

INTERNATIONAL
FOOD POLICY
RESEARCH
INSTITUTE

IFPRI

IFPRI Discussion Paper 02461

November 2026

Social Desirability Bias in Program Evaluation

The Case of a Childhood Nutrition Program in Sri Lanka

Bailey Peterson-Wilhelm

Benjamin Schwab

Dilini Hemachandra

Derek Headey

Development Strategies and Governance Unit

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), a CGIAR Research Center established in 1975, provides research-based policy solutions to sustainably reduce poverty and end hunger and malnutrition. IFPRI's strategic research aims to foster a climate-resilient and sustainable food supply; promote healthy diets and nutrition for all; build inclusive and efficient markets, trade systems, and food industries; transform agricultural and rural economies; and strengthen institutions and governance. Gender is integrated in all the Institute's work. Partnerships, communications, capacity strengthening, and data and knowledge management are essential components to translate IFPRI's research from action to impact. The Institute's regional and country programs play a critical role in responding to demand for food policy research and in delivering holistic support for country-led development. IFPRI collaborates with partners around the world.

AUTHORS

Bailey Peterson-Wilhelm (bailey.petersonwilhelm@usask.ca) was a PhD Candidate in the Department of Agricultural Economics at Kansas State University, Manhattan, KS, when she wrote this work. She is currently an Assistant Professor in the Department of Agricultural and Resource Economics, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

Benjamin Schwab (benschwab@ksu.edu) is an Associate Professor in the Department of Agricultural Economics at Kansas State University, Manhattan, KS.

Dilini Hemachandra (dilinisp@gmail.com) is a Senior Lectureer in the Department of Agricultural Economics and Business Management at the University of Peradeniya, Kandy, Sri Lanka.

Derek Headey (D.Headey@cgiar.org) is a Senior Research Fellow in the Development Strategies and Governance (DSG) Unit at the International Food Policy Research Institute, Colombo, Sri Lanka.

Notices

[1] IFPRI Discussion Papers contain preliminary material and research results and are circulated in order to stimulate discussion and critical comment. They have not been subject to a formal external review via IFPRI's Publications Review Committee. Any opinions stated herein are those of the author(s) and are not necessarily representative of or endorsed by IFPRI.

[2] The boundaries and names shown and the designations used on the map(s) herein do not imply official endorsement or acceptance by the International Food Policy Research Institute (IFPRI) or its partners and contributors.

[3] Copyright remains with the authors. The authors are free to proceed, without further IFPRI permission, to publish this paper, or any revised version of it, in outlets such as journals, books, and other publications.

Contents

LIST OF TABLES iv

LIST OF FIGURES v

ABSTRACT vi

ACKNOWLEDGMENTS vii

ACRONYMS viii

Introduction 1

- Background 4
- UNICEF Program 5

Research Design 8

- Data 8
- Direct Questions 9
  - Direct 9
  - List Experiments 10

Estimation Strategy 12

- List Experiment Assumptions 13
- Social Desirability Bias 15

Results 18

- Robustness Checks 18
- Ability to Meet Basic Food Needs 19
- Dietary Diversity 20

Discussion 24

Conclusions 28

REFERENCES 40

Appendix 44

LIST OF TABLES

Table 1: Summary Statistics & Randomization Assumption Balance Table 32
Table 2: List Experiment Instructions and Questions 33
Table 3: No Design Effect Assumption Tests 34
Table 4: No Liars Assumption: Frequency of Floor & Ceiling Responses by Group & Question 34
Table 5: Vitamin A Foods Consumption by Question Type & Social Desirability Bias 35

LIST OF FIGURES

Figure 1: UNICEF Program Timeline 30
Figure 2: Sarvodaya Vitamin A Foods Themed Calendar 31
Figure 3 : Vitamin A Food Consumption Prevalence & Social Desirability Bias 35
Figure 4: Single List Compared to Dual List: Validity Check 36
Figure 5: Alternative Direct Measure - No Orange/Yellow Vegetables 37
Figure 6 Vitamin A-Rich Food Consumption Prevalence by Level of Ability to Afford Food and Question Type 38
Figure 7: Impact of Social Desirability Bias on Childhood Dietary Diversity Adequacy and Food Groups Consumed 39

ABSTRACT

Social protection programs in low-income settings increasingly follow a "cash-plus" model, combining cash transfers with a complementary training program designed to promote behavioral change. Evaluations of such programs are especially vulnerable to social desirability bias because the "plus" interventions explicitly reveal to beneficiaries the program providers' preferences while relying on self-reported behavior. We measure social desirability bias in childhood feeding reports in a UNICEF nutrition cash-plus program in Sri Lanka. Using list experiments, we test whether participants misreport young children's consumption of vitamin A-rich foods. We find that direct self-reports overstate adherence to program advice by 23 percentage points. Overreporting is concentrated among households least able to afford food. The mismeasurement of one feeding component passes through and affects aggregate measures of dietary diversity. The magnitude of the findings suggests that social desirability bias could serve as a partial explanation for the persistent gap between recalled dietary intake and anthropometric outcomes in cash-plus program evaluations.

Keywords: Social desirability bias, program evaluation, child feeding, Sri Lanka

ACKNOWLEDGMENTS

This paper builds on a study commissioned by UNICEF and undertaken by IFPRI. The analysis and views presented are those of the authors and do not necessarily represent the views of UNICEF.

We thank Leslie Hodges and participants in the 2025 Kansas Health Economics Conference for thoughtful feedback.

ACRONYMS

BCC – Behavior Change Communication
DHS – Demographic Health Survey
DS – Divisional Secretariat
GN – Grama Niladhar
LNS – Lipids-based Nutritional Supplement
SDB – Social Desirability Bias
UNICEF – United Nations International Children's Emergency Fund

INTRODUCTION

Self-reported data are often subject to bias, particularly when the subject matter is sensitive or morally charged. Social desirability bias occurs when a respondent gives what they deem to be the socially "correct" answer rather than the honest answer to a sensitive question. The potential for social pressure to influence self-reported responses has been widely investigated in the empirical social science literature. Social pressures have been shown to bias measures of land rental rates (Abate et al. 2025), child labor use (Jouvin 2023), intimate partner violence (Cullen 2023), sensitive health behaviors (Lépine, Treibich, Ndour, et al. 2020; Lépine, Treibich, and D'Exelle 2020), and voting behavior (Rosenfeld et al. 2016). The pervasiveness of social desirability bias in the context of program evaluation is comparatively understudied.

Cash plus programs, which combine a cash transfer element and one or more complementary inputs, have become increasingly popular in the last decade (Pruce et al. 2025; Sedlmayr et al. 2020; Webb Girard et al. 2020). Such programs are premised on the idea that "plus" elements enhance the impacts of transfers and enable beneficiaries to undertake necessary investments or changes that are required for longer term sustainability of program impacts.

Our study focuses on a cash-plus program aimed at improving childhood nutrition. Nutrition-sensitive cash interventions commonly include a cash transfer paired with Behavior Change Communication (BCC) efforts aimed at fostering greater knowledge and improved child feeding practices. Evaluations of these programs find generally positive effects on measures based on recalled consumption, but evidence for consistent impacts on anthropometric outcomes is mixed (Alderman 2015; de Groot et al. 2017; Little et al. 2021; Olney et al. 2022). The inconsistency between outcomes raises concerns about the reliability of self-reported data.

One potential explanation for the inconsistency is that social desirability bias influences key self-reported metrics, for example consumption and feeding frequency, for nutrition programs that provide a combination of cash transfers and BCC. Though such transfers are unconditional, BCC components

create a social expectation that the provided cash should be spent in service of the program's objective, such as improving the healthfulness of children's diets. The implied, or explicitly stated, correct use of the cash transfer creates a soft conditionality.[1] While the BCC element of cash-plus programs intends to foment behavioral change through imparting knowledge, skills, and an improved social environment, the soft conditions may also incentivize respondents to augment their responses to evaluation surveys to align with program objectives and expectations. Further, independent evaluation is unlikely to reduce the soft conditionality, as it may be difficult for beneficiaries to understand or trust the separation between a scientific evaluation of a program and the implementation of the program itself.

The possibility that social desirability bias may contaminate evaluations of social programs has been previously considered in other social science disciplines. Prior literature has attempted to utilize methods from psychology developed by (Reynolds 1982) to identify people inherently prone to social desirability bias based on personality (Menon et al. 2016; Rawat et al. 2017). However, we argue that topics can become sensitive and therefore subject to social desirability bias in the context of a cash-plus program evaluation. Within that framework, the circumstance, not the person, drives sensitivity to bias (Blair et al. 2020). In evaluating cash-plus programs, if social desirability bias is perfectly correlated with treatment assignment, researchers face a potentially fundamental identification problem.

Blair, Coppock, and Moor (2020) suggest that social desirability bias[2] occurs only when four elements are present. The first is a social referent, a person that the respondent has in mind when responding to a question. The social referent can be the enumerator or the person or group analyzing the survey responses. The second is a perceived risk that the referent will know the respondent's answer to a sensitive question. Third, the respondent assumes to know the referent's preferential answer to the sensitive question. Fourth, there is a perception that the respondent will face social, monetary, or physical

[1] Conditional cash transfers often rely on administrative data to check adherence, which can mitigate potential social desirability bias concerns.

[2] Blair, Coppock, and Moor (2020) use the term sensitivity bias rather than social desirability bias. Sensitivity bias is an alternative definition of social desirability bias that includes monetary and safety threats. In the broader literature, the two terms are often used interchangeably. For continuity, we use social desirability bias in our paper.

repercussions for failing to give the preferred response.[3] While criteria 1, 2 and 4 are arguably present in all evaluation-based surveys,[4] cash-plus programs present a unique social-desirability bias risk because the "plus" treatment unmasks the preferential survey response. Specifically, programs with an information or BCC component make explicit the "correct" answer from the perspective of the implementing agency, satisfying criterion three.

We use a list experiment to estimate the degree of social desirability bias in a measure of child feeding behavior from an evaluation of a childhood nutrition cash-plus program in Sri Lanka. List experiments, a common form of indirect questioning, disrupt Blair, Coppock, and Moor (2020)'s second criterion (the referent must know the response) by allowing the respondent to answer sensitive questions anonymously. Comparing the indirect prevalence rate of a promoted feeding behavior to the direct measure provides an estimate of the level of social desirability bias.

We find significant levels of social desirability bias in child feeding reports of vitamin-A rich foods in a UNICEF cash-plus program in Sri Lanka. Our analysis finds that traditional dietary recall modules overstate feeding of vitamin A-rich foods to children by 23 percentage points.[5] We investigate the association between ability to afford food and level of social desirability bias. Further, we demonstrate downstream implications of using biased data in the calculation of dietary diversity.

We make several contributions to the burgeoning empirical literature assessing cash-plus programming for social protection, as well as the specific case of nutrition-focused interventions. First, we highlight the unique evaluation risks posed by social desirability bias when evaluations rely on self-reports. Our results suggest the positive bias, which may overstate intervention effectiveness, can be high. Second, our results help interpret recent evidence on the importance of integrating BCC into transfer programs aimed at improving nutritional outcomes for vulnerable children. We identify five recent studies

[3] Independent evaluation may be viewed as a means of reducing perceived risk. Though, as noted previously, it is unlikely for respondents to view the agencies separately.

[4] The referent is the enumerator or a representative from the implementing agency that observes or will observe the responses. When asked directly, the respondent must reveal whether they have followed the program advice or not, meeting the second criterion. Respondents face social repercussions, such as shame or embarrassment, and potentially fear being required to return the money not used for food or reduced future benefits—representing financial risks (criterion four).

[5] The 99% confidence interval of our point estimate ranges from 13.1 to 32.9 percentage points.

that use rigorous, large-scale randomized control trials to assess the marginal impact of providing BCC alongside a cash transfer in a low-income country setting: Ahmed et al. (2025), Field and Maffioli (2025) Hirvonen et al. (2025), Premand and Barry (2022), and Soofi et al. (2022). All find substantial positive impacts on recalled measures of child feeding. However, only two find positive impacts on anthropometric outcomes. We hypothesize a typical "plus" component that includes BCC, coaching or training substantially elevates the risk of social desirability bias. While the empirical analysis is consistent with such a problem, the absence of a control group prevents formal testing. As detailed in the Discussion section we discuss the potential role of social desirability bias in partially reconciling these results, and consider implications for the literature, highlighting opportunities for evaluations that include ex-ante randomized designs.

Background

Sri Lanka experienced a major economic crisis in 2022. In the space of just four months, real wages fell by 25-30% for a wide range of low skill, semi-skill and higher skilled occupations effectively reducing income by between one quarter and one third (World Bank 2022). At the same time, rising inflation increased the price of a sufficient food basket by as much as 66.4%, with the largest impacts concentrated among poor households (World Bank 2022). Moreover, fertilizer bans, input shortages, and fuel shortages disrupted agricultural production, leading to a 25-30% reduction in rice yields (Ghose et al. 2026). These compounding effects led to a nearly doubling of the poverty rate from 13.1% in 2021 to 25.6% by 2022 (World Bank 2022).

Among those most impacted were children. As of October 2022, 43.4% of children under five faced some degree of nutritional challenge, the vast majority of which came from undernutrition. The prevalence of wasting increased from 8.2% to 10.1% nationally from 2021 to 2022 (Ministry of Health 2022). Undernutrition trends hold true for all age groups and sectors (urban, rural, and estate). The annual report prepared by the Sri Lanka Ministry of Health in 2022 recognized a need for intervention and the importance of a healthful diet, especially among young children (Ministry of Health 2022). In 2023,

UNICEF began a childhood nutrition program focused around BCC and cash transfers in vulnerable districts.

UNICEF Program

The UNICEF childhood nutrition program provided five monthly cash transfers to eligible households in districts classified as vulnerable. The selected districts have high rates of severe child wasting. Payments were distributed in two waves. The first wave, which began in March 2023 and ended in July 2023, included Anuradhapura, Kegalle, Kilinochchi, Monaragala, Mullaitivu, and Puttalam. The second wave began in July 2023 and ended in November 2023 and included the Rathnapura, Nuwara Eliya, and Vavuniya districts.[6] Figure 1 provides a detailed timeline of program eligibility, implementation, and evaluation.

Households with at least one child born between May 1, 2021, and December 31, 2022, were eligible for the program. All households with a young child in the selected districts received the payments uniformly, with no other needs-based criteria. For each eligible child, the household was given 6,750 rupees every month for five months via bank transfer. The cash transfer amount was non-trivial, making up over half of the average monthly food expenditure for the beneficiary child.[7]

In addition to the cash transfers, a BCC component was incorporated into the program. The BCC implementing partner was Sarvodaya, a large community development non-governmental agency working in Sri Lanka. Each month households were visited by local Sarvodaya trainers. During the visit, caretakers were given a calendar with a monthly nutrition message (Figure 2) along with other relevant Ministry of Health information. The objective of these visits was to improve dietary diversity, increase awareness of available health and nutrition services, and improve the quality of diets for both the child and the mother.

[6] See Appendix Figure 1 for a map of payment dispersion.
[7] The average per capita monthly food expenditure was 10,158 rupees in 2023 (FAO 2023).

The cash transfer was not conditional on participation in the training or implementation of advice. All eligible households received the cash transfer, regardless of involvement in the BCC. Due to implementation challenges, the transfer of cash and the Sarvodaya BCC elements were not uniformly synchronized. The delays potentially reduced the link between the cash transfer and the programmatic advice, though the synchronization was stronger in the second wave of districts that is the focus of the current study.

In addition to the Sarvodaya BCC, relative to other low- and middle-income countries, Sri Lanka has a robust public health system responsible for disseminating childhood nutrition messages. Information on feeding practices and priming to report specific childhood nutrition outcomes could originate both from the program trainings as well as generally available resources. Given the universal nature of the program, we refrain from drawing strong conclusions about the difference between program specific and general social desirability bias.

The criteria for social desirability bias are likely to be met in the context of the UNICEF cash plus program. The social referent is the enumerator, or more broadly the implementing or evaluating agency representatives that may see the data.[8] When answering questions directly, as is standard survey practice, the response is revealed, meeting criterion two. The BCC made the ideal behavior from the perspective of the implementing agencies explicit, so that criterion three holds. Finally, the repercussions for not following program advice are shame reciprocity failure, and potentially fear of diminished future benefits.

There are three important ways the repercussions from criterion four manifest. Respondents are likely to want to appear as willing and engaged participants in the event the program had (1) potential extensions, (2) unofficial requirements, or (3) cultural expectations. Registration for eligible households was publicized with a pamphlet detailing that the cash transfers would be given unconditionally for four months. However, when implemented a total of five monthly transfers were provided, suggesting that the

[8] For this study, the implementing agencies are UNICEF (funding) and Sarvodaya (BCC) and the evaluating agencies are the University of Peradeniya and the International Food Policy Research Institute.

beneficiaries may have considered the potential for extension salient.[9] For the second point, data was not collected based on respondents' understanding of the unconditional nature of the program. So, it is not clear if beneficiaries were aware that the payments were **not** conditional on behavior changes. Third, prior literature has found higher levels of social desirability bias in collective, cooperative South Asian countries than in more individualistic countries, like the United States (Kim and Kim 2016; Ryan et al. 2021). In total, this suggests a reasonable basis to believe that the criteria for social desirability bias to influence responses were met.

[9] The final evaluation survey reflects that 44% of caregivers reported expecting four payments while only 16% expected five payments. In our sample of 'second wave' beneficiaries, the payments were completed only three months prior to the beginning of the survey.

RESEARCH DESIGN

Data

A survey was deployed to evaluate the UNICEF program with measurement experiments embedded.[10] The follow-up survey was administered in-person to 498 randomly selected households in the second wave of the UNICEF transfers, of which 490 were complete and usable responses. The survey was conducted by district from February 2024 to April 2024 by a team of 18 enumerators. At the time of the survey, eligible children would be between 13 and 36 months old. The sample was drawn from the three districts in the second wave: Rathnapura (198 observations), Nuwara Eliya (181 observations), and Vavuniya (111 observations). Sampling was limited to the second wave of program implementation to reduce recall length.[11] The survey's indirect questions, which are the focus of this study and are explained in detail in the List Experiment subsection, were incorporated in module 10 and the direct questions, discussed in the Direct Question subsection were in module 13. The full evaluation survey included a total of 15 modules.

Caregivers, most commonly the mother, were surveyed as they were the most knowledgeable party. Surveyed households were selected based on two levels of stratification. First, the sample was stratified based on rural, urban, and estate areas. Then, four Divisional Secretariat (DS) divisions were selected from each district. The DS division selection was random for Rathnapura and Nuwara Eliya; however, Vavuniya only has four DS divisions, so all were selected. Finally, within each DS division, four Grama Niladhari (GN) divisions were selected, representing the second level of stratification. Vavuniya and the urban populations were slightly oversampled relative to the population to have an adequate sample size. Household and surveyed caregiver characteristics are summarized in Table 1.

It is important to note that this study is not a randomized control trial. All eligible households received the cash and were offered BCC sessions. Participation in BCC was not a condition of payments,

[10] The broader program evaluation study conducted by Headey et al. (2024) is publicly available.

[11] The broader survey focused on program involvement and use of the funds. Because all beneficiaries were a part of the second wave, the recall window is equivalent for the sample, eliminating concerns about variation in the recency of program interaction.

and so households were not required to attend.[12] Any variability in program access did not occur randomly. The universal nature of the cash and BCC sessions means that we cannot disentangle the impact of program intensity on social desirability bias. We leave this to future work.

Direct Questions

Direct

Direct, self-reported behavioral questions are the standard approach to measuring unobservable characteristics. Typical child feeding modules involve recall over a set period of time. Respondents were asked to report if their child had consumed a list of foods in the past 24 hours in snacks, meals, or mixed with other foods. The food module included 34 food and drink options drawn from the Sri Lanka Demographic and Health Survey (Headey et al. 2024).[13] The child feeding module 24-hour recall window matches the 24-hour recall window used for all items, sensitive and non-sensitive, in the list experiment question, discussed in the next section. Matching the recall window length between the direct and indirect questions ensures comparability of questions across question formats.

Within the child feeding module, the food items relevant to the list experiment are dark green and leafy vegetables or orange fruits and vegetables, which are typically grouped as vitamin A-rich fruits and vegetables. The dark green leafy vegetables and orange vegetables and fruits were asked in three separate, non-consecutive questions embedded in the child feeding module. First, respondents are asked if their child had consumed *pumpkin, carrots, squash, or sweet potatoes that are yellow or orange inside*. Then, two questions later, they were asked if their child had consumed *any dark green, leafy vegetables, such as spinach, okra, lettuce, gotukola, mukunuwanna, Kathurumurunga or other dark green, leafy vegetables*. Finally, two questions later they were asked if their child had consumed *ripe mangoes, ripe papayas, guava, or passion fruit*.

[12] We only have self-reported indicators of participation, which could themselves be subject to social desirability bias. Over 80% of participants reported attending at least one BCC session.

[13] The foods are appropriate for consumption for the beneficiary child at the time of the survey.

The UNICEF program's nutritional advice, provided in the form of a calendar (Figure 2), recommended parents feed their children leafy green vegetables or orange vegetables or fruits every day, which again matches the 24-hour recall window. Respondents were considered to have followed the programmatic advice by the direct measure, if they answered yes to one or more of the three questions.

List Experiments

List experiments, also referred to the item count technique, are commonly used to measure sensitive topics that respondents may be averse to answering honestly. By design, the response to the sensitive item in the list experiment is anonymous even to the enumerator, therefore alleviating the external pressure on respondents to provide the socially perceived "correct" response. An indirect response suggests the second criterion is unlikely to hold, substantially reducing the risk of social desirability bias.

In a standard single list experiment, individuals are randomly assigned to two groups, which we refer to as Group A and Group B.[14] Respondents in Group B are read a list of items and asked to report how many are true for them or their household. Individuals in Group A are read the same list of items, with the addition of the sensitive item of interest, and are similarly asked to report how many are true. Importantly, respondents never reveal which statements are true, instead only reporting in aggregate how many are true. The difference in mean number of affirmative responses between Group A and Group B is the estimated aggregate prevalence of the sensitive item. Therefore, we can uncover the prevalence of a sensitive item indirectly, which ensures the privacy of each individual respondent.

While list experiments mitigate the impact of social desirability bias, it comes with a loss in efficiency. List experiments add random noise to the estimation of the prevalence of an item, especially when compared to the low-variance direct measure. Efficiency can be improved by using a dual list experiment, rather than a single list experiment (Droitcour et al. 1991). In the dual list experiment,

[14] Other studies typically refer to these groups as the treatment group and control group. We have elected to use more general language so as to not confuse the measurement group assignment with program assignment.

respondents are asked two list questions, one with the sensitive item and one without. The order in which the respondents see the list with the sensitive item is different for Group A and Group B. The non-sensitive items differ between list questions one and two; however, the sensitive item remains the same. The dual list uncovers the indirect response to the sensitive item for all respondents, rather than half, greatly improving efficiency. The double list requires two questions to estimate the prevalence of one metric, which has the potential to increase fatigue and cognitive burden to respondents (Glynn 2013). Given our relatively limited sample size, the improved efficiency from a dual list is necessary.

Our primary objective is to detect social desirability bias in key programmatic evaluation measures, and so the selection of the sensitive item in the list experiment is key. As noted in the UNICEF Program Section, in addition to the unconditional cash transfer, caregivers received nutritional training with a monthly theme. While all of the themes were informational, only one monthly theme had time-bounded, quantifiable instructions. Parents were given advice to feed their child a dark green leafy or orange vegetable or fruit every day, with a corresponding list of health benefits of doing so. The reference to "daily" is a clear boundary that matched the recall window for the direct questions. "The beneficiary child in my household ate dark green leafy vegetables or orange colored vegetables or fruit in the past 24 hours" was therefore selected as the key item.

Clear instructions are an important tool to reduce the cognitive burden of answering the complex list experiment questions. To account for differing levels of literacy and numeracy, techniques like passing beans between hands (Tadesse et al. 2020; Abate et al. 2025) or counting fingers (Jouvin 2023) have been used. We opted to use the counting fingers strategy as it did not require additional materials. Respondents were instructed to keep track of affirmative statements on their fingers behind their back, out of sight of the enumerator. Doing so lowers the cognitive effort required for the question while still keeping the response to each item anonymous. The full list experiment instructions and questions are detailed in Table 2.

ESTIMATION STRATEGY

Following notation from Tsai (2019), let T_i be the group indicator. If respondent *i* is in Group A then $T_i=1$ and for a respondent in Group B, $T_i=0$. Respondent *i* has potential answer S_i to the sensitive item and $\mathrm{R}_{\mathrm{i,j}}^{1}\left(R_{i,k}^{2}\right)$ to non-sensitive item *j* (*k*) with a total of *J* (*K*) non-sensitive items in list question one (list question two). In the context of our study, $S_i=1$ indicates the respondent *i*'s young beneficiary child was fed a green leafy or orange vegetable or orange fruit in the past 24 hours. For list question one, a response of $R_{i,1}^{1}=1$ would indicate a person in respondent *i*'s household had consumed rice in the past 24 hours. By design, the responses to S_i, $R_{i,j}^{1}\ \forall\ j$, and $R_{i,k}^{2}\ \forall\ k$ are not observable. Instead, we observe $Y_i^1=T_iS_i+\sum_{j=1}^{J}R_{i,j}$, where Y_i^1 is the number of affirmative statements in list question one for respondent *i.* Similarly, we observe $\mathrm{Y_i^2}=(1-\mathrm{T_i})\mathrm{S_i}+\sum_{\mathrm{k=1}}^{\mathrm{K}}\mathrm{R_{i,k}}$, where Y_i^2 is the number of affirmative statements in list question two for respondent *i*.

To estimate the prevalence of the sensitive item $\left(\mathrm{P}(\mathrm{S_i}=1)\right)$, we use a difference-in-means estimator. Using a double list experiment, we estimate the prevalence of the sensitive item separately for list question one and list question two then take the arithmetic mean.

$$P(S_i=1)=\left[\left(\frac{\sum_{i=1}^{n}Y_i^1T_i}{\sum_{i=1}^{n}T_i}-\frac{\sum_{i=1}^{n}Y_i^1(1-T_i)}{\sum_{i=1}^{n}(1-T_i)}\right)+\left(\frac{\sum_{i=1}^{n}Y_i^2(1-T_i)}{\sum_{i=1}^{n}(1-T_i)}-\frac{\sum_{i=1}^{n}Y_i^2T_i}{\sum_{i=1}^{n}T_i}\right)\right]/2 \qquad (1)$$

The estimator in Equation 1 was implemented using the Tsai (2019) dual list procedure, which is estimated by generalized method of moments and reports the estimated prevalence and standard error.

We extend our analysis to include a multivariate analysis to explore the relationship between social desirability bias and ability to afford food. The difference-in-means approach cannot accommodate multivariate analysis without further splitting the sample, so an alternative approach must be used.

Therefore, we implement the Imai (2011) proposed nonlinear least-squares estimator[15] that has been adapted by Tsai (2019) for the dual list application. A simultaneous equation model is used as follows:

$$Y_i^A = J * logit^{-1}(x'\gamma^A) + T_i * logit^{-1}(x'\delta) + \epsilon^A$$

$$Y_i^B = K * logit^{-1}(x'\gamma^B) + (1 - T_i) * logit^{-1}(x'\delta) + \epsilon^B \quad (2)$$

Where x is a vector of covariates and other variables are as defined above. The parameters γ^A and γ^B index the expected number of affirmative responses to non-sensitive items lists A and B, while δ is the common sensitive item parameter. Estimation follows the three-step procedure outline by Tsai (2019), where more details can be found in the appendix.

Our covariate of interest is a measure of a household's ability to cover basic food and nutritional needs (F_i). The question had five potential responses ranging from very poor to very good.[16] In addition, we include an asset index, Z_i, with a mean of zero to account for household wealth, as the program was not targeted. The asset index was created using responses from the asset module and the *swindex* package in Stata (Schwab et al. 2021).

$$P(S_i = 1|F_i, Z_i) = logit^{-1}\left(\hat{\delta_f}F_i + \hat{\delta_z}Z_i\right) \quad (3)$$

which is evaluated at the observed range of the ability to afford food.

List Experiment Assumptions

The validity of list experiments hinges on three key assumptions: randomization, no design effects, and no liars. When these assumptions hold, the difference in mean number of affirmative responses between Group A and Group B for each of the lists captures only the prevalence of the sensitive item.

The randomization assumption requires that Group A and Group B be similar so that the average response to the non-sensitive items is balanced. In Table 1, we see that for nearly all observable characteristics Group A and Group B are statistically similar. The electricity bill is the only characteristic

[15] A linear probability model would allow for a multivariate analysis; however, estimated probabilities may be outside the allowable zero to one interval.

[16] The question appeared in module six, before both the list experiment and direct food consumption questions. The same phrasing was used to ask about five other basic need areas including health care, education, clothing, accommodation, and utilities.

where the groups are statistically different. The difference occurs at the 10% significance level and other measures of household wealth do not vary by group. Therefore, the randomization assumption appears to hold.

For the no design effects assumption to hold, the presence of the sensitive item cannot influence the response to the non-sensitive items. The likelihood of design effects can be proactively reduced by selecting non-sensitive items that are similar in nature to the sensitive item (Blair and Imai 2012). When all of the items in the list are similar, the sensitive item is less likely to stand out and is therefore unlikely to influence the response to other items. In our list experiments, all of the items were in some way related to food, from market to consumption.

Post-survey implementation, we use the Blair and Imai (2012) test for design effects. The results of the formal design effects tests are in Table 3. The test is only suited for a single list experiment, so we have conducted the tests on list question one and two separately. Table 3 Panel A shows the tests of the joint probability of each response. Negative coefficients indicate that the design effects assumption is violated. In practice, adding one sensitive item to a set list of non-sensitive items should not decrease the probability of that response occurring. Table 3 Panel B includes results an additional test to determine if any negative coefficients occurred randomly. There are two tests of stochastic dominance relationships conducted separately with the results combined using the Bonferroni correction. The Bonferroni-adjusted p-values near 1 indicate that we fail to reject the null hypothesis of no design effects. In total, we cannot reject the hypothesis that the no design effects assumption holds.

The final assumption is that respondents do not purposefully give a false answer to the list question, called the no liars assumption. The primary reason that someone would strategically give a non-truthful answer to the list experiment is if, in doing so, they would reveal their answer to the sensitive item. In that case, the list question is functionally similar to the direct question. Respondents would inadvertently reveal the answer to the sensitive item if they gave an answer of zero or four when they had the list that includes the sensitive item. These are commonly referred to as floor and ceiling effects, respectively.

To avoid floor and ceiling effects, we include a non-sensitive item with a low prevalence and an item with a high prevalence (Glynn 2013). For list question one, we included rice consumption to avoid floor effects, as rice is widely consumed. Similarly, purchasing livestock in the past 24 hours would be rare and reduces the chance of ceiling effects. In list question two, consuming tea is a high-prevalence item and preparing food with kerosene is a low-prevalence item. In Table 4 we see zero and three responses among the groups with three item lists are rare. Answers outside of the allowable range were also uncommon. The prospectively designed questions effectively reduce floor and ceiling effects, making it likely that the no liars assumption holds.

The necessary assumptions for a valid list experiment appear to hold. Therefore, the dual list difference-in-means estimation strategy in Eq. 1 is valid. Our indirect estimate gives the share of households where the beneficiary child was fed a dark leafy green or orange vegetable or fruit in the past 24 hours.

Social Desirability Bias

We define social desirability bias as the difference between the direct and indirect measures of the prevalence of vitamin A-rich fruit and vegetable consumption in the past 24 hours.[17] Assuming the anonymous nature of the indirect question relieves social pressures, we expect $\mathrm{P}_{\mathrm{direct}} > \mathrm{P}_{\mathrm{indirect}}$. This gives:

$$\text{Social Desirability Bias} = P_{direct} - P_{indirect} \tag{4}$$

Identifying social desirability bias requires that the only difference between the two questions used to estimate prevalence is the level of social influence. Put another way, both methods would uncover the same prevalence rate in the absence of social pressure. There are three potential threats to our identification strategy, though we argue they are not significant concerns. They are (1) the difference in

[17] We obtain standard errors and confidence intervals by nonparametric bootstrap, resampling households with replacement across 1,000 replications.

the cognitive difficulty of the questions, (2) asking all respondents both questions, and (3) the difference in the wording of the indirect and direct questioning.

The two questions may require different levels of mental effort leading to more non-strategic mismeasurement in the indirect metric than the direct metric. We would expect non-strategic measurement error to manifest in responses outside the allowable range. However, Table 4 shows that we have a relatively low number of responses out of range. The in-person data collection offered an opportunity to clarify any uncertainty in the list questions, reducing the likelihood of non-strategic mismeasurement. Ahlquist (2018) found that the difference in means method is more robust to non-strategic mismeasurement than the maximum likelihood estimator proposed by (Imai 2011). So, to account for the possibility of non-strategic error, we use the difference in means method for our base analysis.

The second concern is that the response to the first question (list question) about the sensitive topic will influence the response to the second question (direct question) on the same topic. While possible, the list experiment questions were unlikely to influence the responses to the direct questions for the following reasons. First, the questions were in different modules, with a nutritional knowledge module separating them, and therefore would not have been asked consecutively. Second, the direct questions were embedded within a child feeding module, and so the questions related to the dark green leafy vegetables or orange-colored vegetables or fruits would not have stood out in relation to the other food items. In fact, even the items that make up the direct measure were asked non-consecutively.

Finally, while the recall window length and the unit of measurement match for both the indirect and direct questions, there are slight differences in how foods are listed. In the direct question, respondents are provided with a listing of potential foods (e.g. spinach), while the indirect question only ask for any foods that fall in the category of dark green leafy vegetables or orange-colored vegetables or fruits food groups. This difference could influence the social desirability bias measure in two different ways. Listing the potential food items, rather than giving generic categories, may artificially reduce the number of affirmative answers to the direct question because a caregiver may not include other items

outside of the listed food items. If this is the case, then the direct measurement would be an underestimate, and the level of social desirability bias would be understated. In this case, the results of this paper are a lower bound. On the other hand, the list of items may be more helpful in recalling the foods given to their child. If true, the indirect measure may be an understatement so that the social desirability bias would also be overstated. However, given the recall window of 24 hours, recall should not be a significant concern, making overstatement of social desirability bias less likely.

In addition, the first direct question includes mention of *sweet potatoes that are yellow or orange inside*. Our indirect measure only captures sweet potatoes that are orange inside. Thus, the direct measure includes one additional food item (yellow sweet potato). The subtle difference is derived from variation in the wording of the traditional child feeding module and the Sarvodaya calendar. The child feeding module was designed to mirror questions asked by Sri Lanka Demographic and Health Survey, while the indirect question sensitive item used the verbatim phrase of the calendar (Headey et al. 2024). Thus, differences in prevalence derived from direct and indirect questioning could potentially be partially driven by yellow sweet potato consumption. We posit that yellow sweet potatoes are unlikely to be the primary food item consumed from the listed food items as they are generally more costly than other sweet potatoes (including the more commonly consumed white sweet potato). Still, we conduct a robustness check (Robustness Check Section) to demonstrate that our findings are unlikely to be driven by this difference.

RESULTS

The results of the list experiment, shown in Table 5 and Figure 3, reveal statistically and nutritionally significant levels of overstatement of feeding of vitamin-A rich foods. When asked directly, 86% of caretakers report that their youngest child ate at least one vitamin A-rich fruit or vegetable in the past 24 hours (Table 5). When asked indirectly, and therefore anonymously, only an estimated 63% of caretakers fed their youngest child a vitamin A-rich fruit or vegetable in the past 24 hours (Table 5). The difference between the direct and indirect measures is significant at the 1% level.

These findings are consistent with social desirability bias driving nutritionally meaningful overstatement of vitamin A-rich fruit and vegetable consumption. While we do not observe the true prevalence of vitamin A consumption in young children, we argue social pressures are more likely to lead to the overstatement of such behavior rather than an understatement. Feeding your child a well-balanced diet, especially vitamin A-rich foods as instructed by the Sarvodaya BCC, is perceived to be the socially "correct" response. When asked directly, caregivers that did not implement this advice may feel pressure to give the "optimal" response. Therefore, the anonymity provided by the indirect question should uncover an estimated prevalence closer to the unobserved true prevalence of vitamin A-rich fruit and vegetable consumption in the recall window.

Overall, our findings suggest that direct measures overstate the proportion of children consuming Vitamin A-rich foods in the previous 24 hours by 23 percentage points, or 26.5%. The observed levels of social desirability bias are consistent with beneficiaries inflating compliance with program advice, though we cannot directly distinguish between program-specific and more general sources of social desirability bias in this study.

Robustness Checks

In the List Experiment Assumptions Section we provided evidence that the necessary list experiment assumptions held. When using a dual list experiment, we can conduct an additional check. Chuang et al. (2021) suggest comparing the prevalence of the key item for each list question used in the dual list

estimation independently. If the assumptions hold, we would expect statistically equivalent estimates of prevalence for each of the lists estimated separately. In Figure 4 we can see that the estimations of the prevalence of the key item are statistically equivalent. Using either list question one or list question two independently, we would still find positive levels of social desirability bias. Figure 4 demonstrates the efficiency gains from using a dual list experiment and also gives further confidence about the validity of the list experiments.

As noted in the Social Desirability Bias Section, the direct measure could potentially capture yellow sweet potatoes that would not be included in the indirect measure. To influence results, yellow sweet potato would need to be the only food item consumed in our direct measures. This scenario is seemingly unlikely and if present would make up a small portion of the social desirability bias. To demonstrate the upper bound of the influence of yellow sweet potatoes, we replicate our results excluding the entire orange vegetable (plus yellow sweet potato) category from the calculation of the direct measure of vitamin A consumption. Assuming that all orange-fleshed vegetable consumption consisted solely of yellow potatoes, the estimate of social desirability bias decreases to 14 percentage points (Figure 5) and remains statistically different from zero. The qualitative conclusions are thus robust to a deliberate under-count of the direct measure.[18]

Ability to Meet Basic Food Needs

We have presented evidence that beneficiaries overstate adherence to nutritional advice. We next examine how differences in the ability of respondents to meet their basic food needs varied with social desirability bias in responses to feeding questions. To do so, we estimate the relationship between both the indirect and direct measures of vitamin-A feeding and a five-point Likert scale where respondents rate their ability

[18] Alternatively, one could bound the potential overestimate using direct measures of the rate of consumption of yellow-fleshed sweet potato as a proportion of carrot, pumpkin, squash and orange-fleshed sweet potato. However, no reliable data on either the consumption of or supply of Sri Lankan sweet potatoes at the variety level exists.

to “cover basic needs” for food. Our baseline estimate relies on the nonlinear logistic specification in Tsai (2019) and controls for a standardized asset bundle.

The analysis, summarized in Figure 6 reveals two important findings. First, the gap between the direct elicitation and indirect measurement is driven by those who report lower ability to meet their food needs. Among the 15 percent of the sample claiming their ability to meet their basic food needs was good or very good, we find no statistical evidence of social desirability bias. For those with ratings ranging from very poor to adequate, the gap between the indirect and direct measures is substantial.

Second, the direct measure strongly understates heterogeneity in compliance with vitamin A feeding. The relationship of the direct measure to the household's ability to afford food is only weakly positive and not statistically distinguishable from zero. The evidence from the list experiment suggests that social desirability bias may mask inequality in nutritional outcomes. Indeed, the indirect results are consistent with the expectation that households who struggle most with basic needs would be least likely to expend resources to comply with micronutrient consumption messaging.

Appendix Table 2 shows that the basic relationship between ability to meet food needs and vitamin A feeding is stable in a variety of model specifications. Nevertheless, the analysis here has several caveats. The list experiment was not stratified by ability to meet food needs, nor did we pre-specify the heterogeneity analysis. While we control for household wealth with an asset index, we cannot fully disentangle the food needs metric from other potential related characteristics. Therefore, the analysis of the association between basic food needs and desirability bias should be treated as exploratory.

Dietary Diversity

In addition to measuring the prevalence of the consumption of vitamin A-rich foods, consumption data may also be used to calculate other metrics of interest. Dietary diversity is commonly measured in childhood nutrition programs. Using the direct measures and an eight-food group dietary diversity

measure[19], 85% of children have adequate dietary diversity.[20] The average number of food groups consumed is 6.05.

The share of children with adequate dietary diversity in our data is high relative to Sri Lanka's national estimate from the 2016 Demographic and Health Survey (DHS). In that sample, the average share of households with adequate dietary diversity across all wealth levels was 71.6%, and only the highest wealth quintile had 85% with adequate dietary diversity (Department of Census and Statistics 2016).[21] While the program was not targeted to exclusively low-income households, it was implemented in three of the poorest districts in Sri Lanka in the year following a major financial crisis. A possible explanation is that the high measured dietary diversity in our data could also be a function of social desirability bias in the consumption data that are used to construct the dietary diversity metric.

To explore how exaggerated self-reports in consumption data can impact aggregate metrics like dietary diversity, we use simulation calibrated to our estimate of the share of caregivers overstating vitamin A-rich food consumption (i.e. 26.5%). We begin under the conservative assumption that social desirability bias affected the consumption of Vitamin-A foods only. Specifically, for each simulation we randomly assign a level of social desirability bias (B) from a normal distribution with a mean 0.265 and a standard deviation of 0.038, drawn from the standard error of our main analysis (Table 5). Then, a randomly selected B% of vitamin A food *consuming* children had their response to the vitamin A food group changed from one to zero. This process was repeated 1000 times with the mean and confidence interval reported in Figure 7. Because only households that reported consuming were eligible to be changed, less than B% of the total sample was changed.[22] As such, our vitamin A foods simulation represents a lower bound of the impact of social desirability bias on dietary diversity measures

[19] The eight food groups are: (1) breastmilk, (2) grains, root, and tubers, (3) legumes and nuts, (4) dairy, (5) flesh foods, (6) eggs, (7) vitamin A rich fruits and vegetables, and (8) other fruits and vegetables. Seven food groups, which excludes breastmilk, is also commonly used. However, 76% of our sample reported that they were currently breastfeeding (Appendix Table 1), so we selected the eight group definition.

[20] Adequate dietary diversity is defined as consuming greater than four food groups.

[21] This sample considered all children under five, which is a wider age range.

[22] This is the monotonicity assumption used in prior list experiment literature, which requires that misreporting is in one direction only (Aronow et al. 2015; Eady 2017).

When only the vitamin A food group is adjusted for social desirability bias, the share of children with adequate dietary diversity dips to 81%. The average number of food groups consumed also decreased to 5.82 food groups. The simulated measures of dietary diversity are lower than the direct counterparts, though the confidence intervals are overlapping.

We selected vitamin A-rich fruits and vegetables as our key item because of the bounded and measurable advice given in the monthly calendar. However, recipients received other general nutrition and feeding advice. We therefore explore how a dietary diversity aggregate may respond to social desirability bias of a similar magnitude in other food groups.

Commonly consumed staples are unlikely to be influenced by social desirability bias, so no adjustments were made to these food groups. To account for other social desirability bias, we do another simulation, this time allowing non-staple foods to have some level of social desirability bias between zero and 26.5%, drawn from a uniform distribution. The selected range and uniform distribution were used to avoid making any assumptions about the social desirability bias of foods we have no indirect measure for.[23] The non-staple foods we allowed to vary are dairy, flesh foods, eggs, and other fruits and vegetables. Vitamin A-rich fruits and vegetables followed the same distributional assumptions and selection process made in the prior scenario. We assumed that the social desirability bias for one food group was independent of another for the same individual. In practice, this is unlikely to hold, and so our simulated results represent a conservative approach to simulating desirability bias in measures of dietary adequacy.[24]

When we deflate reports of non-staple foods consumption based on the simulated level of social desirability bias, the share of children estimated to have adequate dietary diversity declines to 74%, which is a 11 percentage point decrease from the direct measure. The number of food groups consumed is 5.45,

[23] We are implicitly assuming that Vitamin A foods have the highest potential for social desirability bias. If social desirability bias is higher among other food groups, then our results would be an underestimate.

[24] If social desirability bias between food groups is highly correlated, then social desirability bias will be concentrated among a subset of the sample. As a result, individuals would be much less likely to meet the minimum diversity threshold, which would decrease the share with adequate dietary diversity. When social desirability bias is randomly spread across individuals through random selection, then more households will surpass the minimum threshold for dietary diversity. As such, the presented results represent a lower bound of the influence of social desirability bias on a threshold-based metric.

which is a decline of 0.6 from the directly measured number. Thus, social desirability bias in consumption measurement at the levels estimated in our list experiment can have meaningful downstream effects on dietary diversity aggregates.

DISCUSSION

The connection between improved diet quality and health has been well established (Arimond and Ruel 2004). However, while social protection programs have consistently improved dietary outcomes, the linkage with physical health outcomes is much weaker. de Groot et al. (2017) and Alderman (2015) suggest several potential reasons for the inconsistent results, which include supply-side market failures, poor targeting, length of program, and age of children.

We propose mismeasurement in the form of social desirability bias as an additional potential explanation. Consumption measures are self-reported, while anthropometric measures are collected through physical measurement. We hypothesize that positive results based on self-reported measures could be partially due to overstatement by beneficiaries who adjust behavioral reports to appear in line with program objectives. In doing so, respondents exaggerate behavior change in a manner that distorts actual dietary consumption and generates 'surprising' null impacts on physical outcomes.

To what extent does social desirability bias impact the reliability of evaluations of cash-plus and explicitly nutrition-sensitive cash interventions? Higher levels of bias should inflate the effects on self-reported consumption relative to anthropometric and biomarker data. However, because the relationship between dietary diversity measures and anthropometric measures is not well-calibrated, such upward bias is difficult to discern when BCC elements are not prospectively isolated from other intervention benefits. A cleaner empirical prediction can be derived from evaluations in which the impact of training is credibly isolated from the material (e.g. cash) components of the intervention. In that case, we would expect the marginal addition of training to inflate recalled outcomes more than anthropometric ones. We note several elements of the literature consistent with this prediction.

As noted by Gentilini 2022, few published evaluations of childhood nutrition programs include both a prospective design and a mix of recalled and anthropometric outcomes. We limit our discussion to studies within the last five years of unconditional cash transfer focusing on childhood nutrition that

include both self-reported and anthropometric outcomes and treatment arms that contain a separable BCC or "plus" component.[25]

From these criteria, we identify five relevant studies. In all these studies, we identify several results consistent with the presence of potentially meaningful levels of social desirability bias.[26] Specifically, each study contains positive self-reported nutritional outcomes that are not reflected, either in whole or in part, in the anthropometric outcomes. We stress that the pattern we identify is not conclusive evidence of social desirability bias---for example, dietary impacts may simply have been too small to generate physical outcomes---but indicates greater attention is warranted to reporting issues.

Hirvonen et al. (2025), Premand and Barry (2022), and Soofi et al. (2022) all find positive and statistically significant effects on nutritional behavior change yet find no impact on anthropometrics. Soofi et al. (2022) find statistically indistinguishable impacts on anthropometry despite the BCC group self-reporting significantly higher compliance with a lipid-based nutritional supplement (LNS) feeding when compared to the group with the same material components but no BCC. The program targeted poor households with children six months of age in Pakistan for 18 months; providing different combinations of cash (32 USD), lipid-based nutrition supplement, and behavior change communication. Similarly, Premand and Barry (2022) find that the addition of BCC to cash in Niger substantially improves recalled dietary outcomes (0.24 standard deviation increase), but not child anthropometry. Cash alone had no measured effect on recalled dietary diversity nor anthropometry for children. The intervention targeted poor households, provided monthly transfers equating to on average 15% of monthly food expenditure for two years along with nutrition and health focused parental training for 18 months. Finally, Hirvonen et al. (2025) do not find any statistically meaningful impact on anthropometric outcomes in the full sample,

[25] The additional administrative monitoring in conditional cash transfers may alter the scope and incentives for reporting error in ways that differ from the focus of this study. For example, the conditional nature of the cash transfer in (Levere et al. 2024) limits comparability to the other studies referenced. In addition, the design does not isolate the marginal impact of trainings.

[26] The desire to focus on the marginal impact of adding training or 'plus' elements excludes several related studies showing positive overall impacts of cash-plus programs. For example, the designs of Carneiro et al. (2021) and Bouguen and Dillon (2026) do not permit us to separate the impacts of training from other program elements. Similarly, Leight et al. (2026) implement a complex, multi-arm trial with varying levels of material and BCC support in Ethiopia, but we do not include the study in our discussion here because the control group had access to some levels of both supports.

despite finding statistically significant impacts on dietary diversity among the treatment arm receiving BCC and cash transfers. This Ethiopian study is layered on top of the Productive Safety Net Program, providing community based training and an additional payment equating to 13% of the monthly food budget for 24 months during the infancy.

The findings of Ahmed et al. (2025) and Field and Maffioli (2025) indicate that social desirability bias cannot be the sole mechanism explaining positive impacts of integrating BCC with transfers. We note, however, the pattern of results from these studies is also consistent with some degree of biased reporting of recalled diversity measures.

In Field and Maffioli (2025) the impact of Cash+BCC on total food expenditure doubled the impact of cash alone in a study of newly pregnant women in Myanmar.[27] Yet despite similarly large marginal impacts of BCC on key dietary intake and diversity measures, the provision of BCC had no impact on children's weights. The absence of weight impacts can plausibly be attributed to the time lag between the end of transfer receipt and measurements. However, that explanation is difficult to reconcile with the persistence of large food intake and expenditure impacts for the Cash+BCC arm, but could in part be consistent with social desirability bias inflating recalled dietary measures.

In Ahmed et al. (2025) both the Food+BCC and Cash+BCC arms increased reported dietary diversity by similar magnitudes (the null hypothesis that the two groups are equal cannot be rejected), yet only the Cash+BCC arm yielded positive anthropometric effects (decreased stunting and increased weight for age). The Bangladesh study lasted 24 months, targeted children under 24 months old, and provided substantial material support for both arms (25% of monthly consumption). Why did the Food+BCC arm not likewise demonstrate anthropometric impact despite large reported consumption gains? The Ahmed et al. (2025) study takes the potential issue of desirability bias seriously, but ultimately attributes the differences to higher animal-source food consumption in the cash group, primarily egg and choline intake rather than social desirability bias.

[27] The value of the cash transfer equated to three or four days of work at the minimum wage.

The reviewed studies all included the ability to isolate BCC from other benefits, unconditional transfers, intermediate length (18-30 months), and cash transfers (10-35% of monthly food budget) that were targeted at early development (gestation through 24 months). While an effort was made to select comparable and rigorous studies, there are still many factors like sociodemographic and geographic context that underpin each study. Therefore, we make no attempt to infer the magnitude of potential social desirability bias or suggest any program characteristics that may be associated with social desirability bias. Instead, this discussion serves to highlight how mismeasurement could account for part of the gap between self-reported and anthropometric impacts, in addition to other factors already established in the literature. As such, future studies should consider the risk of social desirability bias in self-reported metrics, particularly those related to program-affiliated BCC.

CONCLUSIONS

In this study we estimate the prevalence of social desirability bias in self-reported measures of vitamin A-rich fruit and vegetable consumption among UNICEF program beneficiaries who received a cash transfer and childhood nutrition information. The nutrition BCC specifically advocated for the daily consumption of vitamin A-rich fruits and vegetables, and provided circumstances that met the necessary conditions for social desirability bias. We estimate that direct reports overstate children's daily consumption of vitamin A-rich fruit and vegetable by 23 percentage points. The overstatement indicates non-trivial levels of social desirability bias.

We also find suggestive evidence that social desirability bias is negatively associated with the ability to afford food. By the indirect measure, the prevalence of vitamin A-rich food consumption rises as ability to afford food increases, while the direct measure shows no detectable variation. One possible explanation is that for households with lower ability to afford food, the repercussions of not giving the socially correct answer may be more salient. The direct measure's larger bias at low levels of food affordability suggests self-reports conceal heterogeneity in program effectiveness.

The social desirability bias in consumption passes through, leading to mismeasurement in dietary diversity. Our simulated results indicate that social desirability bias could lead to overstating dietary diversity, a common evaluation metric for nutrition programs aimed at increasing the nutritional status of children. Mismeasurement has the potential to lead to skewed interpretations of program effectiveness.

Our study contributes to existing measurement error literature by suggesting social desirability bias as an additional reason for the regularly observed gap between self-reported and anthropometric measures in nutrition program evaluation. While our study results are limited to nutritional outcomes in the UNICEF program in Sri Lanka, the conditions necessary for social desirability bias to influence data quality are present in a wide variety of programs outside of the realm of nutrition. We leave it to future research to investigate the presence of social desirability bias in other programmatic interventions across topics.

The evidence and discussion of the literature illustrates that social desirability bias plausibly plays a meaningful role in inflating some self-reported measures of food intake in BCC interventions. Similar problems may plague the "cash-plus" literature more generally. How then should researchers and program designers proceed?

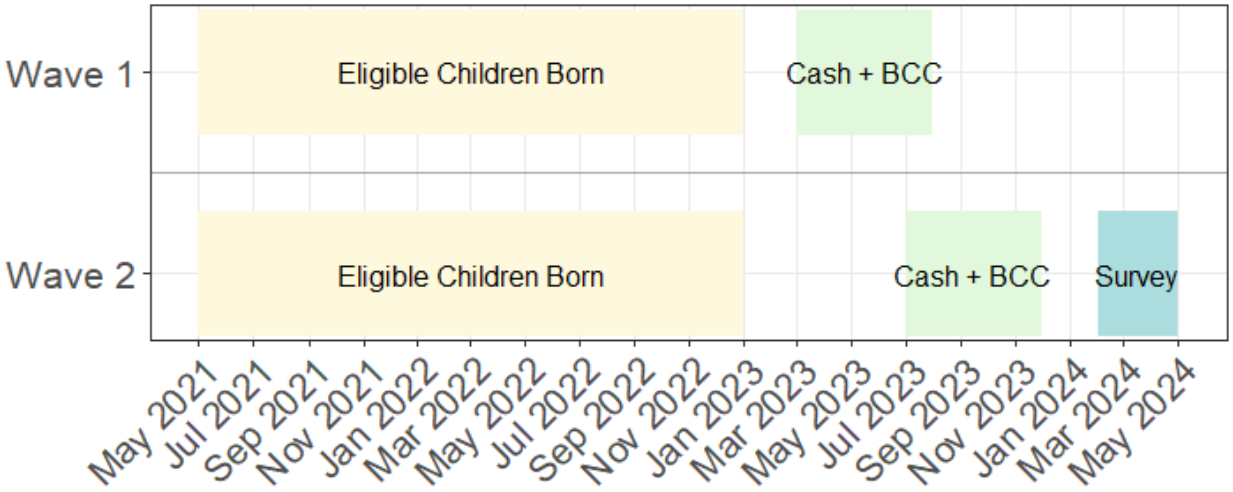
One potential route is to avoid the threat of social desirability bias by focusing on outcomes that do not rely on beneficiary self-reports. For nutrition-based projects, these can include anthropometric outcomes and biomarker data. In some cases, administrative data sourced from program participation records or digital payment transaction records may be viable.

For outcomes that require recall-based reports from beneficiaries, evaluation surveys can incorporate indirect measurement approaches more routinely. However, these measures can be complicated to implement and are inherently less precise. That suggests an opportunity for research to better understand the nature of the social desirability bias problem and test potential solutions.

Future research should attempt to differentiate generally present social desirability bias from program specific social desirability bias. In the case where a program may amplify existing normative behavior, disentangling the source of social desirability bias (program or broader society) is useful. The natural next step for future research is to build tests for social desirability bias into randomized control trials that incorporate BCC into treatment arms. A study with multiple treatment arms with differing levels of BCC intervention could begin to disentangle general and specific social desirability bias.

The more granular understanding of the contours of the problem may uncover mechanisms to reduce social desirability from direct questions. Future work should also investigate ways to reduce social desirability bias with a lighter touch so that more efficient estimation strategies can be employed. Assessing whether potential low-cost mitigation strategies, such as cheap talk or taking repeated alternative measures, may provide high returns. In addition, research could seek to better understand how survey practices that reduce non-strategic error (e.g. higher frequency surveys) interact with social desirability bias.

Figure 1: UNICEF Program Timeline



Source: Figure created by authors using information from Headey et al. (2024).

Figure 2: Sarvodaya Vitamin A Foods Themed Calendar



Note: English translation: Dark green leafy vegetables and dark yellow/orange vegetables and fruits help maintain healthy eyesight and protect against disease. So add these to your child's diet daily.
Source: Headey et al. (2024).

Table 1: Summary Statistics & Randomization Assumption Balance Table

	Total	Group A	Group B	p-value
Household Characteristics	Mean	Mean		
Respondent Age	31.37	31.46	31.28	0.72
Household Size	4.63	4.54	4.73	0.13
Children 4 and under	1.21	1.20	1.23	0.53
Language[a]	Count	Count		
Sinhala	303	152	151	0.34
Tamil	187	102	85	0.34
Education[a]	Count	Count		
Primary or less	7	2	5	0.21
Secondary Grade 6-9	58	28	30	0.56
Secondary O/L level	232	126	106	0.30
Secondary A/L level	136	68	68	0.62
Vocational stream	13	8	5	0.48
Higher education	44	22	22	0.80
Household Wealth	Mean	Mean		
Has Debt	0.54	0.53	0.56	0.54
Electricity Bill (LRK)	2,984	3,233	2,720	0.08
Owns Refrigerator	0.53	0.51	0.55	0.44
Owns Motorcycle or Car	0.48	0.47	0.49	0.55
Owns Land	0.19	0.18	0.19	0.79
Owns Livestock	0.13	0.15	0.12	0.45
Food Consumption Behavior	Mean	Mean		
Produces own food	0.76	0.74	0.78	0.42
Child ate orange vegetables in last 24 hours	0.65	0.63	0.68	0.27
Child ate orange fruits in last 24 hours	0.50	0.50	0.49	0.85
Child ate green leafy vegetables in last 24 hours	0.68	0.67	0.68	0.78
Child ate Vit A rich food in last 24 hours	0.86	0.87	0.86	0.85
N	490	254	236	

Note: A t-test of the differences of the means of Group A and Group B yields the p-value. For variables denoted with [a], the p-values are for tests of no difference between sample shares in Group A and Group B.
Source: Authors' elaborations from primary data.

Table 2: List Experiment Instructions and Questions

List Instructions:

I am going to read a few statements about your household. Do not reply yes or no after I have read each statement. Your answers must remain confidential. Please put one hand behind your back. If the statement is true for your household lift a finger and keep it raised. If the statement is not true, and does not apply to your household, do not lift a finger. After I have read all the statements tell me the number of fingers you have raised. Do not tell me which statements are true, only the total number of statements that are true. Let's begin:

List 1

Group A	Group B
● Someone in the household has consumed rice in the past 24 hours	● Someone in the household has consumed rice in the past 24 hours
● We have purchased livestock in the past 24 hours	● We have purchased livestock in the past 24 hours
● **The beneficiary child in my household ate dark green leafy vegetables or orange colored vegetables or fruit in the past 24 hours**	● Food our household ate in the past 24 hours had been stored in a refrigerator in our house
● Food our household ate in the past 24 hours had been stored in a refrigerator in our house	

List 2

Group A	Group B
● Someone in the household has consumed tea in the past 24 hours	● Someone in the household has consumed tea in the past 24 hours
● Someone in our household travelled by car or tuk tuk to purchase food in the past 24 hours	● Someone in our household travelled by car or tuk tuk to purchase food in the past 24 hours
● Food was prepared using Kerosene cooking in the past 24 hours	● **The beneficiary child in my household ate dark green leafy vegetables or orange colored vegetables or fruit in the past 24 hours**
	● Food was prepared using Kerosene cooking in the past 24 hours

Note: The instructions were read to respondents before each of the two list experiment questions.
Source: Author designed questions.

Table 3: No Design Effect Assumption Tests

Panel A : Joint distributions of the key and non-key items

	List 1				List 2			
	Coef	Robust SE	z	P>z	Coef	Robust SE	z	P>z
Pr(R=0,S=1)	0.00	0.01	0.05	0.52	0.00	0.01	-0.09	0.46
Pr(R=0,S=0)	0.00	0.00	1.00	0.84	0.01	0.01	1.74	0.96
Pr(R=1,S=1)	0.51	0.04	13.05	1.00	0.47	0.04	11.82	1.00
Pr(R=1,S=0)	0.20	0.03	7.81	1.00	0.30	0.03	9.85	1.00
Pr(R=2,S=1)	0.17	0.03	6.78	1.00	0.11	0.02	4.97	1.00
Pr(R=2,S=0)	0.10	0.04	2.70	1.00	0.10	0.03	2.98	1.00
Pr(R=3,S=1)	0.00	0.00	1.00	0.84	0.00	0.00	1.00	0.84
Pr(R=3,S=0)	0.01	0.01	1.06	0.86	0.00	0.01	0.52	0.70

Panel B: Test for design effects (with GMS)

Ha: Pr<0	K[a]	Lambda	P>Lambda	#P>Lambda	K[a]	Lambda	P>Lambda	#P>Lambda
Pr(R ,S=0)	1.00	0.00	0.50	1.00	0.00	0.00	1.00	1.00
Pr(R ,S=1)	1.00	0.00	0.50	1.00	1.00	0.01	0.46	0.93

Note: [a] indicates the number of tests to be conducted. Here, the tests are conducted because each test has a zero coefficient that could statistically be negative. We use the Tsai (2019) package to implement Blair and Imai (2012) design effect test. Panel A tests the joint probability of a given response, where R indicates number of affirmative non-sensitive items and S indicates the sensitive item response. Negative probabilities indicate design effect failure. Panel B tests if any negative probabilities have arisen by chance with the method of generalized moment selection to improve the power of the test.
Source: Authors' elaborations using the Tsai (2019) package and primary data.

Table 4: No Liars Assumption: Frequency of Floor & Ceiling Responses by Group & Question

	List 1				**List 2**			
	Group A		Group B		Group A		Group B	
Response	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
0	1	0.4	1	0.4	3	1.2	3	1.3
1	51	20.1	167	70.8	197	77.6	72	30.5
2	155	61.0	65	27.5	52	20.5	134	56.8
3	46	18.1	3	1.3	2	0.8	26	11.0
4	1	0.4	-	-	-	-	1	0.4
Out of range	-	-	2	0.8	3	1.2	-	-

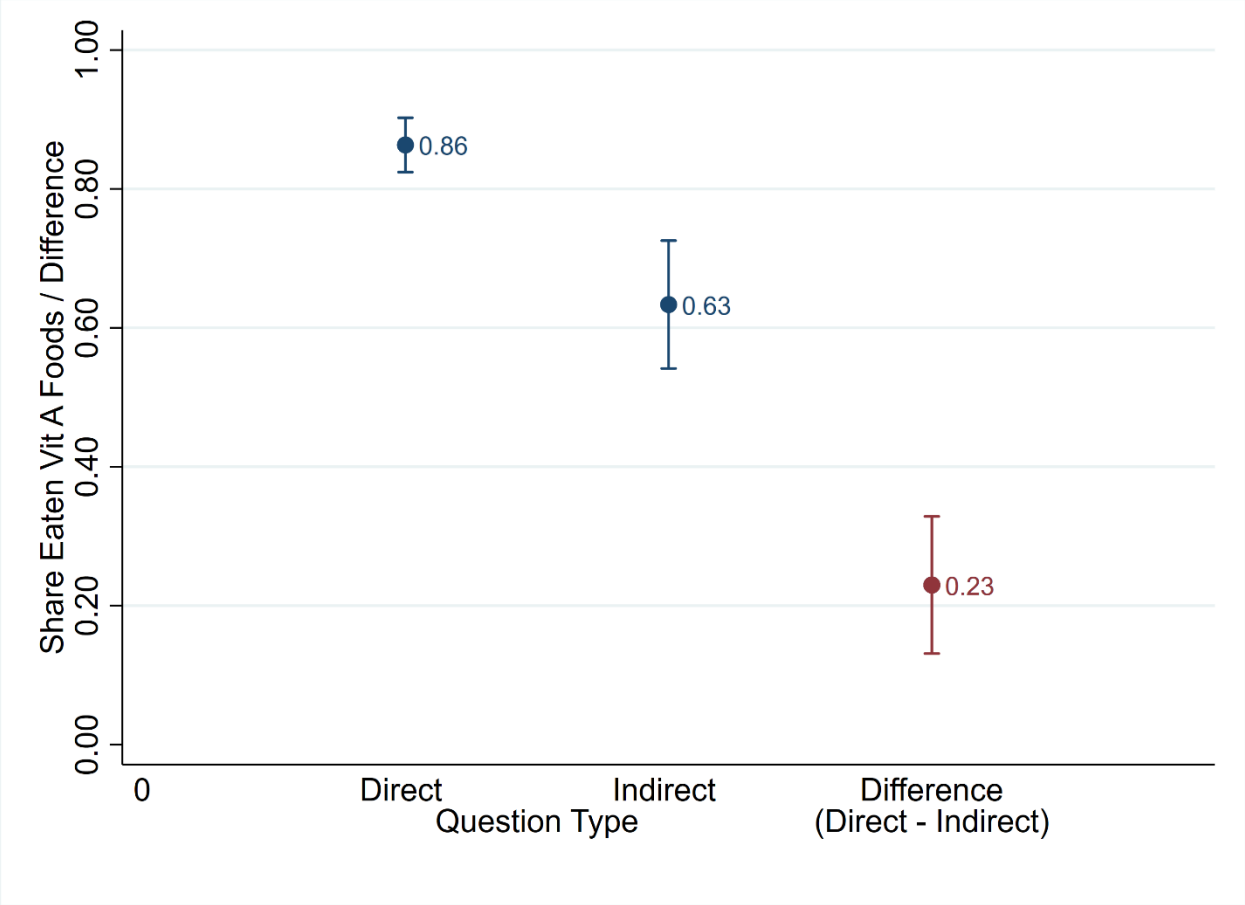
Note: A floor response is zero for both list 1 and list 2. A ceiling response is three for Group B list 1 and Group A list 2 and a response of four for Group A list 1 and Group B list 2. A high prevalence of ceiling responses among the short list group draws into question the validity of the no liars assumption.
Source: Authors' elaborations from primary data. Responses outside the allowable range are excluded from analysis.

Table 5: Vitamin A Foods Consumption by Question Type & Social Desirability Bias

	Direct	Indirect	Difference
Prevalence	86.4%	63.4%	23.0%
SE	0.015	0.036	0.038

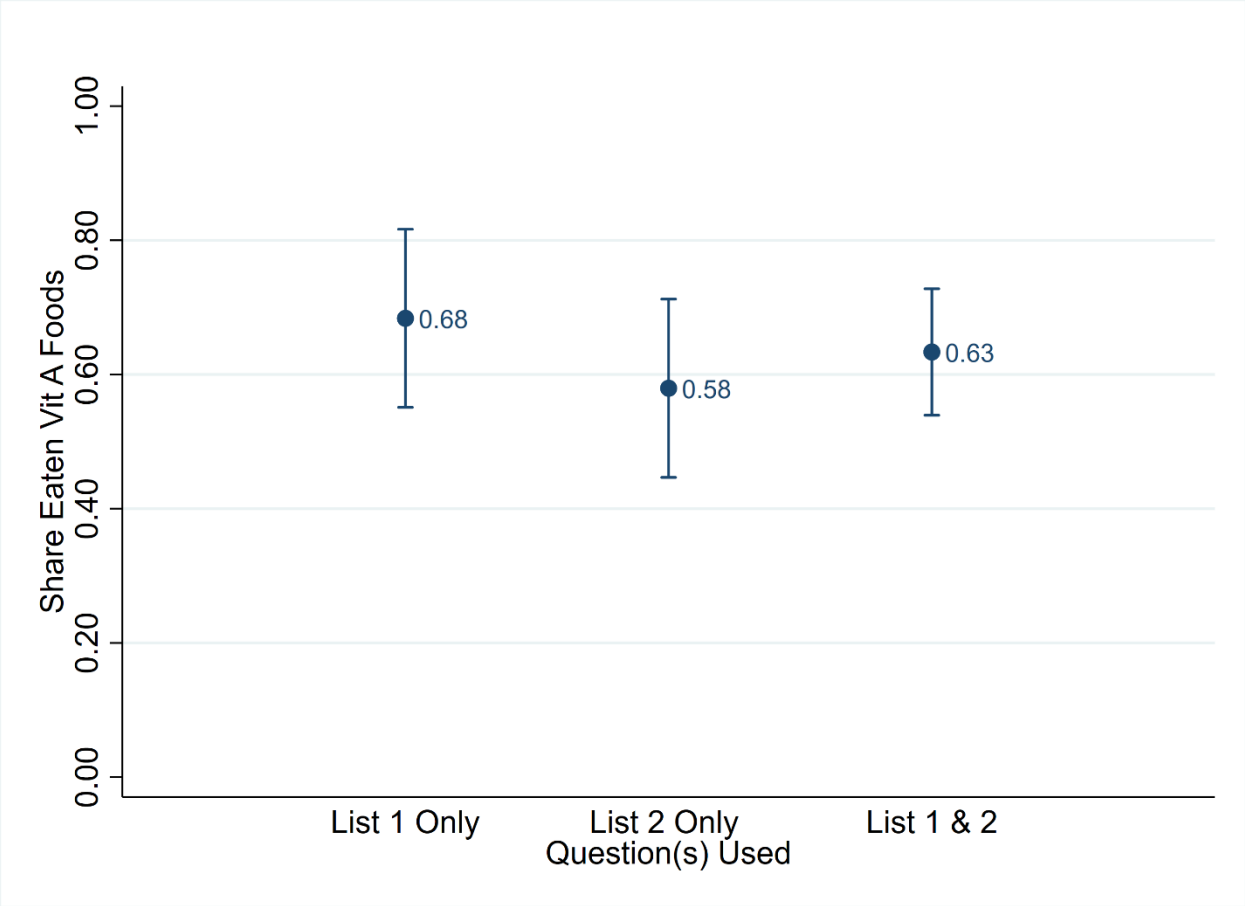
Note: The direct measure is estimated using three direct, business as usual food consumption questions. The indirect measure is estimated using an indirect, dual list experiment. The difference is the direct estimate minus the indirect measurement with 99 percent bootstrap confidence interval drawn from 1,000 samples.
Source: Authors' elaborations from primary data.

Figure 3 : Vitamin A Food Consumption Prevalence & Social Desirability Bias



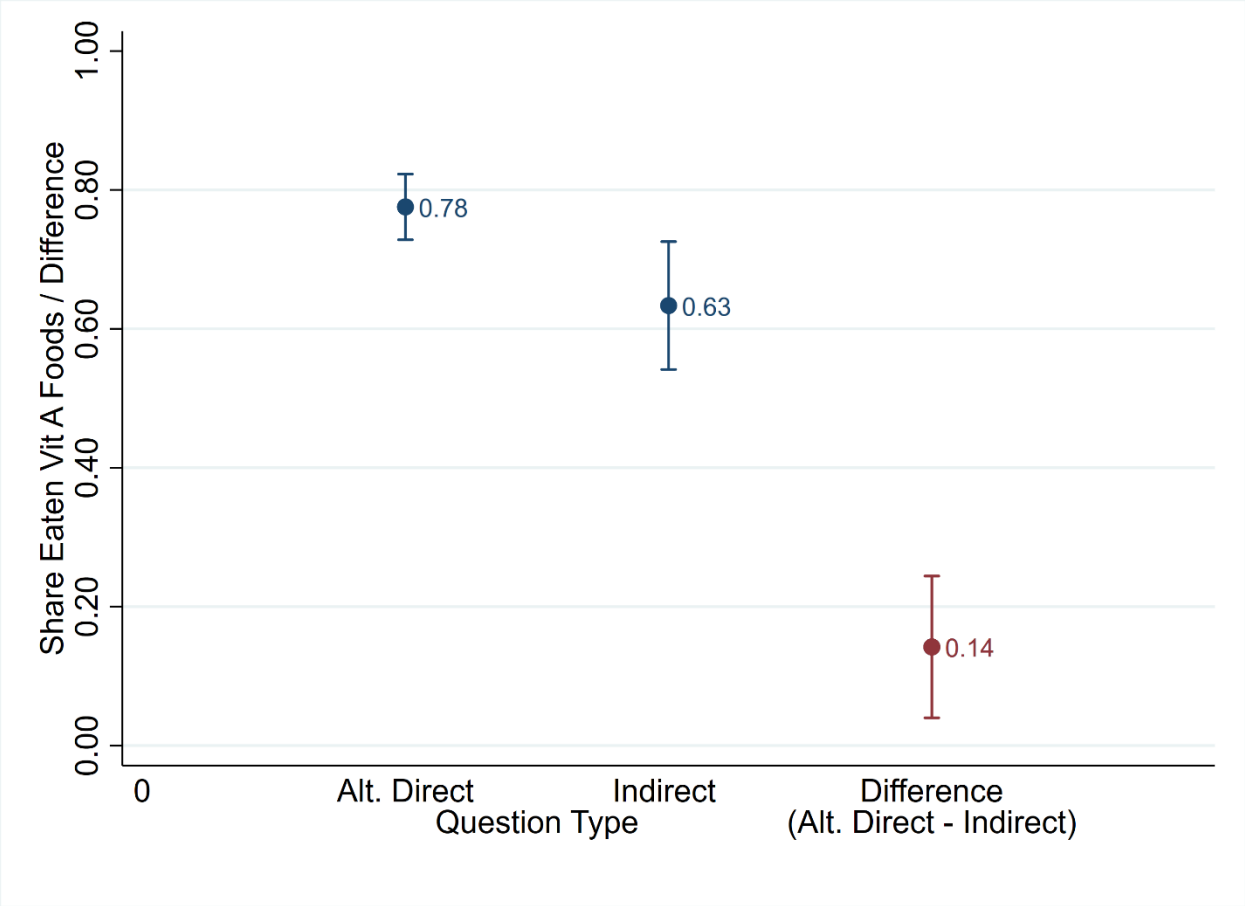
Note: Confidence intervals are reported at the 1% significance level. The direct measure is estimated using three direct, business as usual food consumption questions. The indirect measure is estimated using an indirect, dual list experiment. The difference reports 99% bootstrap confidence intervals using 1,000 draws.
Source: Authors' elaborations from primary data.

Figure 4: Single List Compared to Dual List: Validity Check



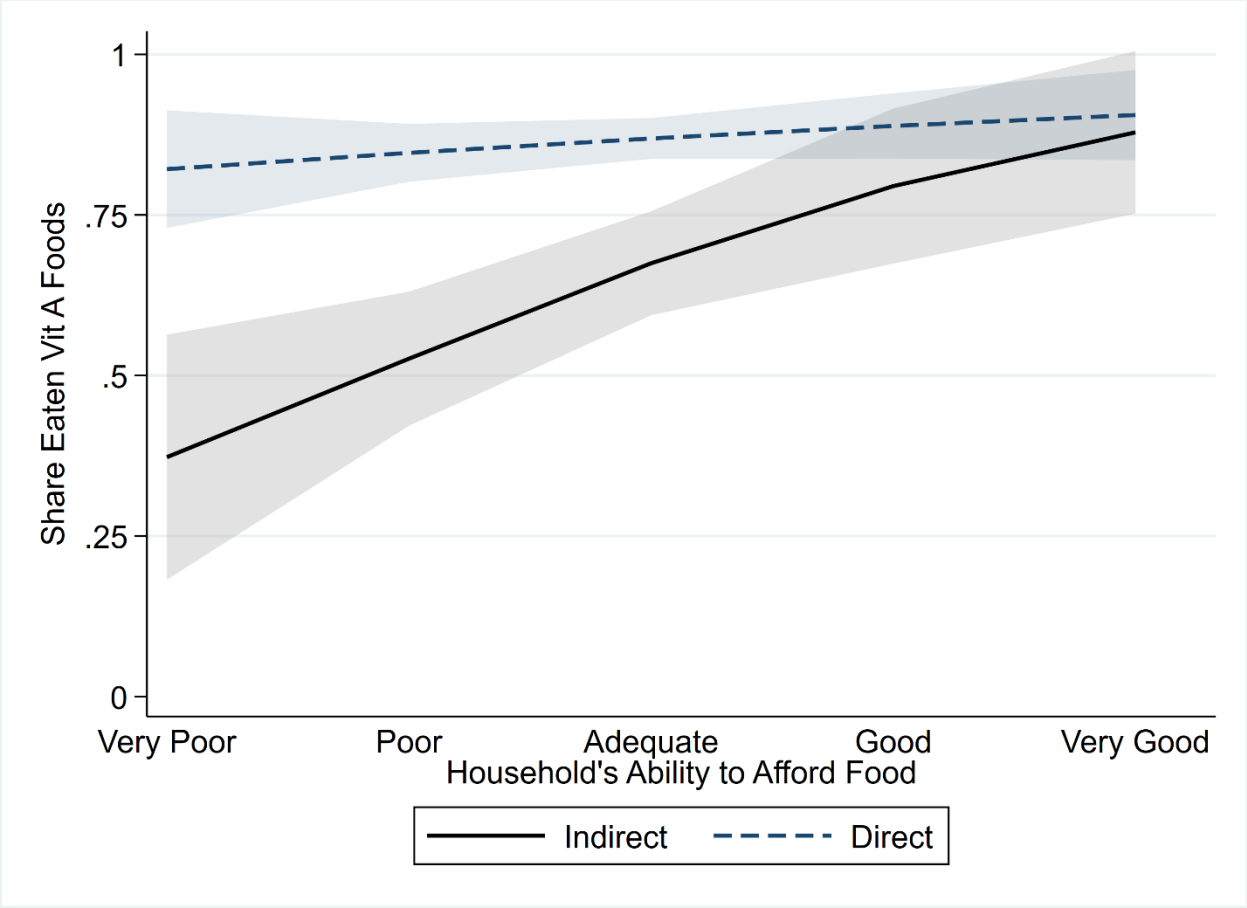
Note: Confidence intervals are reported at the 1% significance level. List 1 and 2 show estimated prevalence using a single list experiment for list question 1 and 2 respectively. List question 1 & 2 uses a dual list experiment to improve efficiency.
Source: Authors' elaborations from primary data.

Figure 5: Alternative Direct Measure - No Orange/Yellow Vegetables



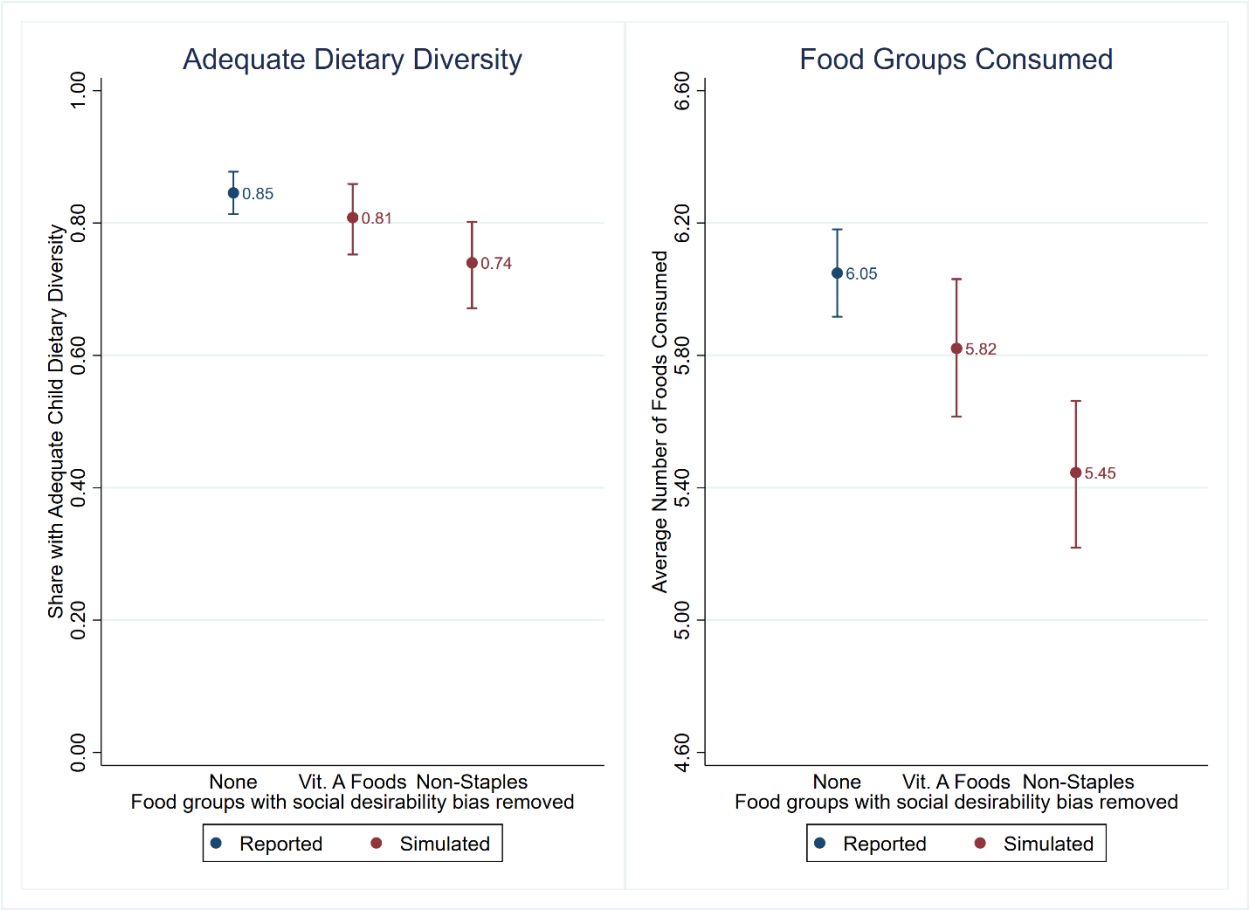
Note: Confidence intervals are reported at the 1% significance level. The alternative direct measure is estimated using *two* direct, business as usual food consumption questions. The indirect measure is estimated using an indirect, dual list experiment as in Figure 3. The difference reports 99% bootstrap confidence intervals using 1,000 draws.
Source: Authors' elaborations from primary data.

Figure 6 Vitamin A-Rich Food Consumption Prevalence by Level of Ability to Afford Food and Question Type



Note: The indirect line is estimated using the Tsai (2019) nonlinear least-squares dual list estimator with ability to afford food (five-item Likert scale) and a mean zero asset index as covariates. The direct question prevalence is estimated with a logistic regression using the same covariates. Vitamin A-rich food consumption was predicted at each level of the covariate for both the indirect and direct measures. The shaded regions indicates the 95% confidence intervals. The vertical distance between the direct and indirect measures is the rate of misreporting at each level.
Source: Authors' elaborations from primary data.

Figure 7: Impact of Social Desirability Bias on Childhood Dietary Diversity Adequacy and Food Groups Consumed



Note: The first metric estimates the business as usual estimation using direct reports and the 95\% confidence interval. The second and third metrics in the left (right) panel simulate the share with adequate dietary diversity (mean number of food groups consumed) if social desirability bias was removed from the vitamin A food group and all non-staple food groups respectively. Non-staple food groups: dairy, flesh foods, eggs, vitamin A foods, and other fruits and vegetables. Confidence intervals for the simulated measures use the 2.5th and 97.5th percentile.
Source: Authors' elaborations from simulation analysis.

REFERENCES

Abate, Gashaw T., Kibrom A. Abay, Jordan Chamberlin, and Samuel Sebsibie. 2025. "Rural Land Rental Markets in Developing Countries: Can Survey Design Innovations Improve Land Market Participation Statistics?" *American Journal of Agricultural Economics* n/a (n/a). https://doi.org/10.1111/ajae.70022.

Ahlquist, John S. 2018. "List Experiment Design, Non-Strategic Respondent Error, and Item Count Technique Estimators." *Political Analysis* 26 (1): 34–53. https://doi.org/10.1017/pan.2017.31.

Ahmed, Akhter, John Hoddinott, and Shalini Roy. 2025. "Food Transfers, Cash Transfers, Behavior Change Communication and Child Nutrition: Evidence from Bangladesh." *The World Bank Economic Review*.

Alderman, Harold. 2015. *Leveraging Social Protection Programs for Improved Nutrition*. Washington, DC: World Ban. https://doi.org/10.1596/2445.

Arimond, Mary, and Marie T. Ruel. 2004. "Dietary Diversity Is Associated with Child Nutritional Status: Evidence from 11 Demographic and Health Surveys." *The Journal of Nutrition* 134 (10): 2579–85. https://doi.org/10.1093/jn/134.10.2579.

Aronow, Peter m., Alexander Coppock, Forrest W. Crawford, and Donald P. Green. 2015. "Combining List Experiment and Direct Question Estimates of Sensitive Behavior Prevalence." *Journal of Survey Statistics and Methodology* 3 (1): 43–66. https://doi.org/10.1093/jssam/smu023.

Blair, Graeme, Alexander Coppock, and Margaret Moor. 2020. "When to Worry about Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments." *American Political Science Review* 114 (4): 1297–315. https://doi.org/10.1017/S0003055420000374.

Blair, Graeme, and Kosuke Imai. 2012. "Statistical Analysis of List Experiments." *Political Analysis* 20 (1): 47–77. https://doi.org/10.1093/pan/mpr048.

Bouguen, Adrien, and Andrew Dillon. 2026. "Beyond Poverty Reduction: Evidence from a Multifaceted Program on Poverty, Nutrition and Child Development." *Journal of Development Economics* 178 (January): 103596. https://doi.org/10.1016/j.jdeveco.2025.103596.

Carneiro, Pedro, Lucy Kraftman, Giacomo Mason, Lucie Moore, Imran Rasul, and Molly Scott. 2021. "The Impacts of a Multifaceted Prenatal Intervention on Human Capital Accumulation in Early Life." *American Economic Review* 111 (8): 2506–49. https://doi.org/10.1257/aer.20191726.

Chuang, Erica, Pascaline Dupas, Elise Huillery, and Juliette Seban. 2021. "Sex, Lies, and Measurement: Consistency Tests for Indirect Response Survey Methods." *Journal of Development Economics* 148 (January): 102582. https://doi.org/10.1016/j.jdeveco.2020.102582.

Cullen, Claire. 2023. "Method Matters: The Underreporting of Intimate Partner Violence." *The World Bank Economic Review* 37 (1): 49–73. https://doi.org/10.1093/wber/lhac022.

Department of Census and Statistics. 2016. *Sri Lanka Demographic and Health Survey*. Department of Census and Statistics.

Droitcour, J., Rachel A. Caspar, Michael Hubbard, T. L. Parsley, Wendy Visscher, and T. M. Ezzati. 1991. *The Item Count Technique as a Method of Indirect Questioning: A Review of Its Development and a Case Study Application*. January 1.

Eady, Gregory. 2017. "The Statistical Analysis of Misreporting on Sensitive Survey Questions." *Political Analysis* 25 (2): 241–59. https://doi.org/10.1017/pan.2017.8.

FAO. 2023. *Special Report – FAO/WFP Crop and Food Security Assessment Mission (CFSAM) to the Democratic Socialist Republic of Sri Lanka: 25 May 2023*. CFSAMs Special Reports, No. 03/2023. Rome, Italy: FAO. https://doi.org/10.4060/cc6202en.

Field, Erica, and Elisa M. Maffioli. 2025. "Are Behavioral Change Interventions Needed to Make Cash Transfer Programs Work for Children? Experimental Evidence from Myanmar." *Economic Development and Cultural Change* 73 (3): 1187–220. https://doi.org/10.1086/730631.

Gentilini, Ugo (Editor). 2022. *Social Protection, Food Security and Nutrition: An Update of Concepts, Evidence and Select Practices in South Asia and Beyond*. World Bank. https://doi.org/10.1596/38210.

Ghose, Devaki, Eduardo Fraga, and Ana Fernandes. 2026. *Trade Policy and Access to Intermediate Inputs*. World Bank Group.

Glynn, Adam N. 2013. "What Can We Learn with Statistical Truth Serum?: Design and Analysis of the List Experiment." *Public Opinion Quarterly* 77 (S1): 159–72. https://doi.org/10.1093/poq/nfs070.

Groot, Richard de, Tia Palermo, Sudhanshu Handa, Luigi Peter Ragno, and Amber Peterman. 2017. "Cash Transfers and Child Nutrition: Pathways and Impacts." *Development Policy Review: The Journal of the Overseas Development Institute* 35 (5): 621–43. https://doi.org/10.1111/dpr.12255.

Headey, Derek, Dilini Hemachandra, and Immacolata Ranucci. 2024. *An Empirical Investigation of UNICEF's Nutrition-Sensitive Short-Term Cash Transfer Programme in Sri Lanka*. August.

Hirvonen, Kalle, Jessica Leight, Daniel O. Gilligan, Hiwot Mekonnen Mesfin, Michael Mulford, and Haleluya Tesfaye. 2025. *The Impact of a Nutrition-Sensitive Graduation Model Program on Child Nutrition*.

Imai, Kosuke. 2011. "Multivariate Regression Analysis for the Item Count Technique." *Journal of the American Statistical Association* 106 (494): 407–16. https://doi.org/10.1198/jasa.2011.ap10415.

Jouvin, Marine. 2023. "Addressing Social Desirability Bias When Measuring Child Labor Use: An Application to Cocoa Farms in Côte d'Ivoire." *The World Bank Economic Review*, September 19, lhad030. https://doi.org/10.1093/wber/lhad030.

Kim, Seung Hyun, and Sangmook Kim. 2016. "National Culture and Social Desirability Bias in Measuring Public Service Motivation." *Administration & Society* (US) 48 (4): 444–76. https://doi.org/10.1177/0095399713498749.

Leight, Jessica, Daniel Gilligan, Melissa Hidrobo, Harold Alderman, and Michael Mulford. 2026. "Can a Light-Touch Graduation Model Enhance Livelihood Outcomes? Evidence from Ethiopia." *Journal of Development Economics* 179 (February): 103682. https://doi.org/10.1016/j.jdeveco.2025.103682.

Lépine, Aurélia, Carole Treibich, and Ben D'Exelle. 2020. "Nothing but the Truth: Consistency and Efficiency of the List Experiment Method for the Measurement of Sensitive Health Behaviours." *Social Science & Medicine (1982)* 266 (December): 113326. https://doi.org/10.1016/j.socscimed.2020.113326.

Lépine, Aurélia, Carole Treibich, Cheikh Tidiane Ndour, Khady Gueye, and Peter Vickerman. 2020. "HIV Infection Risk and Condom Use among Sex Workers in Senegal: Evidence from the List Experiment Method." *Health Policy and Planning* 35 (4): 408–15. https://doi.org/10.1093/heapol/czz155.

Levere, Michael, Gayatri Acharya, and Prashant Bharadwaj. 2024. "The Role of Information and Cash Transfers on Early-Childhood Development: Short- and Long-Run Evidence from Nepal." *Economic Development and Cultural Change* 72 (3): 1267–93. https://doi.org/10.1086/723203.

Little, Madison T., Keetie Roelen, Brittany C. L. Lange, et al. 2021. "Effectiveness of Cash-plus Programmes on Early Childhood Outcomes Compared to Cash Transfers Alone: A Systematic Review and Meta-Analysis in Low- and Middle-Income Countries." *PLOS Medicine* 18 (9): e1003698. https://doi.org/10.1371/journal.pmed.1003698.

Menon, Purnima, Phuong Hong Nguyen, Kuntal Kumar Saha, et al. 2016. "Combining Intensive Counseling by Frontline Workers with a Nationwide Mass Media Campaign Has Large Differential Impacts on Complementary Feeding Practices but Not on Child Growth: Results of a Cluster-Randomized Program Evaluation in Bangladesh." *The Journal of Nutrition* 146 (10): 2075–84. https://doi.org/10.3945/jn.116.232314.

Ministry of Health, . 2022. *Nutrition Month 2022*. Sri Lanka: Family Health Bureau.

Olney, Deanna K., Aulo Gelli, Neha Kumar, Harold Alderman, Ara Go, and Ahmed Raza. 2022. "Social Assistance Programme Impacts on Women's and Children's Diets and Nutritional Status." *Maternal & Child Nutrition* 18 (4): e13378. https://doi.org/10.1111/mcn.13378.

Premand, Patrick, and Oumar Barry. 2022. "Behavioral Change Promotion, Cash Transfers and Early Childhood Development: Experimental Evidence from a Government Program in a Low-Income Setting." *Journal of Development Economics* 158 (September): 102921. https://doi.org/10.1016/j.jdeveco.2022.102921.

Pruce, Kate, Roz Price, and Rachel Sabates-Wheeler. 2025. *Learning from Cash Plus: A Summary of Evidence*. Institute of Development Studies. https://doi.org/10.19088/CPAN.2025.005.

Rawat, Rahul, Phuong Hong Nguyen, Lan Mai Tran, et al. 2017. "Social Franchising and a Nationwide Mass Media Campaign Increased the Prevalence of Adequate Complementary Feeding in Vietnam: A Cluster-Randomized Program Evaluation." *The Journal of Nutrition* 147 (4): 670–79. https://doi.org/10.3945/jn.116.243907.

Reynolds, William M. 1982. "Development of Reliable and Valid Short Forms of the Marlowe-Crowne Social Desirability Scale." *Journal of Clinical Psychology* 38 (1): 119–25. https://doi.org/10.1002/1097-4679(198201)38:1%3C119::AID-JCLP2270380118%3E3.0.CO;2-I.

Rosenfeld, Bryn, Kosuke Imai, and Jacob N. Shapiro. 2016. "An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions." *American Journal of Political Science* 60 (3): 783–802. https://doi.org/10.1111/ajps.12205.

Ryan, Ann Marie, Jacob Bradburn, Sarena Bhatia, et al. 2021. “In the Eye of the Beholder: Considering Culture in Assessing the Social Desirability of Personality.” *The Journal of Applied Psychology* 106 (3): 452–66. https://doi.org/10.1037/apl0000514.

Schwab, Benjamin, Sarah Janzen, Nicholas P. Magnan, and William M. Thompson. 2021. “SWINDEX: Stata Module to Create a Standardized Weighted Index of Multiple Indicator Variables.” *Statistical Software Components*, February 24. https://ideas.repec.org//c/boc/bocode/s458912.html.

Sedlmayr, Richard, Anuj Shah, and Munshi Sulaiman. 2020. “Cash-plus: Poverty Impacts of Alternative Transfer-Based Approaches.” *Journal of Development Economics* 144 (May): 102418. https://doi.org/10.1016/j.jdeveco.2019.102418.

Soofi, Sajid Bashir, Shabina Ariff, Gul Nawaz Khan, et al. 2022. “Effectiveness of Unconditional Cash Transfers Combined with Lipid-Based Nutrient Supplement and/or Behavior Change Communication to Prevent Stunting among Children in Pakistan: A Cluster Randomized Controlled Trial.” *The American Journal of Clinical Nutrition* 115 (2): 492–502. https://doi.org/10.1093/ajcn/nqab341.

Tadesse, Getaw, Gashaw T. Abate, and Tadiwos Zewdie. 2020. “Biases in Self-Reported Food Insecurity Measurement: A List Experiment Approach.” *Food Policy* 92 (April): 101862. https://doi.org/10.1016/j.foodpol.2020.101862.

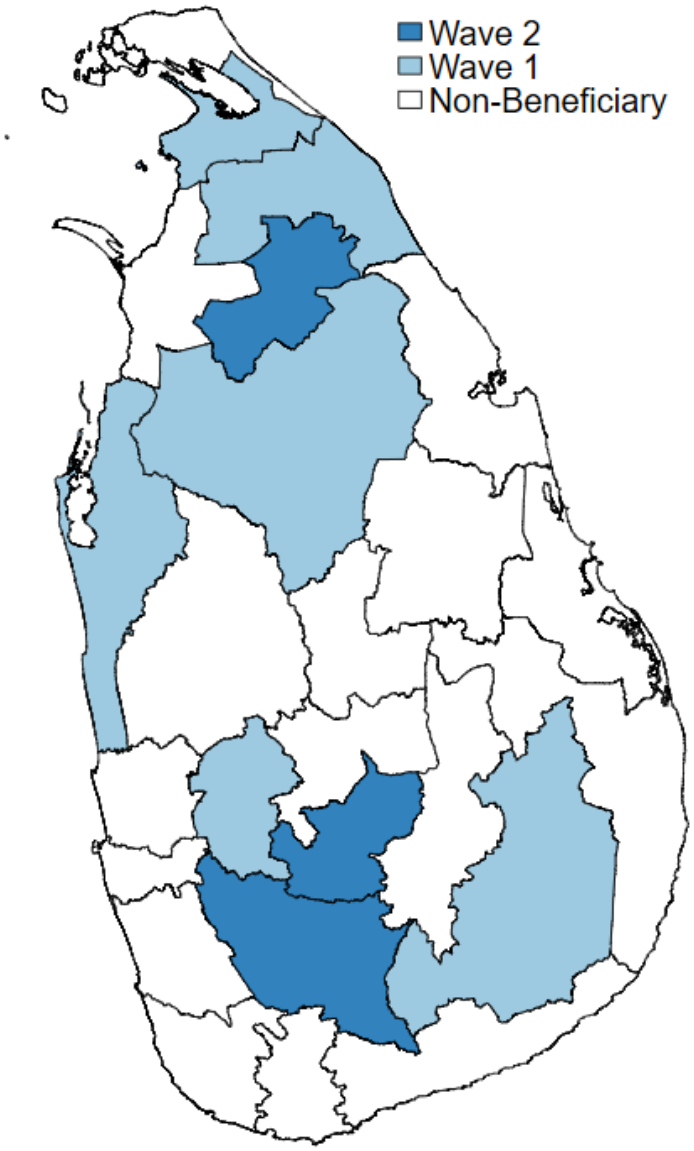
Tsai, Chi-lin. 2019. “Statistical Analysis of the Item-Count Technique Using Stata.” *The Stata Journal: Promoting Communications on Statistics and Stata* 19 (2): 390–434. https://doi.org/10.1177/1536867X19854018.

Webb Girard, Aimee, Emma Waugh, Sarah Sawyer, Lenette Golding, and Usha Ramakrishnan. 2020. “A Scoping Review of Social-Behaviour Change Techniques Applied in Complementary Feeding Interventions.” *Maternal & Child Nutrition* 16 (1): e12882. https://doi.org/10.1111/mcn.12882.

World Bank. 2022. *Sri Lanka Development Update 2022: Protecting the Poor and Vulnerable in a Time of Crisis*. World Bank. https://doi.org/10.1596/35833.

APPENDIX

Appendix Figure 1 Program Implementation Waves



Source: Figure created by authors using information from Headey et al. (2024).

Appendix Table 1 Share of Households Consuming Each Food Group

Food Group	Share Consumed
Breastmilk	75.9%
Grains, roots, & tubers	95.9%
Legumes & nuts	65.7%
Dairy	79.6%
Flesh foods	71.4%
Eggs	59.4%
Vit-A rich fruits & vegetables	86.3%
Other fruits & vegetables	74.4%

Source: Authors' elaborations from primary data.

Appendix Table 2 Food Affordability Coefficients & Robustness Checks

	(1)	(2)	(3)	(4)
Food Affordability	0.625***	0.685**	0.680**	0.667**
	(0.242)	(0.275)	(0.278)	(0.280)
Asset Index	-0.076	-0.047	-0.108	-0.126
	(0.187)	(0.221)	(0.238)	(0.242)
Tamil			-0.706	-0.686
			(0.462)	(0.480)
Over Primary Education				0.619
				(0.569)
District Fixed Effects	No	Yes	Yes	Yes
* p<0.10; ** p<0.05; *** p<0.01				

Note: Coefficients are from the sensitive item estimation and are on the logit scale. Standard errors in parentheses.
Source: Authors' elaborations from primary data.

ALL IFPRI DISCUSSION PAPERS

All discussion papers are available **here**

They can be downloaded free of charge

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
www.ifpri.org

IFPRI HEADQUARTERS
1201 Eye Street, NW
Washington, DC 20005 USA
Tel.: +1-202-862-5600
Fax: +1-202-862-5606
Email: ifpri@cgiar.org