

Pre-Analysis Plan for “Parental Misperceptions on Child Nutrition in India: Implications for Child Feeding Practices and Growth”

15 March 2025

PI: Sneha Nimmagadda, Department of Economics, University of Southern California

1. Background

Child undernutrition is a major public health challenge worldwide and has been shown to have significant effects on mortality, physical and cognitive development, school completion rates and test scores, and earnings in adulthood (Strauss and Thomas 2007, Smith et al. 2012, Hoddinott et al. 2008, Hoddinott et al. 2013, Deshpande and Ramachandran 2022, Galasso and Wagstaff 2019). India accounts for almost one-third of all the world’s stunted children and half the world’s wasted children (Fanzo et al. 2019). The most recent wave of the National Family Health Survey (NFHS-5, 2019-20) showed that 35.5%, 19.3%, and 32.1% of children under 5 years of age were, respectively, stunted, wasted, and underweight, with these numbers representing only modest improvements over the past decade despite significant economic growth and rising real incomes across all sections of the population.

Previous studies have also documented a dissonance between actual and self-reported diet adequacy among individuals in India. Data from the National Sample Survey (2004-05) showed that the share of well-nourished rural households was only 24.2%, but 97.5% of the respondents reported that “everyone in their household got enough food every day” (Deaton and Dreze 2009). A policy report by the International Growth Center suggested that even in households with malnourished children, parents often believe that their children are adequately nourished, treating government supplementary nutrition programs as optional rather than essential (Abraham and Fraker 2014).

This project aims to examine the role of parental misperceptions and information gaps in contributing to the persistently high rates of child undernutrition in India. It is guided by two core hypotheses:

1. **Parents systematically overestimate the nutritional status of their children:** If parents form expectations about how healthy their child is by observing other children around them, then parents in areas with high levels of stunting and wasting may be more likely to believe that their own child is relatively healthy and have a skewed perception of “ideal” height and weight levels.
2. **Parents systematically underestimate the effects of child malnutrition on long-term health, education, and labor market outcomes:** While there is a large literature documenting the effects of child malnutrition on the incidence of infectious and chronic diseases, years of education, test scores, and earnings in

adulthood, this information is most likely not common knowledge among parents in India, particularly in rural areas.

These misperceptions, if proven true, may create a suboptimal equilibrium for child nutrition outcomes, trapping families in a cycle of inadequate nutrition. Moreover, improving child nutrition may have important externalities that are currently not accounted for by parents and researchers alike.

This study is based on the premise that updating mothers' beliefs on their child's true relative nutritional status, and the effects of child malnutrition on health, education, and labor market outcomes, will improve child feeding practices and health outcomes. The specific research questions that the project seeks to answer are:

1. Do mothers systematically overestimate the nutritional status (height- and weight-for-age percentiles) of their children, relative to global WHO standards and other children in their region?
2. Do mothers underestimate the effects of child malnutrition on long-term health, education, and labor market outcomes?
3. What mechanisms could explain the formation of such misperceptions in the first place? Are mothers with higher exposure to undernourished children more likely to overestimate their children's nutritional status?
4. Would updating mothers' beliefs about a) their children's true relative height-for-age and weight-for-age percentiles, and/or b) the long-term effects of child malnutrition, improve child feeding practices, utilization of government nutrition services, and child growth outcomes?

2. Experimental Design

The research design involves an individual-level field experiment with mothers of children aged 7 to 24 months from the Mancherla and Kumarambheem-Asifabad districts of Telangana, India, with two treatment arms and a control arm:

- **Treatment arm 1 (N = 500):** Provide mothers information on the true height-for-age and weight-for-age percentiles of their child relative to a reference group of healthy children based on WHO standards
- **Treatment arm 2 (N = 500):** Provide mothers information on the true height-for-age and weight-for-age percentiles of their child AND information on the effects of child malnutrition on health (risk of chronic and infectious diseases, mortality), education (high school test scores, years of education), and labor market (earnings) outcomes, synthesized from existing literature
- **Control arm (N = 500):** Status-quo, no intervention

These two districts were specifically chosen due to their varying rates of child stunting: Kumarambheem-Asifabad (58.1%) and Mancherial (33.9%). The target age range of children (7-24 months) was chosen as this is a critical window where growth faltering typically occurs among children and where nutritional interventions have been found to be most effective (Behrman 2015, Victora et al. 2010, Alderman and Headey 2018).

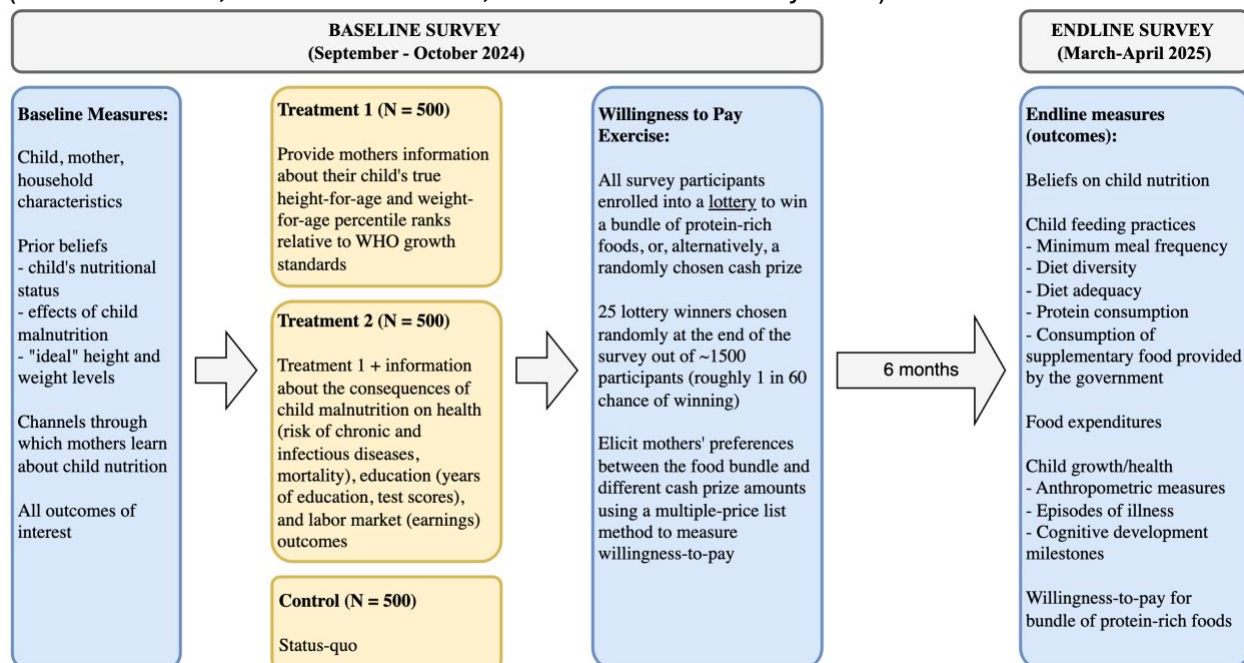


Figure 1. Experimental Design and Timeline

Sampling and Stratification

Government administrative data from the Nutrition and Health Tracking System (NHTS) in the state of Telangana was used to construct the sampling frame for this project. These data contain monthly height and weight measurements for all children under the age of 5 years registered in Anganwadi centers (community health centers) across the state (N > 1.8 million), along with variables such as the date of birth, sex, and caste of the children, and certain Anganwadi worker characteristics.

The sample selection was done in two stages –

Sampling of Anganwadi centers:

Using NHTS data from August 2024, 170 centers were randomly sampled from the universe of Anganwadi centers in the Mancherial and Kumarambheem-Asifabad districts in Telangana. The randomization was stratified by a measure of average child nutrition at the Anganwadi center level, constructed using the administrative child growth monitoring data.

85 “high nutrition” centers and 85 “low nutrition” centers were randomly sampled. This stratification will be used to assess if mothers with higher exposure to undernourished children are more likely to overestimate their children’s nutritional status. The 170 sampled centers were then randomly assigned to either a “high treatment-intensity” group (N = 115 centers) or a “low treatment-intensity” group (N = 55 centers) to evaluate spillover effects.

Sampling of children:

Between 10-20 children were sampled from the roster of children aged 7 months to 24 months at each selected center and assigned to one of the three treatment arms using a block randomization technique. Children were stratified by sex and a nutrition score based on anthropometric z-scores. In the high treatment-intensity centers, ~75% of the children were assigned to one of the two treatment arms on average, and in the low treatment-intensity centers, ~40% of the children were assigned one of the two treatment arms.

Children were oversampled at this stage to account for unavailability during the survey. The target was to achieve up to 10 completed interviews at each Anganwadi center, for a total sample of 1500 children across the three arms.

Exclusion:

1. Children with missing or implausible anthropometric z-scores (height-for-age, weight-for-age, and weight-for-height) in August 2024 were excluded from the sampling frame.
2. Children with height-for-age, weight-for-age, AND weight-for-height z-scores above the WHO median (0) were also excluded from the sampling frame since the interventions are primarily targeted towards children at lower levels of nutrition.
3. Children with chronic medical conditions or special needs were excluded as the nutritional needs of such children may be different and require medical expertise/supervision.

Power Calculations

Power calculations were conducted for the primary and secondary outcomes using data collected from a pilot study conducted in July 2023 and the fifth wave of the National Family Health Survey in India (NFHS-5, 2019-20). I assumed individual-level randomization with measures at baseline and endline, a significance level of 0.05, power of 0.8, and equal allocation between treatment arms. I set the correlation between baseline and follow-up measurements as 0.8 for the anthropometric outcomes and 0.6 for all the other outcomes. With a sample size of 500 children in each arm under the above assumptions, we have 80% power to detect effects between 0.10-0.15 standard deviations from standardized counterfactual levels for all non-anthropometric outcomes, and 0.14-0.20 points for the anthropometric z-score outcomes (relative effect ~0.11 SD). The stratification during sampling and the inclusion of covariates in the analysis should further improve power.

3. Outcomes

The outcomes for this study fall into 6 main domains: (1) Beliefs about child's nutritional status, (2) Beliefs about the effects of child malnutrition, (3) Child Feeding, (4) Child Growth, (5) Willingness-to-pay for protein-rich foods, and (6) Child Health.

Primary Outcomes

A) Beliefs about child's nutritional status

BL1: Difference between true and perceived height-for-age percentile of child relative to WHO standards

BL2: Difference between true and perceived weight-for-age percentile of child relative to WHO standards

The child's true height-for-age and weight-for-age percentiles will be calculated according to World Health Organization guidelines. Child height and weight will be measured during the baseline and endline surveys. Child age will be measured in completed months since birth (integer values), and sex will be defined as biological sex at birth. These percentiles will be compared to mothers' perceived percentile ranks elicited during the baseline and endline surveys.

B) Beliefs about the effects of child malnutrition

BL3: Knowledge score on the effects of child malnutrition on long-term health, education, and labor market outcomes (continuous)

This indicator will be constructed based on mothers' responses to "Likert-scale" questions eliciting beliefs on whether malnutrition (stunting or wasting) before the age of 3 has any effects on (1) years of education, (2) school test scores, (3) likelihood of early death, (4) adult height, (5) likelihood of cardiometabolic diseases in adulthood, (6) and earnings in adulthood. The indicator will equal the number of "correctly" answered prompts.

C) Child Feeding

CF1: Child Dietary Diversity Score (continuous)

CF2: Child Minimum Dietary Diversity (binary)

CF3: Child Minimum Meal Frequency for children aged 6-23 months

CF4: Whether child consumed Balamrutham in last 24 hours

CF5: Quantity of Balamrutham consumed in last 24 hours

CF6: Estimated protein consumption of child

CF7: Number of eggs consumed in the last calendar month

CF1 will be constructed based on the WHO index, with scores ranging from 0 to 8 for children aged 6-23 months, and 0 to 7 for children older than 23 months, equal to the number of defined food groups from which the child has consumed food over the previous day: (1) breast milk (for children aged 6-23 months only); (2) grains, roots, tubers and plantains; (3) pulses (beans, peas, lentils), nuts and seeds; (4) dairy products (milk, infant formula, yogurt, cheese); (5) flesh foods (meat, fish, poultry, organ meats); (6) eggs; (7) vitamin-A rich fruits and vegetables; and (8) other fruits and vegetables. The binary indicator (CF2) will indicate whether the child has received food from five or more of those food groups (or four or more for children over 23 months) over the previous day.

CF3 is a binary indicator for children aged 6-23 months coded “1” if the child was fed the minimum recommended number of times for their age, based on WHO guidelines. The minimum number of times is defined as: - two feedings of solid, semi-solid or soft foods for breastfed infants aged 6–8 months; - three feedings of solid, semi-solid or soft foods for breastfed children aged 9–23 months; and - four feedings of solid, semi-solid or soft foods or milk feeds for non-breastfed children aged 6–23 months whereby at least one of the four feeds must be a solid, semi-solid or soft feed.

Balamrutham is a therapeutic supplementary food provided by the state government to children aged 6 months to 5 years. CF4 is a binary variable coded “1” if the child consumed any Balamrutham the previous days, and “0” otherwise”. CF5 is defined as the quantity of Balamrutham consumed the previous day, measured using a standardized bowl across all respondents (bowlfuls in fractional terms – eg. 0.25 bowl or 1.5 bowls).

For CF6, we measure the types (milk, dairy products, eggs, pulses, nuts and seeds, flesh foods, Balamrutham) and quantities (in standardized bowls) of protein-rich foods that the child has consumed the previous day. Guidelines from the National Institute of Nutrition will be used to calculate the grams of protein consumed by the child. This is intended to be a rough approximate estimate of the child’s daily protein intake.

The government provides 16 eggs each month as supplementary food to children under the age of 3 years. Indicator CF7 will measure the number of eggs consumed by the child in the last calendar month.

D) Child Growth

CG1: Child Height-for-Age Z-score (HAZ)

CG2: Child Weight-for-Age Z-score (WAZ)

CG3: Child Weight-for-Height Z-score (WHZ)

CG4: Child Mid-Upper Arm Circumference (MUAC)

The three z-scores will be measured in standard deviation units following WHO guidelines, and MUAC will be measured in centimeters. Child height, weight, and MUAC will be directly measured during the baseline and endline surveys. Since the gap between the baseline and endline surveys is only 6 months, I don't expect to find significant treatment effects on the height-for-age z-scores within this timeframe. For longer-term follow-up on changes in height, I will also collect monthly height and weight measures for the children from government administrative data for 12 months after the endline survey.

E) Willingness-to-Pay for a bundle of protein-rich foods

WTP1: Willingness-to-Pay for a bundle of protein-rich foods

WTP2: Percentage of mothers with willingness-to-pay equal to or above market price for a bundle of protein-rich foods

Mothers' willingness-to-pay for a bundle of protein-rich foods was measured at the end of the baseline survey after the information treatments were administered, and will be measured again during the endline survey.

All mothers who participated in the baseline survey were entered into a lottery to win a bundle of protein-rich food items for their child or, alternatively, a randomly chosen cash prize (amount ranged Rs. 100 to Rs. 2000). 25 lottery "winners" were chosen randomly at the end of the survey. Mothers were asked to state their preferences between the food bundle and several potential cash prize amounts, using a multiple-price-list elicitation method. One cash prize amount will be randomly chosen for each mother, and their choice for that amount will be implemented in case they win the lottery. WTP1 will be measured by the mid-point of the interval of two cash amounts at which a mother switches from preferring to receive cash to preferring to receive food. WTP2 is a binary indicator coded "1" if the mother's willingness-to-pay is equal to or above the market price, and "0" otherwise.

This exercise will be repeated during the endline survey. The potential cash prize amount will range from Rs. 100 to Rs. 10000 (increased from Rs. 2000 during the first round of the survey).

Secondary Outcomes

A) Beliefs about child's nutritional status

BL4: Percentage of mothers who correctly categorize their child's height status (normal or stunted)

BL5: Percentage of mothers who correctly categorize their child's weight status (normal or underweight/wasted)

BL6: Percentage of mothers who overestimate their child's height-for-age percentile by WHO standards by more than 10 points

BL7: Percentage of mothers who overestimate their child's weight-for-age percentile by WHO standards by more than 10 points

BL8: Mothers' perceived "ideal" height for their child's age

BL9: Mothers' perceived "ideal" weight for their child's age

BL4, BL5: The child's height status will be defined based on their height-for-age z-score (stunted = $HAZ < -2$; normal = $HAZ \geq -2$), calculated based on WHO growth standards. The weight status will be calculated based on weight-for-age and weight-for-height z-scores (underweight/wasted = WAZ or $WHZ < -2$; normal = WAZ and $WHZ \geq -2$).

BL6 and BL7 will be measured in the same way as BL1 and BL2 described above.

BL8 and BL9 measure mothers' beliefs about the "ideal" height and weight for their child's age, in centimeters and kilograms, respectively.

B) Beliefs about the effects of child malnutrition

BL10: Knowledge score on the effects of child malnutrition on short-term and long-term outcomes (continuous)

BL11: Knowledge score on the effects of child malnutrition on short-term outcomes (continuous)

BL12: Knowledge score on the effects of additional protein intake on short-term and long-term outcomes (continuous)

These indicators will be based on an index of questions on the beliefs of mothers about the effects of child malnutrition and additional protein intake during childhood on various short-term and long-term outcomes.

C) Child Feeding

CF8: Child Minimum Meal Frequency for all children (6-30 months)

CF9: Child Minimum Diet Adequacy for all children (6-30 months)

CF10: Quantity of Balamrutham consumed in last calendar month

The WHO guidelines only define the diet diversity, minimum feeding frequency, and diet adequacy indicators for children aged 6-23 months. The children in the study sample were all in this age range at the time of the baseline survey, but older at the time of the endline survey (up to 29 months). I will construct equivalent indicators for children of all ages (6 to 30 months) as secondary outcomes. CF9 is a binary variable

coded "1" if the child meets the minimum meal frequency (as defined in CF3) AND the minimum diet diversity (CF2) criteria defined above.

The government provides a 2.5kg packet of Balamrutham each month as supplementary food to children aged 6 months to 3 years. Indicator CF9 will measure the quantity of Balamrutham consumed by the child in the last calendar month, as a fraction of this packet size (in grams).

D) Child Growth

CG5: Change from baseline in height-for-age z-score

CG6: Change from baseline in weight-for-age z-score

CG7: Change from baseline in weight-for-age z-score

CG8: Change from baseline in mid-upper arm circumference

We will calculate the difference between the baseline survey measure and endline survey measure for each of these indicators. Monthly administrative height and weight data will be collected for 12 months after the endline survey to measure longer-term changes.

E) Child Health

CH1: Any episode of illness in the last 14 days (binary)

CH2: Child cognition z-score

CH1 is a binary indicator coded "1" if the child experienced any episode of illness in the last 14 days, and "0" otherwise.

The child cognition z-scores will be calculated based on the Caregiver-Reported Early Development Instruments (CREDI) Short Form tool, which is a validated set of 20 population-level measures of early childhood development (ECD) for children from birth to age three (0-36 months) using self-reported data from a caregiver (mother). The responses on this 20-point scale (based on age) will be converted to a norm-referenced standardized Z-score for overall development. The z-scores may range from -6 to +6, with larger scores representing better outcomes.

F) Household Food and non-Food Expenditures

HE1: Household expenditure on food in the last calendar month (Rs.)

HE2: Household expenditures on non-food items in the last calendar month (Rs.)

HE3: Fraction on last month's income spent on food

HE1 and HE2 measure food and non-food expenditures in the last calendar month at the household level. These will be used to estimate the fraction of monthly income that is going towards food versus other expenses.

4. Empirical Model

The main analysis will compare outcomes for individuals randomized across the three study arms: Treatment group 1 vs. Treatment group 2 vs. Control group.

Consider an outcome, Y , such as an indicator measuring the difference between the true and perceived height-for-age percentile of a child relative to WHO standards. For child i , the estimating equation is:

$$Y_i = \beta_0 + \beta_1 I(Treatment1)_i + \beta_2 I(Treatment2)_i + \beta_3 X_i + \epsilon_i \quad (1)$$

where $I(Treatment1)_i$ is an indicator variable equal to “1” if the child was randomized to treatment group 1 and “0” otherwise, and $I(Treatment2)_i$ is an indicator variable equal to “1” if the child was randomized to treatment group 2 and “0” otherwise. β_1 and β_2 are the parameters of interest and measures the causal effect of being randomized into each of the treatment groups. Since the randomization is at the individual level, the standard errors will not be clustered.

Linear models will be used to examine causal treatment effects as is standard practice in the economics literature, even for binary outcomes. Additionally, I will check robustness to model specification.

X_i is a vector of control variables, which will include a set of demographic and socio-economic variables (eg. child age in months, child sex, birth order, mother’s education in years, household poverty probability index, household caste category (SC, ST, or OBC)), baseline anthropometric measures (child’s baseline height-for-age, weight-for-age, and weight-for-height z-scores, mother’s height), as well as a lag of the dependent variable when available. The final set of characteristics to be included in X_i will be selected using post-double selection LASSO (Belloni et al. 2014, Cilliers et al. 2024).

As a secondary model, I will also estimate a difference-in-differences (DID) specification as follows:

$$Y_{it} = \beta_0 + \beta_1 I(Treatment1)_i + \beta_2 I(Treatment2)_i + \beta_3 (Endline_t) + \beta_4 (Endline * Treatment1)_{it} + \beta_4 (Endline * Treatment2)_{it} + \beta_3 X_i + \epsilon_{it} \quad (2)$$

where *Endline* is a time dummy variable equal to “1” if the outcome measure is from the endline survey and “0” if from the baseline survey, and Y_{it} is the outcome for individual i at time t . The other variables are as defined above.

Heterogeneity analysis

I will test whether treatment effects are heterogeneous depending on individual child, household, or community characteristics. Relevant interaction terms will be included in the main specifications for these characteristics:

1. Child sex
2. Whether the child was malnourished (stunted, underweight, or wasted) at baseline
3. Baseline beliefs on child’s relative nutritional status (gap between true and perceived height-for-age and weight-for-age percentiles)
4. Baseline beliefs on the effects of child malnutrition
5. Average child nutrition (mean z-scores) at the Anganwadi center level

I will also examine other dimensions of heterogeneity, including child age-group, baseline child diet adequacy, household income/wealth, mother’s education, and household size, but will only report these in the paper if they meaningfully affect the interpretation of the main results.

Attrition

Since it is common for new mothers in rural India to split their time between their natal and marital homes, some mothers interviewed during the baseline survey may not be available at the same place during the endline survey. To minimize attrition, we will contact the mothers by phone 2-3 weeks before the endline survey, inform them about the expected dates of the endline survey, and enquire about their availability. In case they are not available at the original survey location, we will try to visit them at their current location, as long as it is within the state of Telangana.

I will test for differential attrition across the treatment and control groups, and report Lee bounds on the estimated effects in case there is differential attrition.

Exploratory Analyses and Mechanisms

Spillover Effects:

I will examine spillover effects on control group households in the “high treatment-intensity” Anganwadi centers versus control group households in the “low treatment-intensity”

Anganwadi centers. This will help assess information sharing across households within communities and measure potential contamination effects, if any.

Mechanisms:

The study also aims to explore mechanisms through which misperceptions about children's nutritional status are formed. I will specifically examine two possible mechanisms:

- (a) if mothers with higher exposure to other malnourished children are more likely to overestimate their child's nutritional status, and
- (b) if mothers who receive incorrect information about their child's nutritional status from the Anganwadi worker are more likely to overestimate their child's nutritional status.

To measure exposure of mothers to malnourished children, I will construct measures of average child nutrition (mean z-scores) and malnutrition rates across several reference groups:

- (a) Other children within the same Anganwadi center
- (b) Other children of the same caste category within the same Anganwadi center
- (c) Children in the mother's reported peer network
- (d) Children from the mother's natal village

I will also explore if characteristics like child sex, birth order, mother's education, household income/wealth, or the number of registered children in the Anganwadi center are associated with higher misperceptions.

References

Abraham, Ronald, and Andrew Fraker. "Bihar's malnutrition crisis and potential solutions." International Growth Centre Blog (2014).

Alderman, Harold, and Derek Headey. "The timing of growth faltering has important implications for observational analyses of the underlying determinants of nutrition outcomes." PloS one 13.4 (2018): e0195904.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. "Inference on treatment effects after selection among high-dimensional controls." Review of Economic Studies 81.2 (2014): 608-650.

Behrman, Jere R. "Growth faltering in the first thousand days after conception and catch-up growth." The Oxford handbook of economics and human biology (2015).

Cilliers, Jacobus, Nour Elashmawy, and David McKenzie. "Using post-double selection Lasso in field experiments." World Bank (2024).

Deaton, Angus, and Jean Drèze. "Food and nutrition in India: facts and interpretations." Economic and political weekly (2009): 42-65.

Deshpande, Ashwini, and Rajesh Ramachandran. "Early childhood stunting and later life outcomes: A longitudinal analysis." Economics & Human Biology 44 (2022): 101099.

Fanzo, Jessica, et al. "2018 Global Nutrition Report." (2019).

Galasso, Emanuela, and Adam Wagstaff. "The aggregate income losses from childhood stunting and the returns to a nutrition intervention aimed at reducing stunting." Economics & Human Biology 34 (2019): 225-238.

Glewwe, Paul, Hanan G. Jacoby, and Elizabeth M. King. "Early childhood nutrition and academic achievement: a longitudinal analysis." Journal of public economics 81.3 (2001): 345-368.

Hoddinott, John, et al. "Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults." The lancet 371.9610 (2008): 411-416.

Hoddinott, John, et al. "Adult consequences of growth failure in early childhood." *The American journal of clinical nutrition* 98.5 (2013): 1170-1178.

International Institute for Population Sciences (IIPS) and ICF. "National Family Health Survey (NFHS-5), 2019-21: India: Volume II." Mumbai: IIPS (2021).

Smith, James P., et al. "The effects of childhood health on adult health and SES in China." *Economic development and cultural change* 61.1 (2012): 127-156.

Strauss, John, and Duncan Thomas. "Health over the life course." *Handbook of development economics* 4 (2007): 3375-3474.

Victora, Cesar Gomes, et al. "Worldwide timing of growth faltering: revisiting implications for interventions." *Pediatrics* 125.3 (2010): e473-e480.

Westfall, Peter H., and S. Stanley Young. "Resampling-based multiple testing: Examples and methods for p-value adjustment." John Wiley & Sons (1993).