14), Uganda (2016), Zambia (2013-14) and Zimbabwe (2015). The data on the stunting levels for India, the SC-ST and upper caste Hindus is from the NFHS-4, 2015-16.

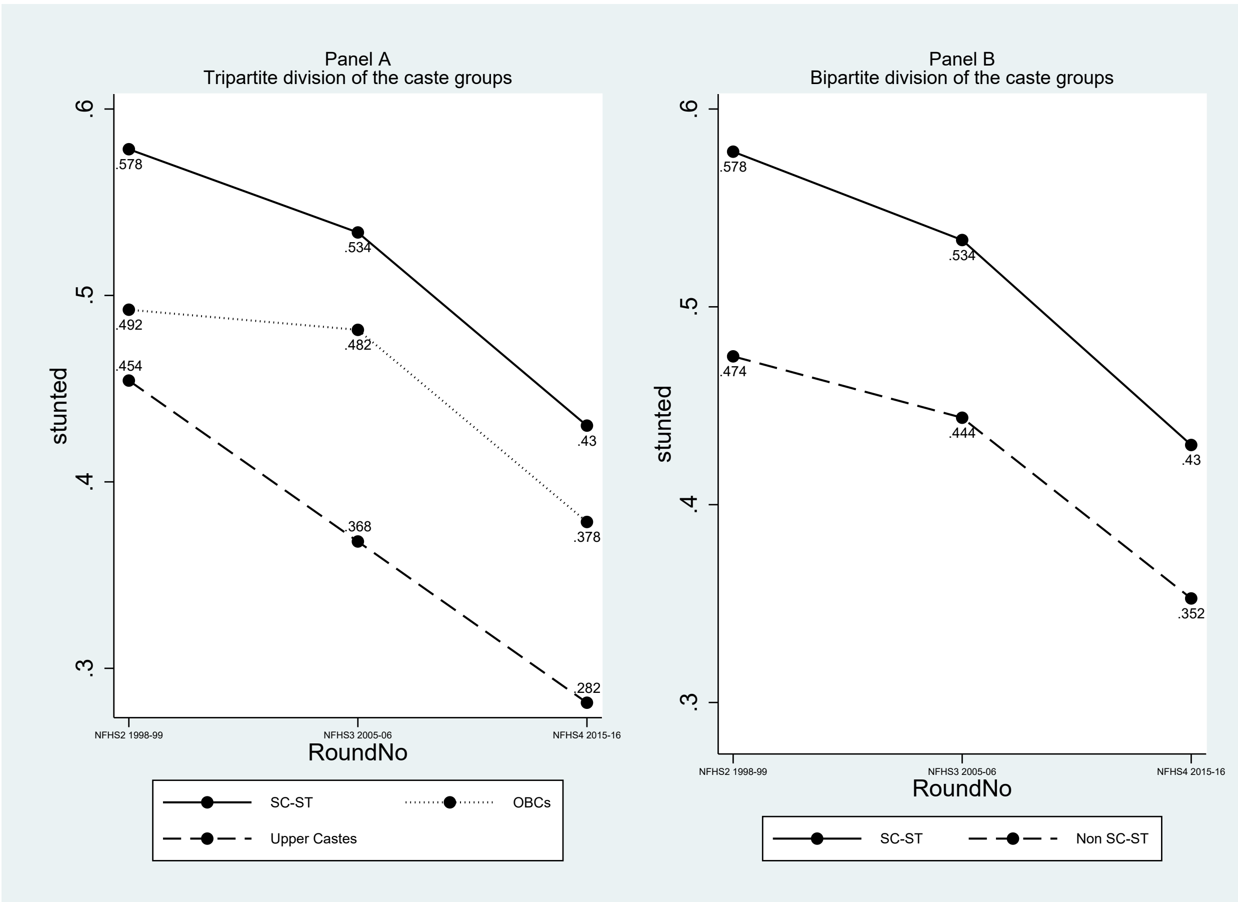


Figure 3: Evolution of stunting by caste groups in India

The data on stunting comes from the three rounds of the NFHS - 1998-99, 2005-06 and 2015-16. The sample is restricted to all individuals whose information on caste is available and who report as belonging to the Hindu denomination. The graph is based on a total of 213,665 observations over the three rounds. The left side panel considers the tripartite division of the caste system into the upper caste Hindus (19.04%), the OBCs (44.88%) and the SC-ST (36.08%), whereas, the right side panel considers only the bipartite division of the caste system into the SC-ST (36.08%) and the non SC-ST (63.92%).

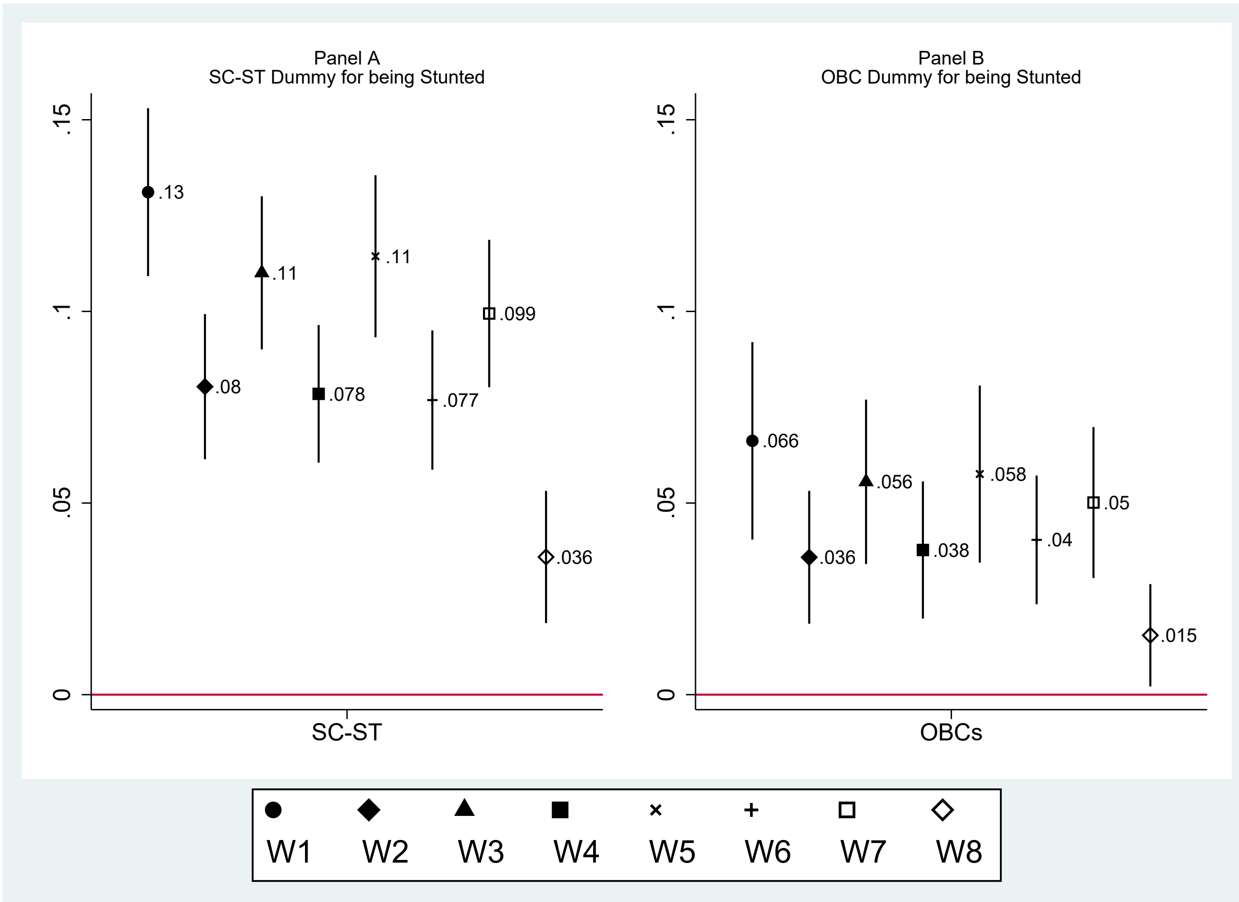


Figure 4: Caste gaps in stunting and the moderating role of socioeconomic factors

The estimates W1-W8 are based on a total of 157,467 observations. Estimates W1 refers to baseline estimates, which include controls for sex of the child, age in months, rural residence, and a set of state dummies and dummies for the NFHS round and result from estimating Equation 1. The estimates W2 adds the years of education of the mother as an additional control. Estimates W3 accounts for the health inputs at birth - (i) place of delivery; (ii) prenatal access to doctor; (iii) breast fed immediately after birth. Estimates labeled W4 accounts for the mother's health status - (i) weight; (ii) height; (iii) height-for-age percentile; (iv) body mass index; (v) Rohrer index; (vi) age; (viii) age at first marriage. Estimates labeled W5 accounts for the mother's birth history - (i) number of of children who have died; (ii) total children born in last 5 years; (iii) age at first birth; (iv) birth order of child. Estimates labeled W6 accounts for the factor score of the wealth index of the mother's household. Estimates labeled W7 accounts for sanitation proxied by - (i) source of drinking water; (ii) type of toilet facility. Estimates labeled W8 account for all the above specified controls (W2-W7) in a single regression. The corresponding regression results are shown in Table 7.

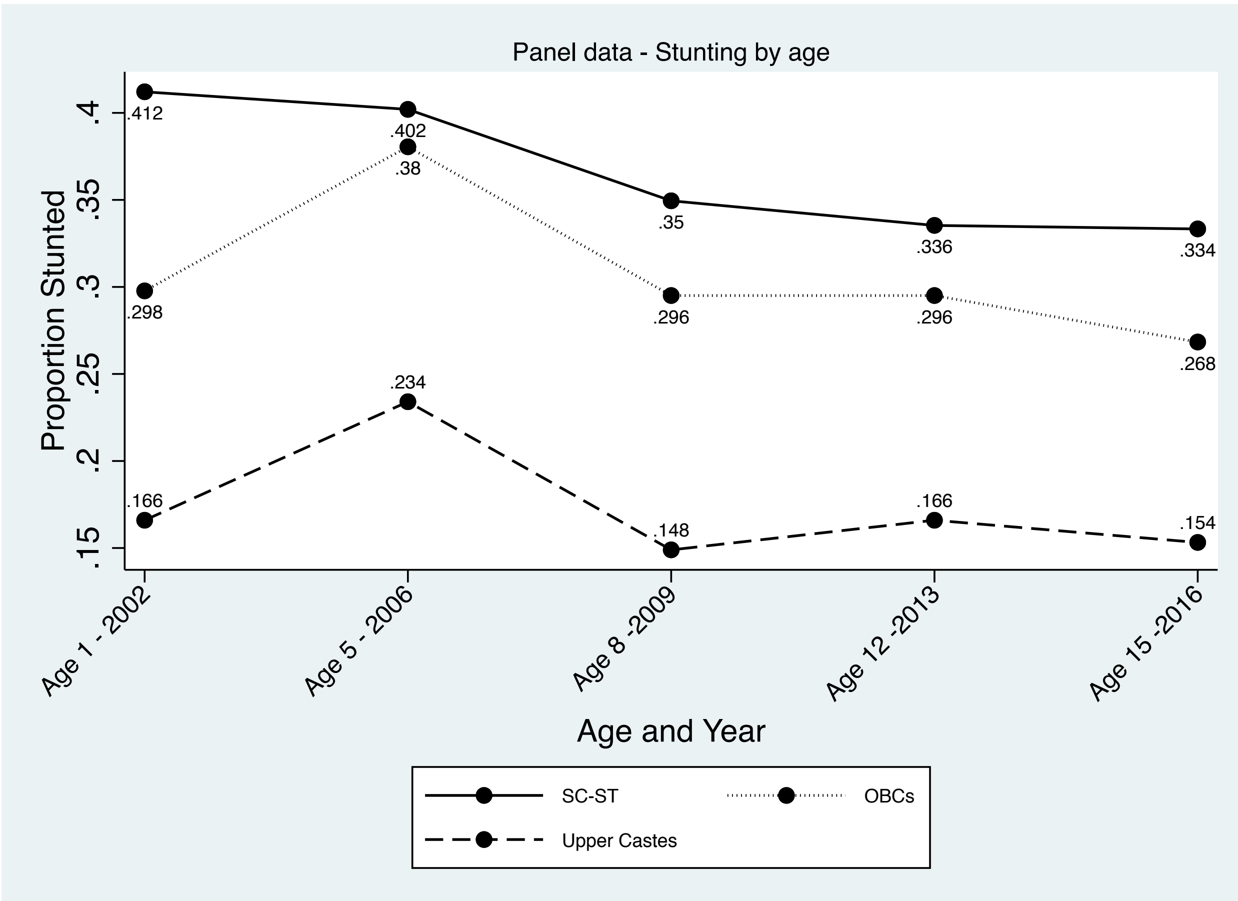


Figure 5: Evolution of stunting by age and caste groups in Andhra Pradesh - Panel Data

The data comes from a panel of children in Andhra Pradesh who were followed as part of the Young Lives Project from the age of 1 in 2002 to the age of 15 in 2016. The anthropometric measurements were made at each point of time: 2002, 2006, 2009, 2013 and 2016. The data is restricted to a sample of 1479 children for whom the anthropometric measures are available at all five points of time.

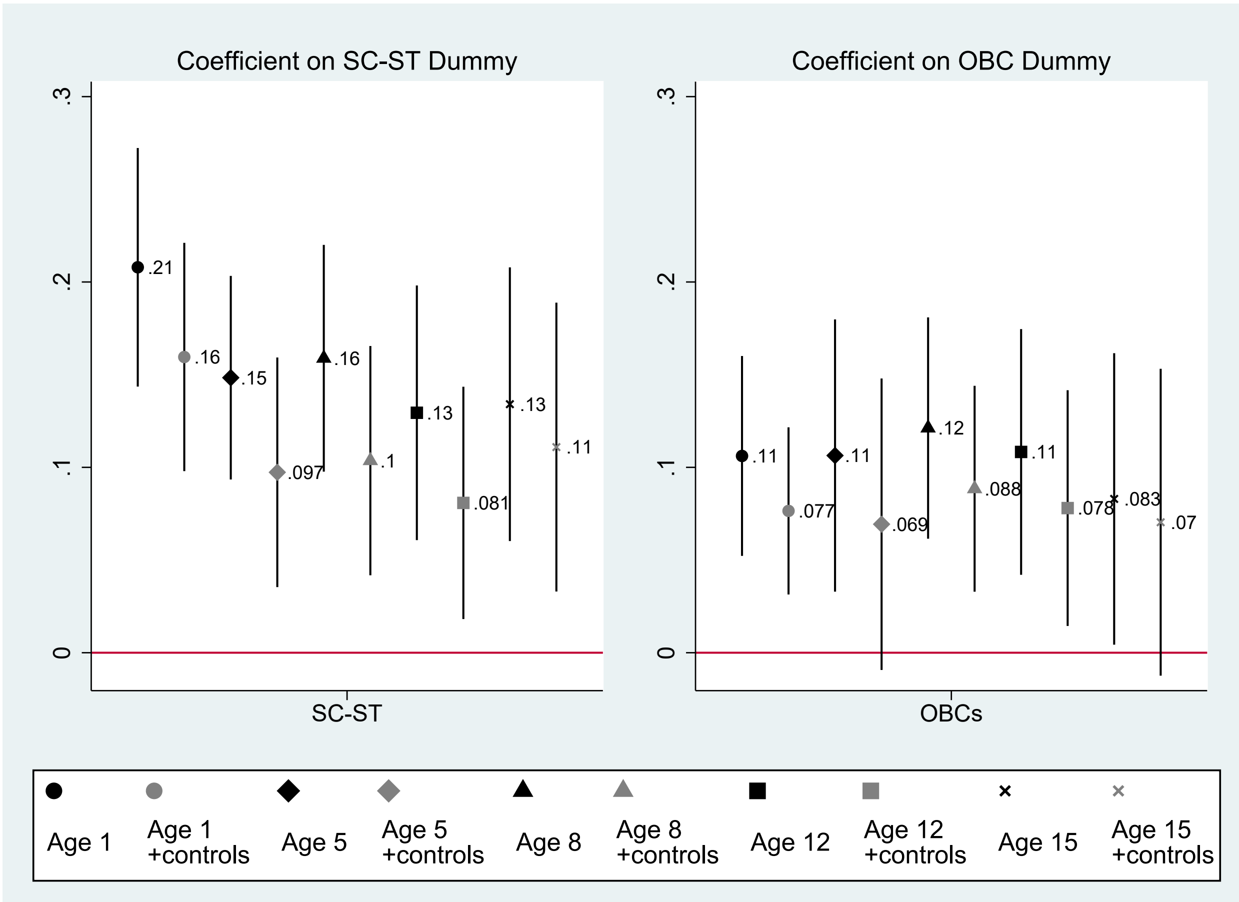


Figure 6: The durability of caste gaps in stunting outcomes - Panel Data

The data comes from a panel of children in Andhra Pradesh who were followed as part of the Young Lives Project from the age of 1 in 2002 to the age of 15 in 2016. The anthropometric measurements were made at each point of time: 2002, 2006, 2009, 2013 and 2016. The data is restricted to a sample of 1479 children for whom the anthropometric measures are available at all five points of time. It shows the results of an OLS for each of the five rounds where the stunting dummy has been regressed on a set of caste dummies, child's age in months, sex, rural dummy and dummies for the 21 sentinel sites from which the data was collected. The regressions with controls besides the forementioned additionally include measured at each point in time a wealth index, household size, mothers education in years, whether has access to safe drinking water and access to sanitation. The proportion stunted from the upper castes for the 5 rounds are 0.17, 0.23, 0.15, 0.17 and 0.15, respectively. The standard errors are clustered at the level of the 21 sentinel sites. The corresponding regression table is shown in Table 8.

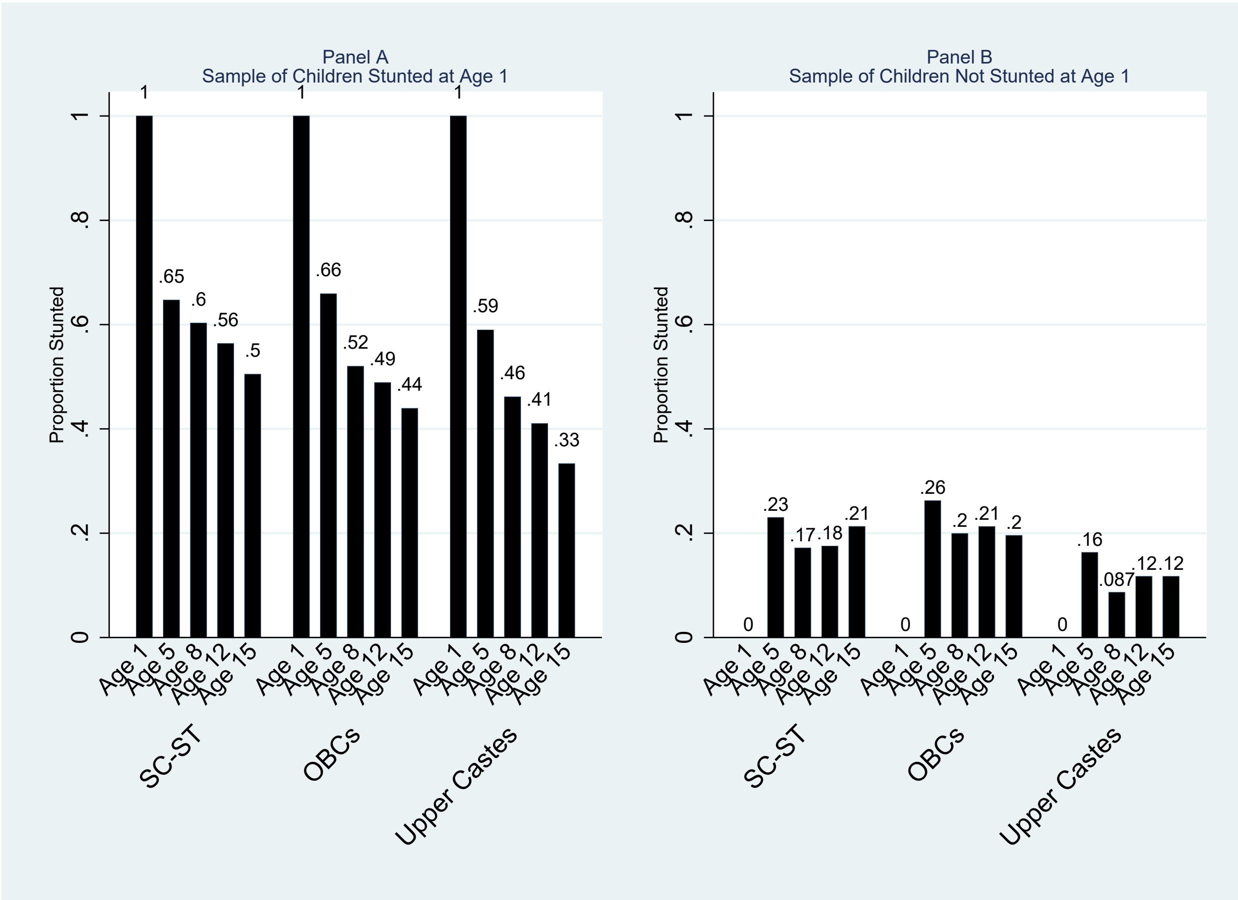


Figure 7: The two-facets of disadvantage for the Schedule Castes and Tribes in stunting outcomes - Panel Data

The data comes from a panel of children in Andhra Pradesh who were followed as part of the Young Lives Project from the age of 1 in 2002 to the age of 15 in 2016. The anthropometric measurements were made at each point of time: 2002, 2006, 2009, 2013 and 2016. The data is restricted to a sample of 1479 children for whom the anthropometric measures are available at all five points of time. Panel A compares the evolution of stunting in children across caste groups who were recorded as stunted at age 1, when the data was collected in 2002. It thus shows the probability of transiting out of stunting status by age 15 conditional on being stunted at age 1 for the three caste groups. Panel B, in turn, compares the evolution of stunting in children across caste groups who were recorded as not being stunted at age 1, when the data was collected in 2002. It thus shows the probability of transiting into the status of stunted at age 15 conditional on not being stunted at age 1 for the three caste groups.

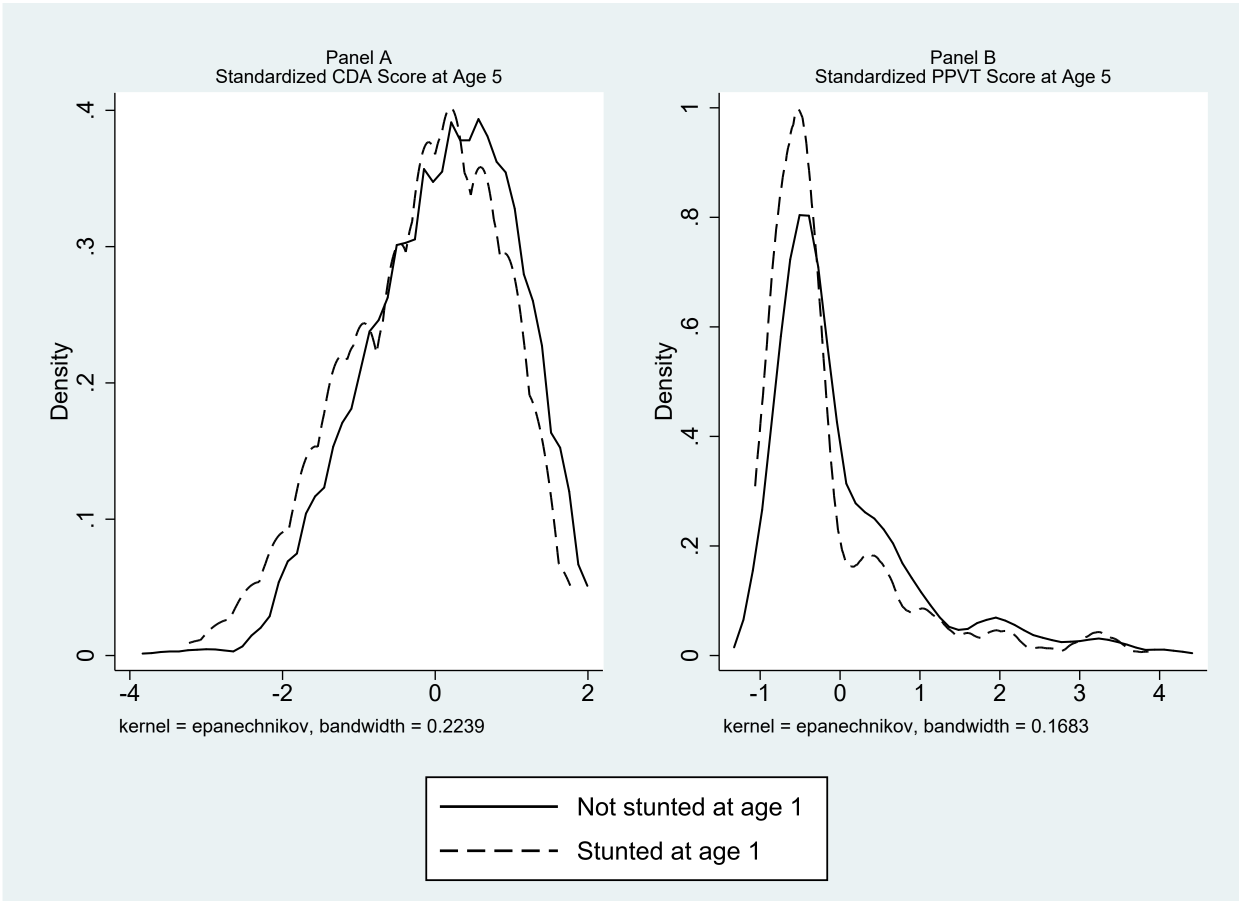


Figure 8: Cognitive outcomes at age 5 by stunting status at age 1

The data comes from a panel of children in Andhra Pradesh who were followed as part of the Young Lives Project from the age of 1 in 2002 to the age of 15 in 2016. The anthropometric measurements were made at each point of time: 2002, 2006, 2009, 2013 and 2016. The data is restricted to a sample of 1479 children for whom the anthropometric measures are available at all five points of time. Panel A and B shows the kernel densities by stunting status at age 1 score on the cognitive development assessment (CDA) test and the Peabody Picture and Vocabulary Test (PPVT), respectively. The PPVT intends to measure vocabulary acquisition in persons from the age of two and half into adulthood. The main task involves the test taker selecting the picture that best represents the meaning of a stimulus word presented orally by the examiner. The CDA was only conducted on the quantitative sub-scale of the CDA and “Notions such as a few, most, half, many, equal, a pair, etc. are assessed with statements such as: ‘Point to the plate that has a few cupcakes’.” For futher details on the PPVT and the CDA refer to https://www.younglives.org.uk/content/psychometric-characteristics-cognitive-development-and-achievement-instruments-round-2-young.

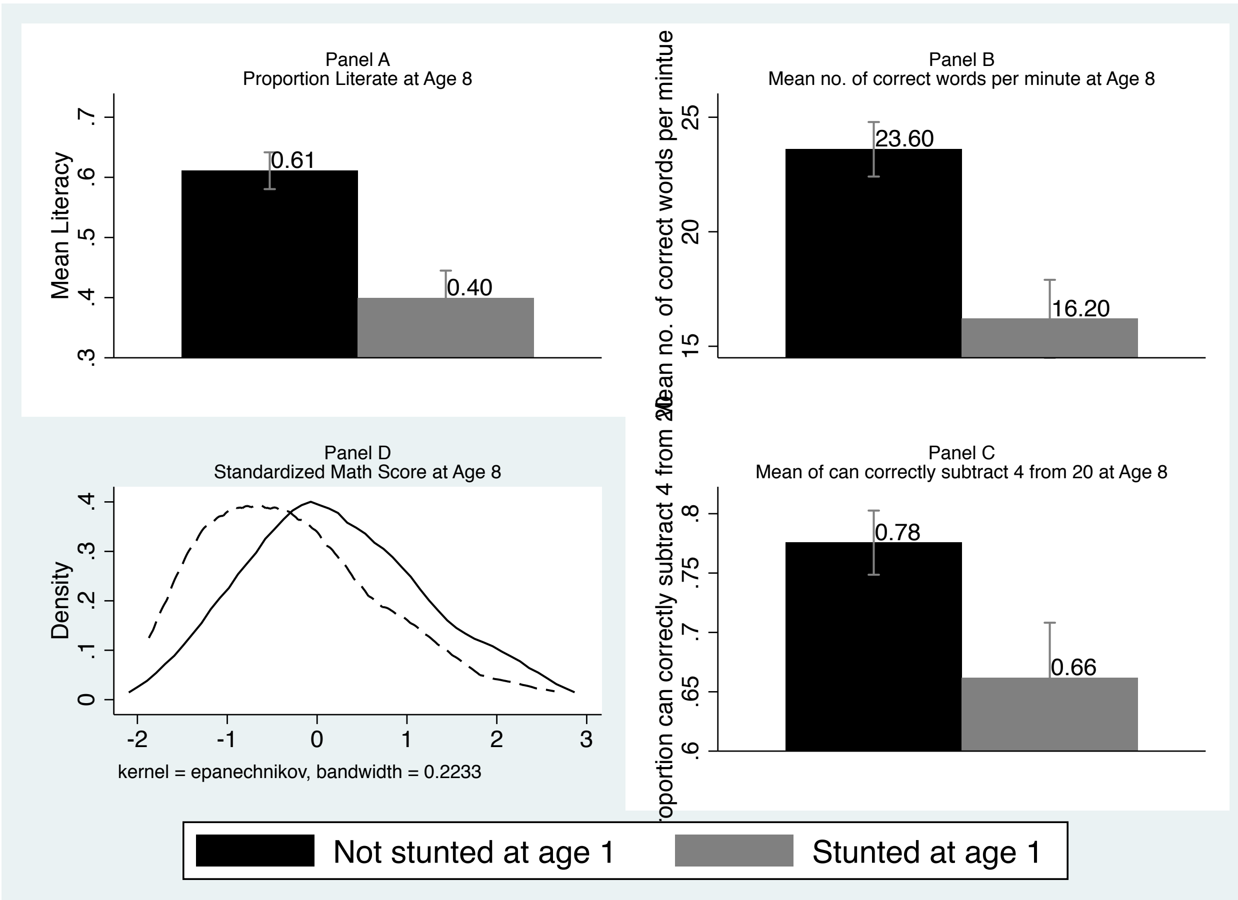


Figure 9: Cognitive outcomes at age 8 by stunting status at age 1

The data comes from a panel of children in Andhra Pradesh who were followed as part of the Young Lives Project from the age of 1 in 2002 to the age of 15 in 2016. The anthropometric measurements were made at each point of time: 2002, 2006, 2009, 2013 and 2016. The data is restricted to a sample of 1479 children for whom the anthropometric measures are available at all five points of time. Panel A shows the proportion literate at age 8 by stunting status at age 1. Panel B shows the average number of correct words read per minute at age 8 by stunting status at age 1. Panel B shows the kernel density of the standardized score on the mathematics test at age 8 by stunting status at age 1. Finally, Panel D shows the proportion of children who can subtract 4 from 20 at age 8 by stunting status at age 1.

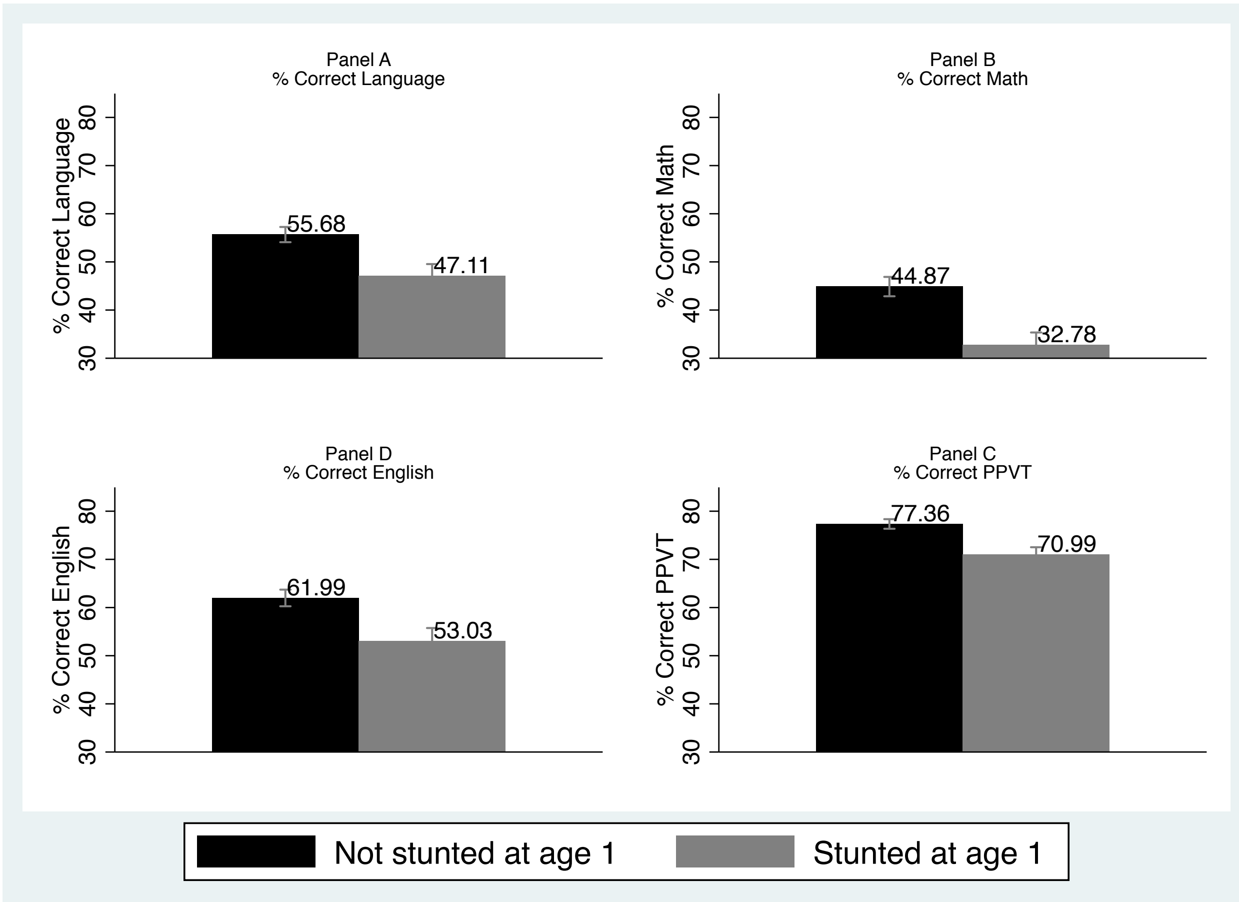


Figure 10: Cognitive outcomes at age 12 by stunting status at age 1

The data comes from a panel of children in Andhra Pradesh who were followed as part of the Young Lives Project from the age of 1 in 2002 to the age of 15 in 2016. The anthropometric measurements were made at each point of time: 2002, 2006, 2009, 2013 and 2016. The data is restricted to a sample of 1479 children for whom the anthropometric measures are available at all five points of time. Panel A shows the percentage score in languages at age 12 by stunting status at age 1. Panel B shows the percentage score in mathematics at age 12 by stunting status at age 1. Panel C shows the percentage score in English at age 12 by stunting status at age 1. Panel D shows the percentage score in PPVT at age 12 by stunting status at age 1.

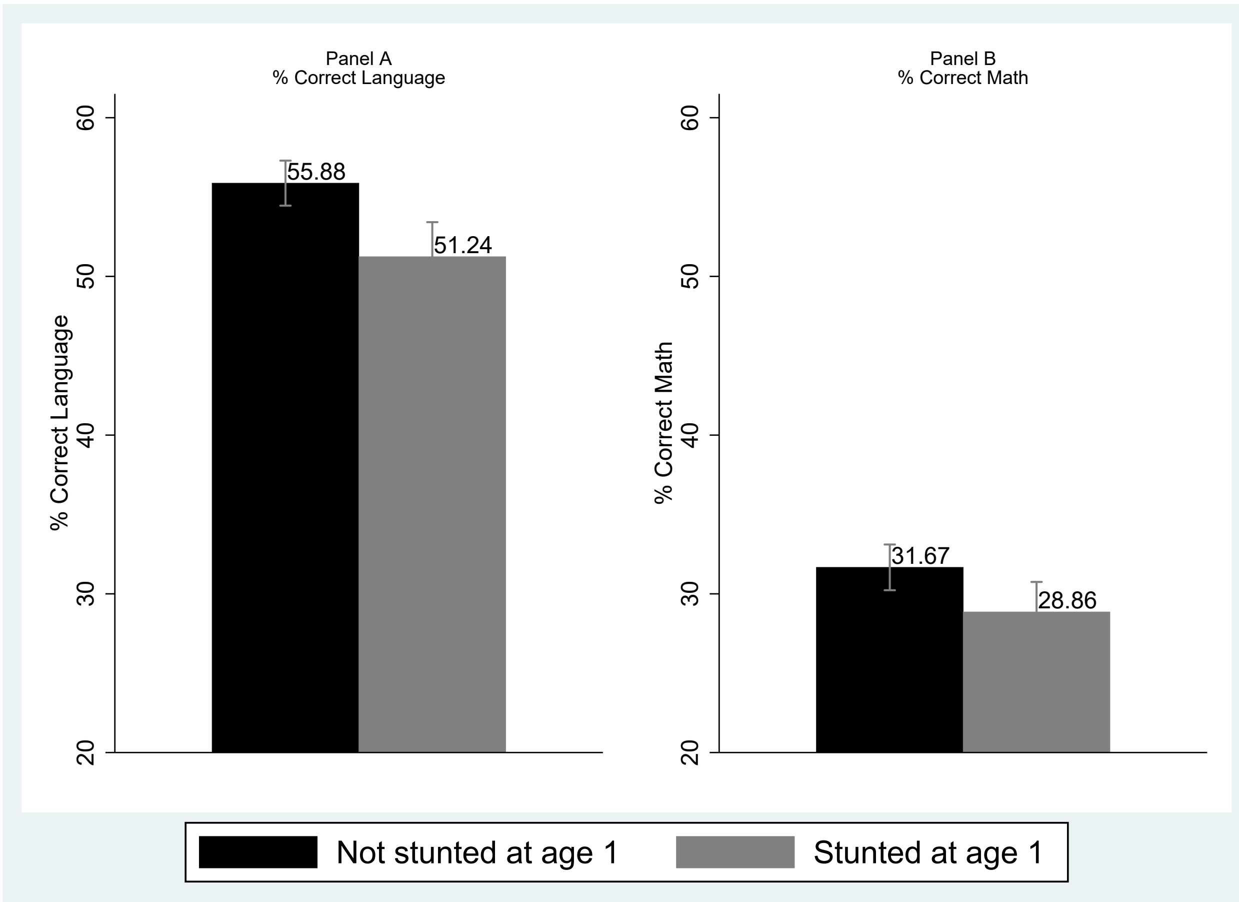


Figure 11: Cognitive outcomes at age 15 by stunting status at age 1

The data comes from a panel of children in Andhra Pradesh who were followed as part of the Young Lives Project from the age of 1 in 2002 to the age of 15 in 2016. The anthropometric measurements were made at each point of time: 2002, 2006, 2009, 2013 and 2016. The data is restricted to to a sample of 1479 children for whom the anthropometric measures are available at all five points of time. Panel A shows the percentage score in languages at age 12 by stunting status at age 1. Panel A shows the percentage score in languages at age 15 by stunting status at age 1. Panel B shows the percentage score in mathematics at age 15 by stunting status at age 1.

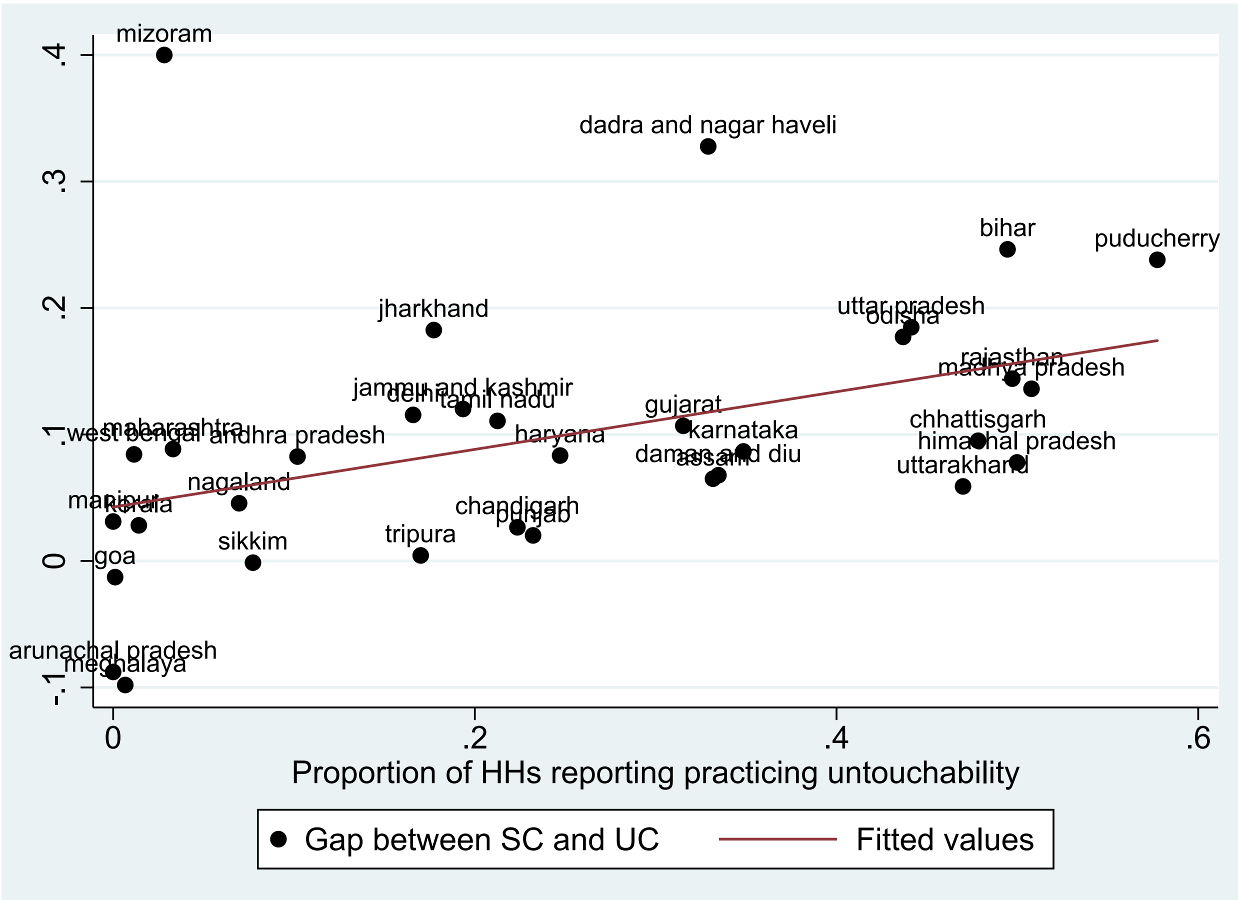


Figure 12: Caste gaps in stunting and the prevalence of untouchability

The prevalence of untouchability at the state level is calculated from the second round of the Indian Human Development Survey (IHDS) conducted in 2011-12. The data on stunting gaps between the schedule castes and the upper castes is based on the NFHS-IV.

Table 1: Stunting gaps and their evolution among caste groups in India

| | height-for-age Z-score | | Dummy for Stunting | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| SC-ST | -0.50*** | -0.45*** | 0.13*** | 0.12*** |
| | (0.048) | (0.031) | (0.012) | (0.0100) |
| | [-0.59 - -0.40] | [-0.51 - -0.39] | [0.11 - 0.16] | [0.098 - 0.14] |
| OBCs | -0.28*** | -0.17*** | 0.072*** | 0.033** |
| | (0.055) | (0.060) | (0.013) | (0.014) |
| | [-0.39 - -0.16] | [-0.29 - -0.045] | [0.046 - 0.098] | [0.0042 - 0.063] |
| SC-ST*NFHS4 | | -0.049 | | 0.017 |
| | | (0.058) | | (0.018) |
| | | [-0.17 - 0.070] | | [-0.021 - 0.054] |
| OBCs*NFHS4 | | -0.12** | | 0.043*** |
| | | (0.051) | | (0.015) |
| | | [-0.22 - -0.014] | | [0.014 - 0.073] |
| SC-ST*NFHS3 | | -0.067 | | 0.026 |
| | | (0.077) | | (0.024) |
| | | [-0.22 - 0.090] | | [-0.022 - 0.074] |
| OBCs*NFHS3 | | -0.13** | | 0.045** |
| | | (0.055) | | (0.017) |
| | | [-0.24 - -0.021] | | [0.011 - 0.079] |
| Age in months | -0.022*** | -0.022*** | 0.0037*** | 0.0037*** |
| | (0.0012) | (0.0012) | (0.00048) | (0.00048) |
| | [-0.024 - -0.019] | [-0.024 - -0.019] | [0.0027 - 0.0046] | [0.0027 - 0.0046] |
| Female dummy | 0.040*** | 0.040*** | -0.0064 | -0.0063 |
| | (0.014) | (0.014) | (0.0047) | (0.0047) |
| | [0.012 - 0.068] | [0.012 - 0.068] | [-0.016 - 0.0032] | [-0.016 - 0.0032] |
| Rural residence dummy | -0.26*** | -0.26*** | 0.071*** | 0.071*** |
| | (0.029) | (0.029) | (0.0086) | (0.0085) |
| | [-0.32 - -0.20] | [-0.32 - -0.20] | [0.054 - 0.089] | [0.054 - 0.088] |
| NFHS3 2005-06 dummy | 0.49*** | 0.56*** | -0.097*** | -0.12*** |
| | (0.041) | (0.058) | (0.013) | (0.019) |
| | [0.41 - 0.57] | [0.44 - 0.68] | [-0.12 - -0.070] | [-0.16 - -0.084] |
| NFHS4 2015-16 dummy | 0.88*** | 0.94*** | -0.19*** | -0.21*** |
| | (0.034) | (0.036) | (0.012) | (0.014) |
| | [0.81 - 0.95] | [0.86 - 1.01] | [-0.22 - -0.17] | [-0.24 - -0.18] |
| State Fixed Effects | Yes | Yes | Yes | Yes |
| Mean of Upper Caste Hindus | -1.23 | -.1.23 | 0.32 | 0.32 |
| Observations | 213,665 | 213,665 | 213,665 | 213,665 |
| R-squared | 0.088 | 0.088 | 0.058 | 0.058 |

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

Notes: The tables show the results of a pooled OLS regression. The dependent variable in the first two columns is the height-for-age-z-score (HAZ score) based on the WHO anthropometric guidelines. The dependent variables in columns (3)-(4) is a dummy for being stunted in case the individual has a $HAZ < -2$. All regression include fixed effects for 36 states and union territories and the standard are clustered at the same level, and thus have a total of 36 clusters.

Table 2: Stunting gaps and the role of wealth and mother's education

| | Dependent Variable - Dummy for Stunted | | | | |
|---|---|---|---|---|---|
| | Baseline Sample | Wealth above 95th Percentile of Sample | Wealth below 5th Percentile of Sample | Education above 95th Percentile of Sample | Education below 5th Percentile of Sample |
| | (1) | (2) | (3) | (4) | (5) |
| SC-ST | 0.13*** | 0.058*** | 0.049** | 0.044** | 0.056*** |
| | (0.012) | (0.013) | (0.019) | (0.019) | (0.0087) |
| OBCs | 0.072*** | 0.0041 | 0.028 | 0.028** | 0.018* |
| | (0.013) | (0.011) | (0.018) | (0.013) | (0.0098) |
| Mean DV for Upper Castes | 0.32 | 0.19 | 0.52 | 0.19 | 0.51 |
| Other Controls | Yes | Yes | Yes | Yes | Yes |
| State Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 213,665 | 10,440 | 12,232 | 17,372 | 72,286 |
| R-squared | 0.058 | 0.012 | 0.042 | 0.012 | 0.043 |

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

Notes: The tables show the results of a pooled OLS regression. The dependent variable in all five columns is a dummy for being stunted in case the individual has a $HAZ < -2$. All regression include fixed effects for 36 states and union territories and the standard are clustered at the same level, and thus have a total of 36 clusters. The other controls include sex of the child, age in months, rural residence, and a set of dummies for the NFHS round. Column (1) considers the entire sample of Hindus for whom information on caste is available; columns (2) restricts the sample to households whose score on the wealth index is above or equal to the 95th percentile; columns (3) restricts the sample to households whose score on the wealth index is below or equal to the 5th percentile; columns (4) restricts the sample to mother's whose syears of schooling is above or equal to the 95th percentile; and finally, column (5) restricts the sample to mother's whose syears of schooling is below or equal to the 5th percentile.

Table 3: Effect of stunting at age 1 on stunting at older ages - Panel Data

| | Dummy for Stunting at | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Age 5 | | Age 8 | | Age 12 | | Age 15 | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Dummy Stunted at Age 1 | 0.45*** | 0.43*** | 0.39*** | 0.38*** | 0.30*** | 0.30*** | 0.23*** | 0.22*** |
| | (0.10) | (0.097) | (0.076) | (0.074) | (0.097) | (0.094) | (0.050) | (0.052) |
| SC-ST*Dummy Stunted at Age 1 | -0.016 | -0.0029 | 0.046 | 0.046 | 0.10 | 0.11 | 0.064 | 0.072 |
| | (0.11) | (0.10) | (0.088) | (0.086) | (0.092) | (0.087) | (0.057) | (0.057) |
| OBCs*Dummy Stunted at Age 1 | -0.049 | -0.038 | -0.072 | -0.071 | -0.030 | -0.024 | 0.016 | 0.023 |
| | (0.090) | (0.085) | (0.082) | (0.081) | (0.098) | (0.095) | (0.049) | (0.050) |
| SC-ST | 0.055 | 0.021 | 0.053 | 0.020 | 0.021 | -0.015 | 0.059 | 0.048 |
| | (0.038) | (0.040) | (0.033) | (0.037) | (0.030) | (0.033) | (0.036) | (0.040) |
| OBCs | 0.072* | 0.048 | 0.10*** | 0.081** | 0.086*** | 0.065** | 0.054 | 0.049 |
| | (0.037) | (0.040) | (0.030) | (0.032) | (0.024) | (0.028) | (0.032) | (0.036) |
| Basic Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Extended Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Sentinel site fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mean of children - Not stunted round 1 | 0.23 | 0.23 | 0.16 | 0.16 | 0.18 | 0.18 | 0.19 | 0.19 |
| Observations | 1,479 | 1,479 | 1,479 | 1,479 | 1,479 | 1,479 | 1,479 | 1,479 |
| R-squared | 0.204 | 0.220 | 0.213 | 0.220 | 0.154 | 0.160 | 0.113 | 0.117 |

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

Notes: The tables show the results of regressing stunting status at age 1 on the stunting status at ages of 5, 8, 12 and 15. The data comes from a panel of children in Andhra Pradesh who were followed as part of the Young Lives Project from the age of 1 in 2002 to the age of 15 in 2016. The anthropometric measurements were made at each point of time: 2002, 2006, 2009, 2013 and 2016. The data is restricted to children for whom the anthropometric measures are available at all five points of time. The dependent variable in columns(1)-(2); (3)-(4); (5)-(6); and (7)-(8) is a dummy for being stunted at ages of 5, 8, 12 and 15, respectively. The basic controlss include sex of child, age of child in months and rural urban residence. The extended controls besides the basic controls include the wealth index, household size and mothers years of education. All regression include fixed effects for 21 sentinel sites from which data was collected and the standard are clustered at the same level, and thus have a total of 21 clusters.

Table 4: Effect of stunting on learning outcomes at age 5 and 8 - Panel Data

| | Basic Controls (1) | Extended Controls (2) | Basic Controls (3) | Extended Controls (4) | Basic Controls (5) | Extended Controls (6) | Basic Controls (7) | Extended Controls (8) |
|---|---|---|---|---|---|---|---|---|
| | Learning outcomes at Age 5 | | | | | | | |
| | Std. PPVT Score | | PPVT Score Below 25thP | | Std. CDA Score | | CDA Score Below 25thP | |
| Dummy Stunted Age 1 | -0.17** | -0.096 | 0.12*** | 0.10*** | -0.25*** | -0.18*** | 0.083*** | 0.064*** |
| | (0.065) | (0.062) | (0.026) | (0.026) | (0.050) | (0.049) | (0.024) | (0.022) |
| Mean of DV - Children Not Stunted (Round 1) | 0.04 | 0.04 | 0.16 | 0.16 | 0.10 | 0.10 | 0.20 | 0.20 |
| Observations | 1,221 | 1,221 | 1,221 | 1,221 | 1,219 | 1,219 | 1,219 | 1,219 |
| R-squared | 0.309 | 0.392 | 0.084 | 0.120 | 0.135 | 0.196 | 0.095 | 0.122 |
| | Learning outcomes at Age 8 | | | | | | | |
| | Literacy Dummy | | No. of Correct Words per Minute | | Std. Math Score | | Word problem Subtract 4 from 20 | |
| Dummy Stunted Age 1 | -0.14*** | -0.12*** | -3.94*** | -3.16*** | -0.33*** | -0.28*** | -0.082** | -0.065** |
| | (0.027) | (0.027) | (0.92) | (0.90) | (0.062) | (0.059) | (0.031) | (0.031) |
| Std. CDA score Age 5 | | 0.082*** | | 3.56*** | | 0.24*** | | 0.087*** |
| | | (0.013) | | (0.42) | | (0.021) | | (0.020) |
| Mean of DV - Not Stunted (Round 1) | 0.61 | 0.61 | 23.5 | 23.5 | 0.20 | 0.20 | 0.77 | 0.77 |
| Observations | 1,416 | 1,416 | 1,348 | 1,348 | 1,418 | 1,418 | 1,316 | 1,316 |
| R-squared | 0.234 | 0.255 | 0.319 | 0.347 | 0.371 | 0.416 | 0.154 | 0.184 |

Notes: The tables show the results of regressing stunting status at age 1 on learning outcomes at age 5 and 8. The basic controls at age 5 include sex of child, age of child in months, rural urban residence, type of school, whether child is enrolled in pre-school and language of the PPVT test. The extended controls besides the basic controls include the wealth index, household size, monthly expenditure on food and non food items measured in real terms and whether the child has been enrolled in caste-based welfare program. At age 8 the basic controls include sex of child, age of child in months, rural urban residence, language of instruction in schools, the distance to school in minutes, whether the child has ever dropped out of school, the wealth index, household size, monthly expenditure on food and non food items measured in real terms. The extended controls includes the standardized score from the Cognitive Developmental Assessment conducted at age 5. All regression include fixed effects for 21 sentinel sites from which data was collected and the standard are clustered at the same level, and thus have a total of 21 clusters.

Table 5: Effect of stunting on learning outcomes at ages 12 and 15 - Panel Data

| | Basic Controls (1) | Extended Controls (2) | Basic Controls (3) | Extended Controls (4) | Basic Controls (5) | Extended Controls (6) | Basic Controls (7) | Extended Controls (8) |
|---|---|---|---|---|---|---|---|---|
| | Learning outcomes at Age 12 | | | | | | | |
| | Language Z-Score | | Math Z-Score | | English Z-Score | | PPVT Z-Score | |
| Dummy Stunted Age 1 | -0.33** | -0.20* | -0.33*** | -0.17* | -0.29*** | -0.19** | -0.22*** | -0.12* |
| | (0.11) | (0.10) | (0.087) | (0.087) | (0.059) | (0.064) | (0.064) | (0.058) |
| Literacy Status Age 8 | | 0.42*** | | 0.23** | | 0.38*** | | 0.24** |
| | | (0.093) | | (0.088) | | (0.092) | | (0.087) |
| Std. Math Score Age 8 | | 0.29*** | | 0.39*** | | 0.19*** | | 0.23*** |
| | | (0.044) | | (0.044) | | (0.033) | | (0.048) |
| Observations | 594 | 594 | 587 | 587 | 590 | 590 | 596 | 596 |
| R-squared | 0.285 | 0.400 | 0.412 | 0.530 | 0.474 | 0.540 | 0.330 | 0.407 |
| | Learning outcomes at Age 15 | | | | | | | |
| | Reading Z-Score | | | | | | Math Z-Score | |
| Dummy Stunted Age 1 | -0.32*** | -0.19** | | | | | -0.15 | -0.032 |
| | (0.091) | (0.063) | | | | | (0.098) | (0.069) |
| Std. Reading Score Age 12 | | 0.30*** | | | | | | 0.14*** |
| | | (0.036) | | | | | | (0.028) |
| Std. Math Score Age 12 | | 0.31*** | | | | | | 0.45*** |
| | | (0.049) | | | | | | (0.052) |
| Observations | 602 | 602 | | | | | 603 | 603 |
| R-squared | 0.246 | 0.458 | | | | | 0.352 | 0.550 |

Notes: The tables show the results of regressing stunting status at age 1 on learning outcomes at age 12 and 15. The basic controls at age 12 include sex of child, age of child in months, rural urban residence, type of school, language of instruction in school, the wealth index, household size, monthly expenditure on food and non food items measured in real terms. The extended controls additionally includes a dummy for whether the child could read at age 8 and the standardized score from the math test conducted at age 8. At age 15 the basic controls include include sex of child, age of child in months, rural urban residence, type of school, whether is currently enrolled in school, whether reports difficulty in getting to school, the wealth index, household size, monthly expenditure on food and non food items measured in real terms. The extended controls additionally includes standardized score from the language and math test conducted at age 12. All regression include fixed effects for 21 sentinel sites from which data was collected and the standard are clustered at the same level, and thus have a total of 21 clusters.

Table 6: The practice of untouchability and caste gaps in stunting

| | Dummy for Stunting | | | | | | |
|---|---|---|---|---|---|---|---|
| | SC Sample | | | | | | UC Sample |
| | Baseline Estimates (W1) | Mother Education (W2) | Mother and Health Inputs (W3, W4, W5) | Family Assets (W6) | Sanitation Proxies (W7) | All Controls (W8) | All Controls (W8) |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Practice untouchability | 0.31*** | 0.23*** | 0.18*** | 0.24*** | 0.23*** | 0.15*** | 0.052 |
| | (0.062) | (0.055) | (0.045) | (0.054) | (0.051) | (0.053) | (0.039) |
| | [0.19 - 0.44] | [0.12 - 0.34] | [0.090 - 0.27] | [0.13 - 0.35] | [0.12 - 0.33] | [0.047 - 0.26] | [-0.027 - 0.13] |
| Baseline Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Mother's Education | No | Yes | No | No | No | Yes | Yes |
| Health Inputs and other MotherControls | No | No | Yes | No | No | Yes | Yes |
| Asset Index | No | No | No | Yes | No | Yes | Yes |
| Sanitation Proxies | No | No | No | No | Yes | Yes | Yes |
| Observations | 27,925 | 27,925 | 27,925 | 27,925 | 27,925 | 27,925 | 21,192 |
| R-squared | 0.031 | 0.050 | 0.081 | 0.050 | 0.045 | 0.089 | 0.069 |

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

Notes: The tables show the results of regressing stunting status on the proportion of households who report practicing untouchability at the state level. The prevalence rates range from a minimum of 0 to a maximum 0.57. Estimates W1 refers to baseline estimates, which include controls for sex of the child, age in months and rural residence. The specification W2 adds the years of education of the mother as an additional control. The specification W3, W4, W5 accounts for the health inputs at birth - (i) place of delivery; (ii) prenatal access to doctor; (iii) breast fed immediately after birth; the mother's health status - (i) weight; (ii) height; (iii) height-for-age percentile; (iv) body mass index; (v) Rohrer index; (vi) age; (viii) age at first marriage; and mother's birth history - (i) number of of children who have died; (ii) total children born in last 5 years; (iii) age at first birth; (iv) birth order of child. The specification labeled W6 accounts for the factor score of the wealth index of the mother's household. The specification labeled W7 accounts for sanitation proxied by - (i) source of drinking water; (ii) type of toilet facility. Estimates labeled W8 account for all the above specified controls (W2-W7) in a single regression.

## Appendix

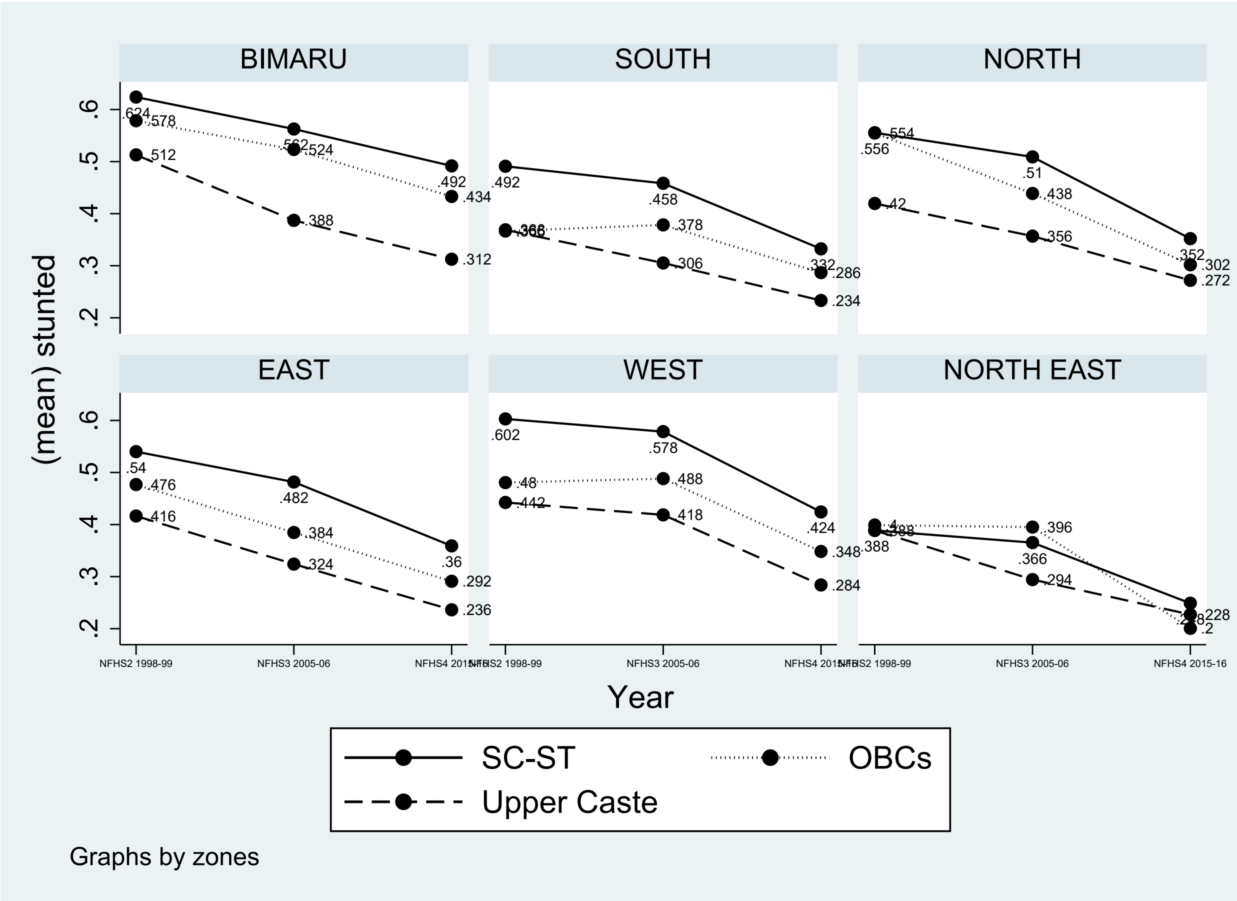


Figure 13: Evolution of stunting by caste groups and zones in India

The data on stunting comes from the three rounds of the NFHS - 1998-99, 2005-06 and 2015-16. The sample is restricted to all individuals whose information on caste is available and who report as belonging to the Hindu denomination. The graph is based on a total of 213,665 observations over the three rounds. The 29 states and 7 union territories are classified into the six regions as follows: (1) BIMARU comprising of Bihar, Chhattisgarh, Jharkhand, Madhya Pradesh, Rajasthan, Uttar Pradesh and Uttaranchal; (2) SOUTH comprising of Andaman and Nicobar Islands, Andhra Pradesh, Karnataka, Kerala, Pondicherry, Tamil Nadu and Telangana; (3) NORTH comprising of Chandigarh, Delhi, Haryana, Himachal Pradesh, Jammu & Kashmir and Punjab; (4) EAST comprising of Assam, Orissa and West Bengal; (5) WEST comprising of Dadra & Nagar Haveli, Daman & Diu, Goa, Gujarat, Lakshdweep and Maharashtra; and finally (6) NORTH-EAST comprising of Arunachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim and Tripura.

Table 7: Socioeconomic factors and caste gaps in stunting

| | DV: Dummy for Stunting | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Baseline Estimates | Mother's Education | Health Inputs | Mother's Health Status | Mother's Birth History | Family Assets | Sanitation Proxies | All Controls |
| | (W1) | (W2) | (W3) | (W4) | (W5) | (W6) | (W7) | (W8) |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | |
| SC-ST | 0.13*** | 0.080*** | 0.11*** | 0.078*** | 0.11*** | 0.077*** | 0.099*** | 0.036*** |
| | (0.011) | (0.0093) | (0.0098) | (0.0088) | (0.010) | (0.0089) | (0.0095) | (0.0085) |
| OBCs | 0.066*** | 0.036*** | 0.056*** | 0.038*** | 0.058*** | 0.040*** | 0.050*** | 0.015** |
| | (0.013) | (0.0085) | (0.011) | (0.0088) | (0.011) | (0.0083) | (0.0097) | (0.0066) |
| Yrs. of Education | | -0.014*** | | | | | | |
| | | (0.00070) | | | | | | |
| No Toilet facility | | | | | | | 0.13*** | |
| | | | | | | | (0.0099) | |
| Public open well | | | | | | | 0.084*** | |
| | | | | | | | (0.018) | |
| Mother's weight | | | | 0.0035*** | | | | |
| | | | | (0.0011) | | | | |
| Mother's height | | | | -0.0019*** | | | | |
| | | | | (0.00020) | | | | |
| ht/a percentile | | | | -1.8e-07 | | | | |
| | | | | (3.6e-06) | | | | |
| Mother's BMI | | | | -0.0022*** | | | | |
| | | | | (0.00068) | | | | |
| Mother's rohrer's index | | | | 0.0020*** | | | | |
| | | | | (0.00065) | | | | |
| age at first marriage | | | | -0.0081*** | | | | |
| | | | | (0.00100) | | | | |
| current age | | | | 0.0019*** | | | | |
| | | | | (0.00055) | | | | |
| place of delivery | | | -0.0035*** | | | | | |
| | | | (0.00042) | | | | | |
| prenatal: doctor | | | -0.057*** | | | | | |
| | | | (0.0068) | | | | | |
| Wealth index factor scores | | | | | | -8.8e-07*** | | |
| | | | | | | (5.7e-08) | | |
| Observations | 157,467 | 157,464 | 157,467 | 157,467 | 157,467 | 157,467 | 157,467 | 157,464 |
| R-squared | 0.060 | 0.077 | 0.066 | 0.102 | 0.066 | 0.076 | 0.070 | 0.114 |

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

Notes: The tables show the results of a pooled OLS regression The dependent variables in columns (1)-(8) is a dummy for being stunted in case the individual has a $HAZ < -2$. All regression include fixed effects for 36 states and union territories and the standard are clustered at the same level, and thus have a total of 36 clusters.The tables evaluates the effects of including a set of socioeconomic controls W7-W8 outlined in Figure 4 on the caste dummies. The table reports the coefficients on the caste dummies, as well as the coefficients on a subset of the control variables outlined in Section 2.1.

Table 8: Caste gaps in stunting by age - Panel Data

| | Dummy for Stunting at | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Age 1 | | | Age 5 | | Age 8 | | Age 12 | | Age 15 | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| SC-ST | 0.21*** | 0.16*** | 0.14*** | 0.15*** | 0.097*** | 0.16*** | 0.10*** | 0.13*** | 0.081** | 0.13*** | 0.11*** |
| | (0.031) | (0.029) | (0.036) | (0.026) | (0.030) | (0.029) | (0.030) | (0.033) | (0.030) | (0.035) | (0.037) |
| OBCs | 0.11*** | 0.077*** | 0.073*** | 0.11*** | 0.069* | 0.12*** | 0.088*** | 0.11*** | 0.078** | 0.083** | 0.070* |
| | (0.026) | (0.022) | (0.022) | (0.035) | (0.038) | (0.029) | (0.027) | (0.032) | (0.030) | (0.038) | (0.040) |
| Wealth index | | -0.28*** | -0.31** | | -0.38*** | | -0.35*** | | -0.28*** | | -0.27*** |
| | | (0.092) | (0.11) | | (0.098) | | (0.059) | | (0.087) | | (0.083) |
| Number of antenatal visits of mother | | | -0.0053 | | | | | | | | |
| | | | (0.0062) | | | | | | | | |
| Mother was attended by skilled health personnel | | | -0.047* | | | | | | | | |
| | | | (0.024) | | | | | | | | |
| Mother recd. at least two injections for tetanus | | | -0.056 | | | | | | | | |
| | | | (0.10) | | | | | | | | |
| Child received BCG vaccination | | | -0.14 | | | | | | | | |
| | | | (0.080) | | | | | | | | |
| Child received vaccination against measles | | | 0.030 | | | | | | | | |
| | | | (0.039) | | | | | | | | |
| Child received vaccination against polio | | | 0.00087 | | | | | | | | |
| | | | (0.11) | | | | | | | | |
| household size | | 0.0038 | 0.0029 | | -0.018*** | | -0.0048 | | 0.0020 | | -0.0047 |
| | | (0.0052) | (0.0060) | | (0.0051) | | (0.0039) | | (0.0050) | | (0.0071) |
| Mother's level of education | | -0.0040 | -0.0028 | | -0.0011 | | -0.0025 | | -0.0039 | | 0.0012 |
| | | (0.0031) | (0.0039) | | (0.0036) | | (0.0023) | | (0.0031) | | (0.0031) |
| Girl Dummy | -0.068*** | -0.068*** | -0.064*** | -0.030 | -0.027 | -0.032 | -0.036 | 0.016 | 0.013 | -0.031 | -0.033 |
| | (0.021) | (0.020) | (0.018) | (0.026) | (0.026) | (0.028) | (0.027) | (0.023) | (0.021) | (0.026) | (0.025) |
| Child's age (in months) | 0.017*** | 0.017*** | 0.018*** | -0.00049 | -0.00013 | -0.0038 | -0.0031 | 0.0016 | 0.0022 | 0.0013 | 0.0018 |
| | (0.0033) | (0.0033) | (0.0040) | (0.0046) | (0.0044) | (0.0037) | (0.0038) | (0.0039) | (0.0039) | (0.0037) | (0.0038) |
| Rural Dummy | 0.44*** | 0.33*** | 0.32*** | -0.18* | -0.21** | -0.022 | -0.048 | 0.013 | -0.037 | 0.034 | 0.0070 |
| | (0.029) | (0.042) | (0.043) | (0.091) | (0.077) | (0.047) | (0.035) | (0.032) | (0.036) | (0.061) | (0.057) |
| Mean DV- Upper castes | 0.17 | 0.17 | 0.17 | 0.23 | 0.23 | 0.15 | 0.15 | 0.17 | 0.17 | 0.15 | 0.15 |
| Sentinel site fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 1,479 | 1,479 | 1,269 | 1,479 | 1,479 | 1,479 | 1,479 | 1,479 | 1,479 | 1,479 | 1,479 |
| R-squared | 0.112 | 0.122 | 0.129 | 0.058 | 0.081 | 0.079 | 0.093 | 0.047 | 0.056 | 0.047 | 0.052 |

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Notes: The tables show the results of regressing stunting status on caste dummies and a set of controls at age of 1, 5, 8, 12 and 15. The data comes from a panel of children in Andhra Pradesh who were followed as part of the Young Lives Project from the age of 1 in 2002 to the age of 15 in 2016. The anthropometric measurements were made at each point of time: 2002, 2006, 2009, 2013 and 2016. The data is restricted to children for whom the anthropometric measures are available at all five points of time. The dependent variable in columns(1)-(2); (3)-(4); (5)-(6);(7)-(8), and (9)-(10) is a dummy for being stunted at ages of 1, 5, 8, 12 and 15, respectively. All regression include fixed effects for 21 sentinel sites from which data was collected and union and the standard are clustered at the same level, and thus have a total of 21 clusters.