DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 50, ARTICLE 32, PAGES 929–966
PUBLISHED 7 MAY 2024
https://www.demographic-research.org/Volumes/Vol50/32/
DOI: 10.4054/DemRes.2024.50.32

Research Article

Gone and forgotten? Predictors of birth history omissions in India

Sharan Sharma

Sonalde Desai

Debasis Barik

O.P. Sharma

Contents

Demographic Research: Volume 50, Article 32
Research Article

Gone and forgotten? Predictors of birth history omissions in India

Sharan Sharma[1]

Sonalde Desai[2]

Debasis Barik[3]

O.P. Sharma[4]

Abstract

BACKGROUND
Fertility histories are subject to measurement errors such as incorrect birth dates, incorrect birth orders, incorrect sex, and omissions. These errors can bias demographic estimates such as fertility rates and child mortality rates.

OBJECTIVE
We focus on births missing in fertility histories. We estimate the prevalence of such omissions and study their associated factors.

METHODS
We leverage a panel survey (the India Human Development Survey) where the same women were interviewed in two waves several years apart. We compare data across waves and identify omitted births. Omissions in the second wave are modeled as a function of several child, mother, household, and survey interviewer variables. Models are fit separately to omissions reported alive or dead in the first wave.

RESULTS
We conservatively estimate the prevalence of omissions at 4%. A large majority of omitted births are those of dead children, especially infants, with children in poorer households at greater risk of being omitted. For children alive in wave 1, female children are much more likely to be omitted in wave 2 compared to male children. Interviewers can detect respondent behaviors associated with omissions.

[1] Assistant Research Professor, Department of Sociology and Joint Program in Survey Methodology, University of Maryland, College Park, Maryland, USA, and Nonresident Fellow, National Council of Applied Economic Research, New Delhi, India. Email: snsharma@umd.edu.
[2] Distinguished University Professor, Department of Sociology, University of Maryland, College Park, Maryland, USA, and Professor, National Council of Applied Economic Research, New Delhi, India.
[3] Senior Fellow, National Council of Applied Economic Research, New Delhi, India.
[4] Fellow, National Council of Applied Economic Research, New Delhi, India.

CONCLUSIONS
Omissions in fertility histories are non-ignorable. They do not randomly occur, and they affect some population subgroups and some interview contexts more than others.

CONTRIBUTIONS
We investigate the understudied but important phenomenon of omitted births in fertility histories. We bring attention to possible biases in demographic estimates. We shed light on the survey process and propose strategies for minimizing bias through improved survey design.

1. Introduction

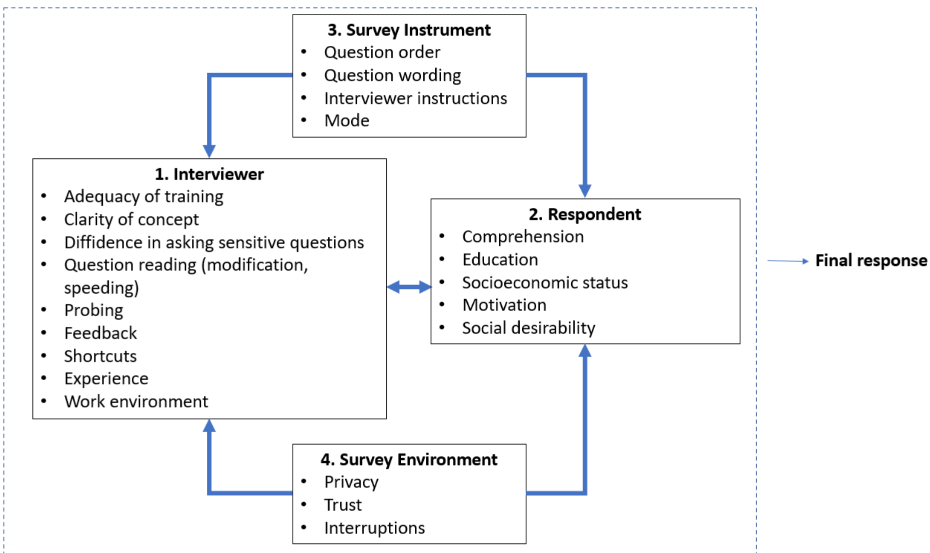
Despite the enormous contributions made by the World Fertility Survey (WFS) and its successor, the Demographic and Health Survey (DHS), concerns about the quality of data have persisted (United Nations 1984; Meekers 1991; Pullum and Becker 2014; Bardley 2016). Helleringer et al. (2020) outline four possible errors when reporting birth history data: date displacement (misreporting birth dates), age error (misreporting ages of children), omissions (births not reported), and misclassifications (e.g., stillbirths reported as neonatal deaths). While methods to correct errors such as age displacement are available, the problem of omissions is "much more intractable" (Blacker 1994; Choi, Li, and Zachary 2018). If births are omitted in the collection of fertility/birth histories, estimates of fertility and mortality may be understated. Moreover, if specific types of births are omitted (e.g., dead children), estimates of fertility and mortality differentials may be biased.

Studies evaluating the quality of fertility history data are uncommon (Murphy 2009; Helleringer et al. 2020). A major reason is that such evaluations require making some assumptions about what "good" data would look like. For example, comparing fertility history data from a given survey to those obtained from other sources, such as registries or censuses (Espeut and Becker 2015), assumes that the latter are the "gold standard," which may not always be valid (Michael et al. 2013). Pullum and Becker (2014) assess birth history omissions using deviations from expected values of the following three criteria: sex ratio at birth, sex ratio of neonatal births, and proportion of infant deaths that are neonatal. Methods using expected values may assume that sex ratios at birth or sex differences in infant mortality rates are biologically determined, which may not be helpful in societies characterized by preferences for sons (Arnold, Choe, and Roy 1998) and the prevalence of sex-selective abortions (Robitaille and Chatterjee 2018). We take a different approach in this paper by comparing fertility histories of the same women in the India Human Development Survey (IHDS) – a panel survey with two waves conducted seven years apart – to examine the prevalence of omitted births. This unique analysis contributes to the literature in three ways. First, it provides an estimate of birth omissions in fertility histories. Second, it examines factors leading to bias; we explore the role of child, mother, and household characteristics in shaping birth omissions and therefore differences in estimates of fertility and infant and child mortality. Finally, we explore the role of interviewers in influencing the quality of fertility history data.

2. Sources of omissions

Figure 1 suggests a framework for understanding omissions based on roles played by the interviewer, the respondent, the survey instrument, and the environment during the survey process.

Figure 1: Role of the interviewer, respondent, survey instrument, and survey environment in impacting omissions in birth history data



Demographic surveys often organize intensive training for interviewers before fieldwork commences. However, many such surveys have long and complicated instruments; training covers a lot of content resulting in only “one day of two to three weeks of training . . . to train interviewers on collecting birth histories” (Leone, Sochas, and Coast 2021; ICF 2020; ICF Macro 2009). This might result in interviewers not fully absorbing the concept of a fertility history (e.g., mistaking it for a list of children present in the household) or forgetting the concept, a phenomenon that is not uncommon (Billiet and Loosveldt 1988). Training programs tend to focus on the questionnaire rather than developing soft skills, such as asking sensitive questions. This impacts birth histories since respondents are asked about children who have died.

Moreover, retraining interviewers during fieldwork is not common (Olson and Peytchev 2007), resulting in persistent deficient interviewing behavior. Examples include rewording a question (assuming it's a properly tested question) so that its meaning is lost – for example, "I am now going to ask you about the birth of your children" rather than "I am now going to ask you about all the births you have had. Please report all live births, whether they are still alive or not." Deficient interviewing behavior also includes reading a question too quickly, leading to a lack of comprehension by the respondent (Holbrook et al. 2020); not paying attention to interviewing instructions (a lack of clear interviewing instructions is a problem in itself), such as not proceeding chronologically; and a lack of probing (Mangione, Fowler, and Louis 1992) – for example, if the gap between two births is large, probing would involve gently and neutrally checking whether there were any births in that gap. Other deficient interviewing behavior includes improper probing (Schober and Conrad 1997), including leading probes such as, "So I guess you did not have any births after [name] was born"; using shortcuts, such as referring to the household roster to prompt the respondent (resulting in dead children and married-out children being omitted); not including recent births, so that questions on topics like vaccination for young children are skipped altogether; and improper feedback (e.g., saying, "Oh, how sad that you had *two* children who passed away."). While interviewers' skills can improve as they do more interviews (Olson and Smyth 2020), fatigue can also set in (Japec 2007; Koch et al. 2009; DeMatteis et al. 2020: 6), leading to careless interviewing and potentially more birth omissions (Cannell and Fowler 1964).

At the respondent's end, since a birth is a salient event (Githens et al. 1993), unintentional recall errors in birth histories are likely to be uncommon. However, several other things may cause a birth to go unreported. In developing countries, the concept of a demographic survey can be novel, especially for people from rural areas, and can be confused with official data collection exercises such as censuses, leading to misreporting of births (Feeney and Jianhua 1994). Respondents may fail to understand the need to report all live births (whether the child is currently living with the respondent or not), thereby specifically impacting the inclusion of nonresident children and dead children. Comprehension issues can be exacerbated by a linguistically complex environment where the interviewer and respondent are not fluent in each other's dialect. Respondents with lower education levels may be associated with more omissions since education is often a marker of cognitive sophistication (Knäuper 1999; Krosnick and Alwin 1987), which is in turn associated with question comprehension.

Perinatal deaths (including infant deaths) are traumatic events (Gold, Dalton, and Schwenk 2007; Murphy, Shevlin, and Elkit 2014), with mothers often experiencing guilt and self-blame (Gold, Sen, and Leon 2018); memories of such deaths could be suppressed, distressing to recall, and/or painful to report. These sensitivities would also make probing for these children particularly challenging for interviewers. In countries like India where a large proportion of births have historically taken place at home (58.5% of birthing took place at home in 2005–2006; Patel et al. 2021), these factors perhaps drive a "severe underreporting of negative reproductive events in surveys" (Haws et al. 2010). Social desirability (Johnson and Van de Vijver 2003; Krumpal 2013) can result in omitting more daughters than sons from a birth history in cultures with a preference for sons, which might impact rural areas more than urban areas (Clark 2000). Finally, respondents could be multitasking (Aizpurua et al. 2018) or be interrupted during the interview, causing them to satisfice (Krosnick 1999; Krosnick, Narayan, and Smith 1996) by omitting some births. Satisficing can also be caused by fatigue from responding to long questionnaires.

3. Study survey and data

We use data from the IHDS, a large-scale pan-India panel survey. Fieldwork for the first wave of IHDS was completed in 2004–2005 with a sample size of 41,554 households. The second wave was completed in 2011–2012 and involved interviewing 34,621 of the wave 1 households and any households split from the root households that still resided in the same locality. (A refresher sample was also recruited to compensate for non-contactable wave 1 households.) Apart from a household questionnaire typically administered to the head of the household, both waves had a separate set of questions administered to "eligible women" (EWs) in the household, defined as ever-married women between 15 and 49 years of age. EW-specific questions covered several domains, such as health beliefs, gender relations, marital history, and fertility history, and were administered by female interviewers. Interviewers were trained to ask for a private space to conduct the interview, but such a setting was not always available. Of the 33,539 EWs interviewed in the first wave, 25,475 were reinterviewed in the second wave. A large part of this attrition was due to sample non-contact (rather than refusals), especially due to the long time between waves (seven years). For example, some households in government employee colonies had moved out due to transfers or retirement. In 2004 the proportion of households with cellphones was also relatively low, which prevented follow-up.

Both IHDS waves used a pen-and-paper personal interview (PAPI) approach and captured birth histories identically, using the same question wording and format. Appendix A contains a screenshot of the birth history question, which occurs toward the end of the EW instrument. Interviewers were trained not only to ask respondents to chronologically report birth history but also to clearly explain that all live births must be reported, irrespective of whether the child was still alive or not. The birth history table was split into two parts. Initially, the interviewer focused only on collecting the children's names and checking whether a child was still a household member. After this, the interviewer went back to the first birth and went through all births in succession, collecting more information from the respondent, such as the child's sex, birth date, present location, and age when they died (if applicable). The interviewer also checked if there were any live births between the two births just recorded. Collecting data in this fashion was done to reduce satisficing; if respondents knew that detailed information was going to be collected for each birth, they might omit some births.

4. Data preparation and exploratory analysis

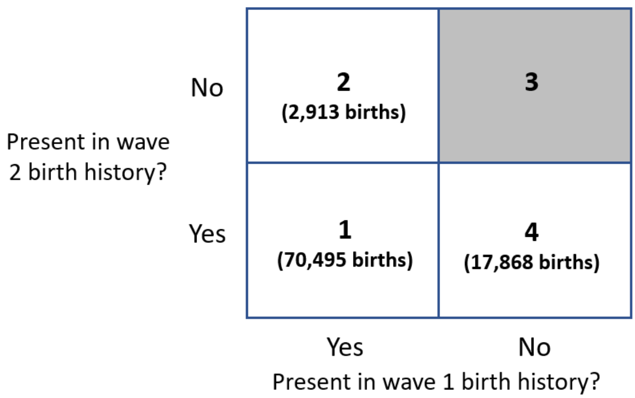
Our analysis is centered around examining the consistency of fertility histories reported by the same woman in both waves of the survey. The separation of waves by seven years is an advantage, since a respondent would be very unlikely to remember what she reported in the prior interview; each interview gives an opportunity to collect fresh fertility history data. However, this separation of waves posed a problem for data preparation. Some communities in India do not name a child until a naming ceremony occurs, several months or even years after the birth. Many yet-to-be named children in wave 1 were recorded as "baby girl" or "baby boy," but wave 2 would have these children's actual names. In many other cases, young children were recorded by a family name or nickname (e.g., Chhoti or Tinu, both meaning "young one"). Matching only by age would not work since age-related data are commonly misreported in India (Pardeshi 2010; Borkotoky and Unisa 2014). With all these issues, if the data were used as is, many births recorded in wave 1 would be misleadingly classified as missing in wave 2. We undertook an intensive exercise, which included calling up the respondent in many cases, to make sure that the 70,000-plus birth records in wave 1 were correctly matched to wave 2 records or classified as omitted in wave 2. Specifically:

a) We first matched births electronically using the child's name, birth date, and sex. Birth records that did not pass this automatic matching were matched using the following rules:

b) When mismatched birth dates were within six months of each other across the two waves, we assumed it was the same child. In many cases, the names of such children were also different, but these differences were often due to spelling issues (e.g., “Suneeta” vs. “Sunita”) or the use of common nicknames or truncations.
c) When mismatched birth dates were more than six months apart, records were cross-checked with the tracking and roster questionnaires, which are separate instruments typically administered to the male household head (not to the EW). We did this exercise even if the names were the same across two waves, since these could be common nicknames applied to two different children across waves. The tracking questionnaire was administered in wave 2 and sought to verify the location of members in wave 1. The roster instrument was administered in both waves and recorded all current household members and their relationships. As an example of how these data were used, take the case of an EW who reported in wave 1 that Raju was born to her in May 2002 and reported in wave 2 that Rajesh was born to her in May 2003 but did not mention Raju. The wave 2 tracking file shows that Raju is still in that household. Further, the wave 2 roster shows that Rajesh was born in May 2002, implying that Raju and Rajesh are the same person. Raju would therefore not be classified as omitted. (Birth records clearly suffer from other measurement errors, but this is not the focus of this paper.)
d) For more complicated cases, we called up the household and confirmed the details.
e) Finally, we dropped cases of adopted children, those of the husband’s first marriage, or children whose mothers were listed in the household roster as other women from the same household.

A schematic representation of our data is represented in Figure 2. A total of 88,363 births were recorded in wave 2 (quadrants 1 and 4 in Figure 2) from panel EWs – that is, EWs who were interviewed in both waves. Of these births, 17,868 births were those that were not reported in wave 1. These births would be a mix of those that occurred between the two waves (and so would not be omitted births) and those that occurred before wave 1 but were not reported in wave 1 (these would be omitted births). We could have used ages reported in wave 2 to identify just the omitted births among quadrant 2 births, but given the uncertainties surrounding birth data in India noted earlier, conclusions regarding omissions would have been contaminated with errors in age data. On the other hand, births not reported in wave 2 but reported in wave 1 (quadrant 2 births in Figure 2) are those we are certain were omitted among all wave 1 reported births (quadrant 1 and 2 births in Figure 2). We therefore focus on quadrant 1 and 2 births for our analysis, with an associated omission rate of 4% [2913/(2913 + 70495)].

Figure 2: Distribution of IHDS birth history panel data across waves by omission status



Note: Quadrant 3 is shaded gray since we would not know of births omitted in both waves.

We then undertook some exploratory analyses to inform our formal modeling. This includes Figure 3, which looks at the share of omissions by mortality and sex. Most of the children omitted in wave 2 were those who had already died by the time the wave 1 interview took place. These dead children represent 73% of all omitted children and are approximately evenly distributed between the two sexes. Among these deaths, 37% were neonatal deaths (as reported by the mother in wave 1), 34% were infants between a month and a year old, and the remaining 29% were children more than a year old. (These estimates are not shown in Figure 3.)

The remaining 27% of the children missing in wave 2 were those reported alive in wave 1. Almost all of them did not live in the respondent's household in wave 2, which is known since except for nine children, none of them were listed in the household roster. So either these omitted children died between wave 1 and wave 2 or were no longer living with the EW. The latter scenario is more likely on average since approximately three-quarters of these alive-in-wave 1 children were female. According to marriage practices in India, daughters who marry and move out are commonly viewed as belonging to other households (Desai et al. 2010); this likely caused such children to be omitted.

Figure 3: Sex and survival status of births reported in wave 1 (2004–2005) but not reported in wave 2 (2011–2012) based on 2,913 missing births (4% of all reported wave 1 births)

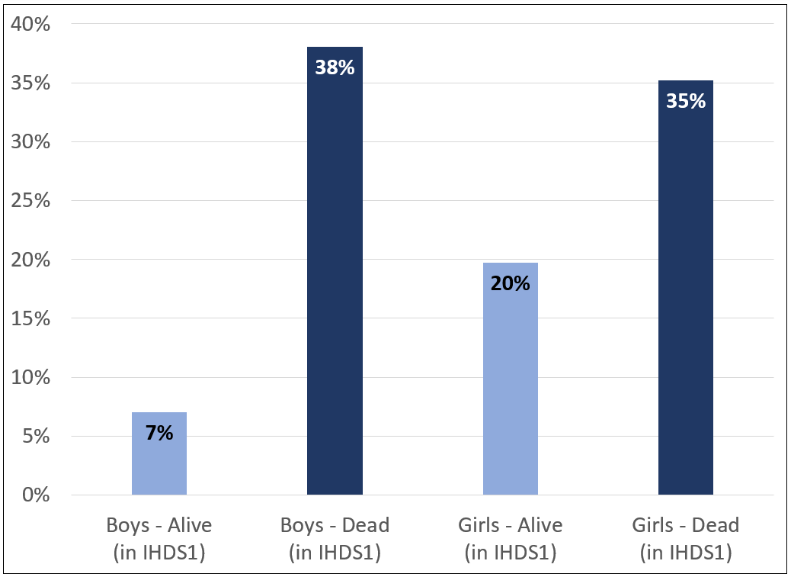


Table 1 displays counts of births by mortality status recorded in wave 1 and omissions in wave 2. The omission prevalence among children reported dead in wave 1 is 31.2% (2,113/6,762) compared to only 1.2% (800/66,646) among those reported alive in wave 1. On average, the relative risk for a child reported dead in wave 1 being omitted in wave 2 is 26 times more than that for a child reported alive in wave 1.

Table 1: Reported mortality of children in wave 1 and their omissions in wave 2

	# births	# births omitted in wave 2	Omitted %
Reported alive in wave 1	66,646	800	1.2
Reported dead in wave 1	6,762	2,113	31.2
Total	73,408	2,913	3.9

5. Analytical approach

5.1 Statistical model

Our exploratory analyses show that a key variable in explaining omissions is likely to be a child's mortality reported in wave 1 and that the factors impacting omissions of alive and dead children are likely to be different. To formalize this, we first fit a model predicting omissions in our full study sample, including child mortality as an explanatory variable. Since the mortality variable remains important in predicting omissions after adjusting for other child, maternal, household, and interviewer characteristics, we fit separate models – the alive-omitted and dead-omitted models – to the sample stratified by mortality status. Doing so gives us flexibility in modeling these separate subpopulations. Our general model specification is as follows:

$$\log\left(\frac{p_{ijklm}}{1-p_{ijlklm}}\right) = \beta_0 + u_{0i} + u_{oj} + \boldsymbol{R}^T_{ijklm}\boldsymbol{\beta}_R + \boldsymbol{W}^T_{ijkl}\boldsymbol{\beta}_W + \boldsymbol{H}^T_{ijk}\boldsymbol{\beta}_H + \boldsymbol{Z}^T_{ij}\boldsymbol{\beta}_Z$$

$$\left(u_{0i}, u_{0j}\right) \sim N(\boldsymbol{0}, \boldsymbol{D}), \text{ where } \boldsymbol{D} = \begin{bmatrix} \sigma^2_{state} & 0 \\ 0 & \sigma^2_{iwer} \end{bmatrix}$$

The outcome variable is a Bernoulli variable, $\boldsymbol{y_{ijklm}} \sim \boldsymbol{BER}\left(\boldsymbol{p_{ijklm}}\right)$, equal to 1 when the child ($\boldsymbol{m}$) was reported as alive in wave 1 and omitted in wave 2 for the alive-omitted model (or reported dead in wave 1 and omitted in wave 2 for the dead-omitted model) and equal to 0 when the child was not omitted in wave 2. We account for the clustering of cases (i.e., children, $\boldsymbol{m}$) within interviewer ($\boldsymbol{j}$) and state ($\boldsymbol{i}$) by introducing state and interviewer random effects. We ignore the clustering of children ($\boldsymbol{m}$) within mothers ($\boldsymbol{l}$) due to estimation issues owing to small cluster sizes. We also ignore the clustering of mothers ($\boldsymbol{l}$) within households ($\boldsymbol{k}$), since a relatively small fraction of homes have multiple EWs. We investigate the omission of births as a function of four vectors: child-level variables ($\boldsymbol{R}$), mother-level variables ($\boldsymbol{W}$), household variables ($\boldsymbol{H}$), and interviewer-level variables ($\boldsymbol{Z}$). We also include cross-level interactions in the model (not shown in the above specification for simplicity of expression).

5.2 Covariates

Based on our response process framework and exploratory analysis, we use the following variables in the covariate vectors:

5.2.1 Child characteristics

a) Sex of the omitted child. ("Male" was used as the reference level.)
b) Age of the child in wave 2. For the omitted children, this is the purported wave 2 age computed from wave 1 data. For dead children, a larger magnitude of this variable would mean that the birth occurred at a more distant time. Past research on cohabitation and fertility histories (Hayford and Morgan 2008; Wu, Martin, and Long 2001) shows that distance from an event reduces the quality of recall. In the case of children recorded alive in wave 1, it is easier for respondents to forget to mention older children who have left home (Marckwardt 1973).
c) Whether the child was an infant when they died (applicable only to the dead-omitted model).

5.2.2 Mother (EW) characteristics

a) The number of years of formal education (mean: 4.3 years; standard deviation: 4.7 years). Education is often used as a proxy for the respondent's cognitive sophistication (Krosnick 1991), and less educated respondents have been found to be more difficult to interview (Loosveldt 1997).
b) Age at wave 2 (mean: 40.4 years; standard deviation: 8.1 years). In a study using the British General Household Survey, Murphy (2009) found evidence of a higher level of deliberate misreporting of birth histories among older women compared with younger women. The plausible explanation was that younger respondents have less scope to misreport since they have children present in the household.
c) Any miscarriages or stillbirths experienced. Questions on miscarriages and stillbirths are sensitive. We hypothesize that if respondents are willing to report such stigmatized events (Haws et al. 2010), the psychological barrier to reporting dead children would be reduced.

At the end of the IHDS interview, interviewers were asked to make observations on the interview process. We include three such EW observation variables:

d) Understanding the purpose of the interview, an indicator of whether the interviewer found it difficult to communicate the survey's purpose at the beginning of the interview. Interviewers reported that EWs had at least some difficulty in understanding the purpose of the interview in 11.5% of interviews.

e) Understanding interview questions, an indicator of whether the interviewer felt that the EW understood questions (or not) in general (i.e., questions not specifically about birth history). Interviewers reported that the EW had at least some difficulty in understanding the questions in 15.2% of interviews.
f) Response confidence, an indicator of whether the interviewer felt that the EW was confident (or not) in her responses in general. In approximately 5.5% of interviews, interviewers reported that the EW was "rarely confident" (in which case, this variable possibly showed a lack of confidence in answering the birth history questions as well). In a further 17% of interviews, EWs were reported to be "sometimes confident." Respondents in the remaining 77.5% of interviews were reported to be "usually confident" (reference level).

5.2.3 Household characteristics

a) Urban residence (urban households, the reference level, 33.2%; rural households, 66.8%).
b) Caste-religious identity. Caste often forms a proxy for social class and quality of education in India (Desai et al. 2010). Religion has also become a focal point in population discourse (Basu 2004) and may lead to biased reporting of fertility. This variable has six categories: Hindu upper caste (the reference level; 22.3% of households), other backward caste (OBC; 33.9% of households), Dalit (scheduled castes; 21.5% of households), Muslim (11.3% of households), Adivasi (scheduled tribes; 8.2% of households), other religion (2.9% of households). The Adivasi community has a higher level of infant mortality than the rest of the population even after controlling for wealth (Das, Kapoor, and Nikitin 2010; Sahu et al. 2015).
c) Asset ownership at wave 2. IHDS provides a derived variable in the public use dataset that classifies a household in one of five quintiles based on asset ownership: Q1 (poorest, 17% of our analytic sample), Q2 (16% of our sample), Q3 (25% of our sample), Q4 (21% of our sample), and Q5 (richest, the reference category, 21% of our sample). We use this variable to indicate the household's economic status; birth registration rates are known to be correlated with economic status (Mohanty and Gebremedhin 2018), and some of the same processes may be at play here.
d) A variable indicating whether multiple EWs were interviewed in the home in wave 2 (18% of households have more than one EW). All things being equal, for the alive-omitted model, we expect to see multiple EWs associated with more omissions since the interviewer may be pressured for time. For the dead-

omitted model, an added factor is that the respondent may not want to talk about her dead children in the likely presence of other EWs.

5.2.4 Interview and interviewer characteristics at wave 2

a) Workload (mean, 55 households; standard deviation, 47.5 households). While IHDS interviewers did not know how many interviews they were going to do in advance, poorly motivated interviewers tended to drop out earlier than their more motivated counterparts. Also, field agencies tended to use good interviewers more. Thus the workload variable can be seen as an indicator of better-quality interviewers.
b) Interview sequence. This is a sequential serial number representing the order in which the interview occurred within an interviewer's workload. We hypothesize that later interviews within an interviewer's workload will see more omitted births, as interviewers might experience more fatigue and start taking shortcuts.
c) Missing age at death. For dead children, respondents were asked about the age at which the child died (the last four columns in the question screenshot in Appendix A). We created an indicator variable that took a value of 1 for an interviewer who had at least one interview with a missing age-at-death value (27% of all interviewers). We used this as an indicator of interviewers who were wary of probing or asking for sensitive information from respondents.

All continuous variables were used in their scaled form (i.e., each value was divided by the standard deviation of the raw values after subtracting the mean of the raw values from it) to improve estimation stability and comparability of effect sizes. We do not include survey weights in our models; we are not interested in population-level inferences on missing children as much as the demographic and interviewing factors associated with omitted births. (However, state of residence and urban residence, the two primary axes of sample stratification in the IHDS, are included in the model.) We account for multiple comparisons in our inferences by adjusting p-values and nominal 90% confidence intervals using the Benjamini-Hochberg (B-H) false discovery rate (FDR) method (Benjamini and Hochberg 1995; Benjamini and Yekutieli 2005). Analysis was conducted in R software (R Core Team 2021) using the lme4 (Bates et al. 2015) package.

5.3 Subset models and AUC curve analysis

Apart from studying the effect sizes for each variable, we also wanted to get a sense of how different sets of variables perform in predicting omissions. For this, we fit separate sub-models that included (a) only the child variables, (b) only the mother variables, and (c) only the household and interviewer variables, and we obtained predicted probabilities of omissions for all cases in our dataset when the subset models were applied. This was done for both the alive-omitted and dead-omitted models.

6. Results

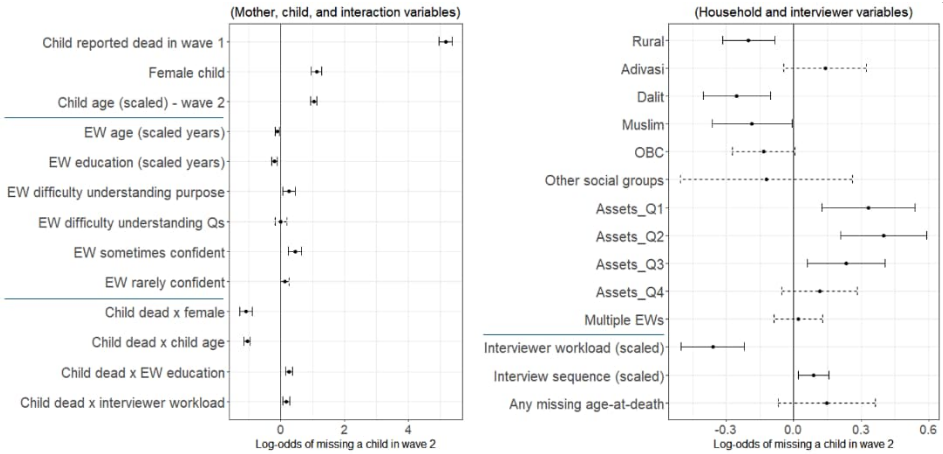
6.1 Results for the initial full sample model

Log odds of omitting a child in wave 2 (who was reported in wave 1) are plotted in Figure 4. A table of coefficient estimates, their standard errors, and B-H–adjusted p-values is in Appendix B. Figure 4 shows that the variable with the largest impact on omissions, by far, is the child's mortality status in wave 1. This variable enters into several interactions with other variables, and the mechanisms underlying omissions likely differ based on mortality. A check also shows that 95.7% of all households with at least one omitted child either omit only children reported alive in wave 1 or omit only children reported dead in wave 1. We therefore fit separate models for wave 2 omissions among children reported alive (alive-omitted model) or dead (dead-omitted model) in wave 1.

6.2 Results for the alive-omitted model

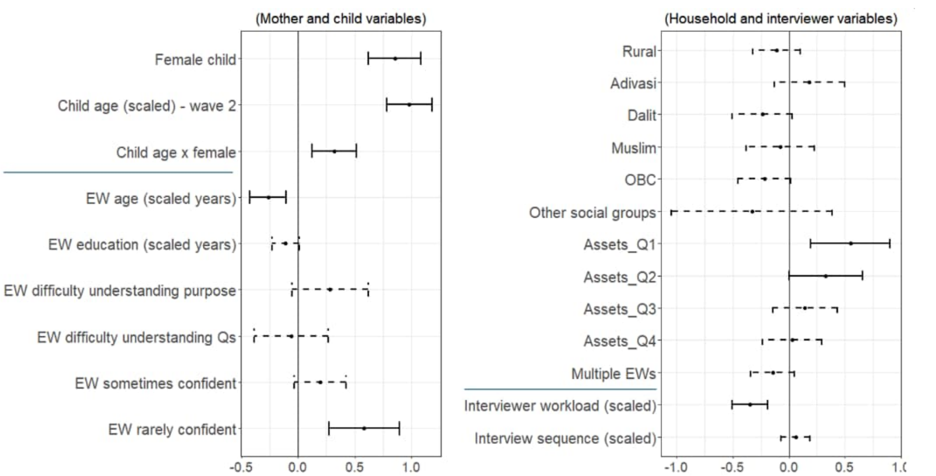
Log odds of a child who was reported as alive in wave 1 being omitted in wave 2 are plotted in Figure 5. A table of coefficient estimates, their standard errors, and B-H–adjusted p-values is in Appendix C.

Figure 4: Log odds (dots) of omitting a child along with 90% B-H confidence intervals



Notes: Solid lines correspond to the confidence intervals that do not contain **0.** The child, mother, and interaction variables are plotted in the left panel, and the household and interviewer variables are plotted in the right panel. (The separator line among the vertical axis labels differentiates these variable groups.) Note the difference in horizontal axis scales between the two panels; n = 72,678 children.

Figure 5: Log odds (dots) of a child who was reported as alive in wave 1 being omitted in wave 2 (alive-omitted model)

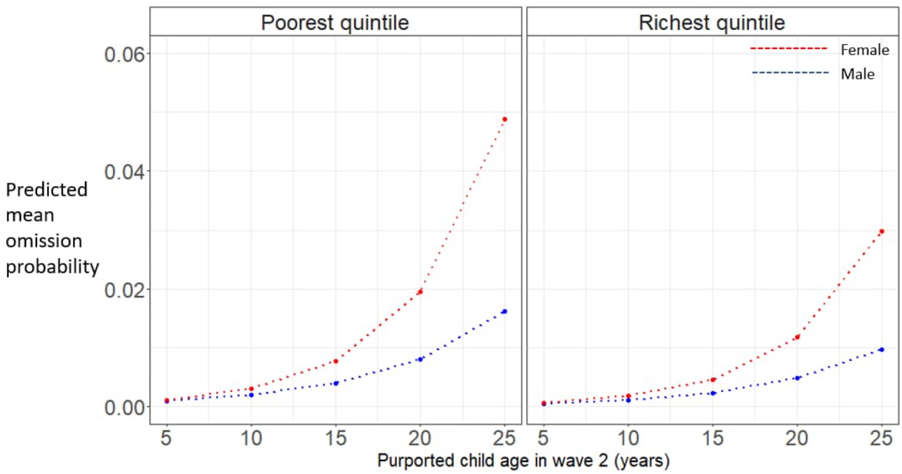


Notes: The 90% B-H–adjusted confidence intervals are shown via horizontal bars; solid lines are those intervals that do not include 0. The child and mother variables are plotted in the left panel, and the household and interviewer variables are plotted in the right panel. (The separator line among the vertical axis labels differentiates these variable groups); n = 66,314 children.

6.2.1 Child variables

Even after adjusting for household, mother, and interviewer variables, the two child-level variables (sex and age, the first three terms in the left panel of Figure 5) are the most impactful variables in predicting omissions. To aid our interpretation, we plot the omission probabilities for different values of these variables in Figure 6. We compute these probabilities separately for children in the poorest and richest households (from the right panel of Figure 5, the poorest quintile households have 73% greater odds of omission [$e^{0.55}$] compared to the reference group of the richest quintile households), all else being equal.

Figure 6: Marginal effect of a child's sex and age on omissions in terms of mean predicted probabilities (vertical axis)



Notes: The child's age (horizontal axis) is the purported age in wave 2 based on wave 1 information. Red lines represent females and blue lines represent males. The panels on the left and right are those of households in the poorest and richest asset quintiles, respectively. For these predictions we assumed a 45-year-old EW respondent with a primary school education in a rural Hindu forward-caste household, with whom the interviewer had no trouble communicating the purpose of the interview, who was usually confident in her responses, and who was the 50th interview among 100 interviews conducted by the interviewer. We did not condition on the random effects in conducting these predictions.

For either the poorest households (left panel of Figure 6) or richest households (right panel of Figure 6), mean omission probabilities for female and male children are similar at younger ages. However, the risk of being omitted rises dramatically for females (especially for children of poorer households) compared to males after 20 years of age. This is approximately the age when a large proportion of females get married (Singh, Shekhar, and Shri 2023) and move out of the natal home. The mean omission probability for a 25-year-old female in the richest quintile household is greater than the mean omission probability for a male of the same age in the poorest quintile household.

6.2.2 Maternal variables

Older respondents (left panel of Figure 5) are associated with lower odds of omission. One possible reason for this is that with younger respondents, the interviewer may just note the children they see in the household. On the other hand, for older women where there are no children around, the birth history question may be asked more methodically, thereby leading to fewer omissions. Of the three interviewer observation variables, omissions are associated with a visible and frequent lack of confidence by respondents in answering questions; the other observation variables do not have any association with omissions.

6.2.3 Household variables

All else being equal, except for the asset quintile variable (discussed above), none of the household variables (right panel of Figure 5) have an association with omissions.

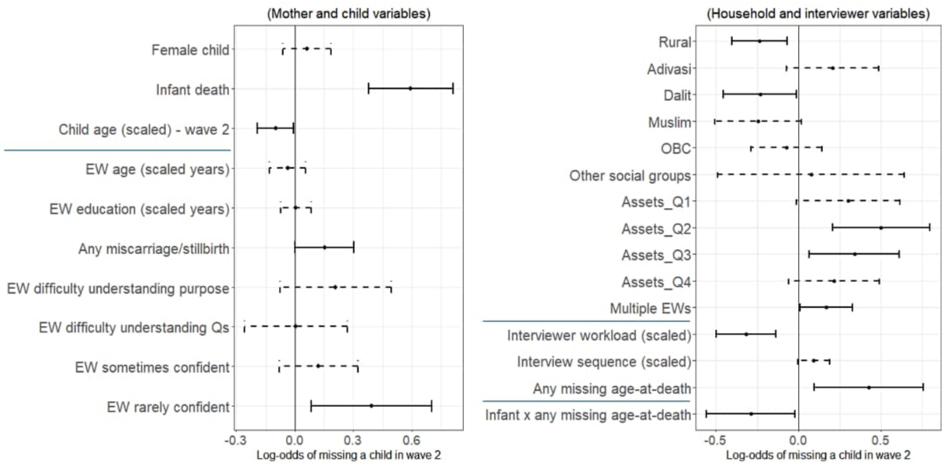
6.2.4 Interview and interviewer variables

While the specific sequential position of the interview in the interviewer's workload does not seem to be associated with omissions, workload magnitude is associated with omissions. For every 1 standard deviation increase in the interviewer's workload (47 interviews), the odds of a child being omitted in the birth history reduces by 29% ($1 - e^{-0.35}$), which aligns with our hypothesis. To evaluate the impact of interviewer workload, we fit a model with the random state and random interviewer effects along with household-level variables and the EW's age, education, and number of children alive. (These fixed effects are meant to adjust for diverse respondents among interviewers.) This yielded a between-interviewer variance estimate of 0.628. If we add the workload variable to this model, the variance reduces to 0.454, implying that this variable alone explains 28% of the between-interviewer variance of omissions in the alive-omitted model.

6.3 Results for the dead-omitted model

The results of the dead-omitted model (children who were reported dead in wave 1 and omitted in the birth history in wave 2) are plotted in Figure 7. A table of coefficient estimates, their standard errors, and B-H–adjusted p-values is in Appendix D.

Figure 7: Log odds (dots) of a child who was reported dead in wave 1 being omitted in wave 2 (dead-omitted model)



Notes: The 90% B-H–adjusted confidence intervals are shown via horizontal bars; solid lines are those intervals that do not include 0. The child and mother variables are plotted in the left panel, and the household and interviewer variables are plotted in the right panel. (The separator line among the vertical axis labels differentiates these variable groups); n = 6,364 children.

6.3.1 Child variables

Children who died as infants (left panel of Figure 7) have 81% higher odds ($e^{0.59}$) of being omitted compared to children who survived infancy. Death in infancy is the variable with the largest impact on omissions. The sex of the child does not matter for omissions (left panel of Figure 7), but older children are at a lower risk of being omitted compared to younger children, all else being equal.

6.3.2 Maternal variables

While the respondent's age and education are not associated with omissions (left panel of Figure 7), if the respondent reported ever having a miscarriage or a stillbirth, it increases the odds of omitting a child in the birth history by 16% ($e^{0.15}$) compared to a respondent who did not report having such an event, all other variable values being the same. Respondents who are observed to be rarely confident in their responses have a much higher odds of omissions (48%, $e^{0.39}$) compared to respondents noted to be usually confident.

6.3.3 Household variables

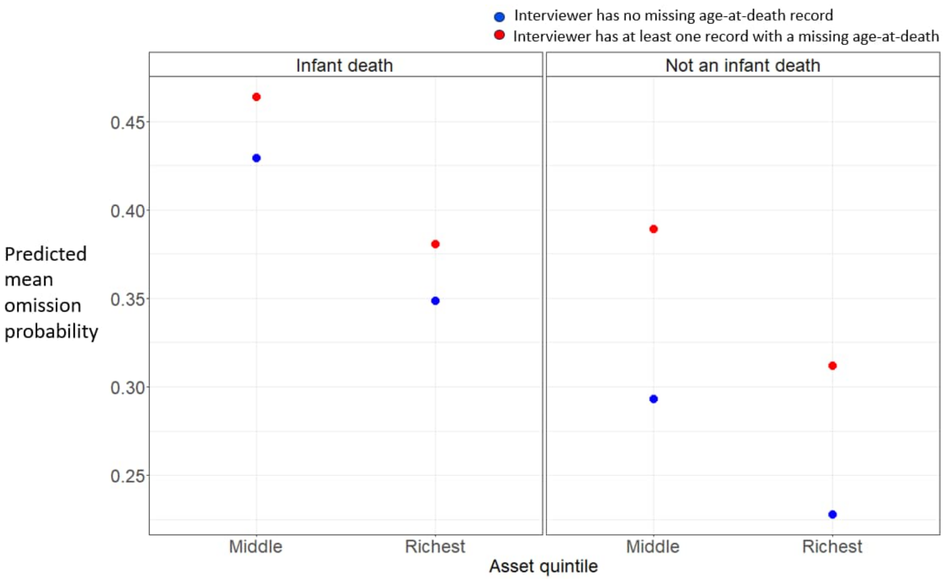
Among the household variables (right panel of Figure 7), rural households have a 21% ($1 - e^{-0.24}$) lower odds of omissions compared to urban households and even after adjusting for this variable, Dalit households have 20% lower odds of omissions ($1 - e^{-0.24}$) compared to forward-caste Hindus. In general, poorer households have higher odds of omitting dead children compared to the richest quintile. One reason for this could be that poorer households are those with less access to good-quality health care and therefore are subject to higher mortality rates compared to richer households (Asaria et al. 2019); the higher prevalence of dead children exposes these households to a greater risk of omissions.

6.3.4 Interview and interviewer variables

Interviews conducted by interviewers with higher workloads are associated with fewer omissions, but the interview sequence does not seem consequential for omissions. Since the missing age-at-death variable has an interaction with the infant variable, we plot predicted mean probabilities involving these two variables along with the household's economic status in Figure 8.

Focusing on the middle asset quintile households, there is a 0.46 mean probability of omitting an infant in the birth history when the interviewer is someone who has at least one missing age-at-death record (leftmost red dot in Figure 8). When the interviewer is someone who has no missing age-at-death record, the mean omission probability for infants drops marginally to 0.43 (leftmost blue dot in Figure 8). But this difference in omission probabilities is exacerbated in the case of non-infant deaths (right panel of Figure 8); the mean chance of a non-infant birth being missed is 10 percentage points more when the interviewer has a missing age-at-death record (mean omission probability of 0.39) compared to when the interviewer has no missing age-at-death record (mean omission probability of 0.29). All else being equal, middle asset quintile households have an approximately 7 percentage point higher omission chance compared to the richest households.

Figure 8: The marginal effect of infant deaths (left panel) or non-infant deaths (right panel) and household economic status in terms of asset quintile (horizontal axis) on mean predicted omission probabilities (vertical axis)



Notes: Red dots represent cases where the interview was conducted by an interviewer with at least one missing age-at-death record, and the blue dots are those where the interviewer had no missing age-at-death records. For these predictions, we assumed a 45-year-old EW respondent with a primary school education in a rural Hindu forward-caste household, with whom the interviewer had no trouble communicating the purpose of the interview, who was usually confident of her responses, and who was the 50th interview among 100 interviews conducted by the interviewer. We did not condition on the random effects in conducting these predictions.

6.4 Subset models

The detailed results of the subset models and AUC analyses are given in Appendix E. They show that for the alive-omitted model, the child-level variables are more effective than the household and maternal variables in predicting omissions. The full alive-omitted model that uses all the variables suggests good predictive power. In contrast, in predicting omissions among dead children, the model with only the household and interviewer variables (which also included the interaction between the infant and the missing age-at-death variable) is more effective than the models using only the child or mother-level variables. The full dead-omitted model does not have strong predictive properties.

7. Discussion

The utility of demographic data can be seriously compromised when respondents do not report important events (Lindberg et al. 2020). This paper focuses on omissions in fertility history data and demonstrates that: (a) the prevalence of omissions is not ignorable; (b) omissions are not random but, as hypothesized by our response process framework, are associated with child, maternal, household, and interviewer factors; and (c) omission of births itself is not a homogenous outcome; mechanisms impacting omissions of children reported alive in wave 1 differ from mechanisms impacting omissions of children reported dead in wave 1.

Omissions in the number of reported births lead to underestimating fertility rates, and as our results show, these biases are not random but affect some groups more than others. We found that undercounts impact girls more than boys, thereby biasing sex ratio estimates, which may well have implications for estimates of missing girls (e.g., Klasen and Wink 2002). Since we find that dead children are more prone to omissions, we are also faced with artificially deflated infant mortality estimates. While this paper focuses on estimating the prevalence of birth omissions and exploring the underlying mechanisms, quantifying the impact of omissions on demographic estimates is a topic of future research.

The analysis indicates that omissions of alive children are not associated with any specific socioreligious group but are likely a combination of respondents' misunderstanding of the birth history question (reporting only current household members) and lack of interviewer probing. We see signs of possible misunderstanding in the case of dead children too, since omissions are associated with reporting a miscarriage or stillbirth by the mother. In the IHDS, questions asking if the respondent experienced a stillbirth or miscarriage (and if so, how many) appear before the fertility history question. Since past research shows that the concepts of stillbirth and neonatal mortality are often confused (Liu et al. 2016; Helleringer 2020), the respondent might think she has already accounted for neonatal deaths in previous questions and fail to report them in the birth history. Interviewers should be trained to be more careful when they encounter a report of a miscarriage or stillbirth by probing and emphasizing to the respondent when administering the birth history questions that stillbirths are not the same as live births.

For both the alive and dead children, interviewer observations about respondents' lack of confidence in their answers are associated with omissions. This lack of confidence could stem from such respondents not understanding what is expected from them and/or from issues such as social undesirability. On the latter aspect, we see that the odds of omitting dead children increase when multiple EWs are interviewed in the household; the reporting of dead children when others are around can be especially unpleasant. We also see that after adjusting for economic status, urban households have a higher risk of omitting dead children compared to rural households; fertility rates in urban India are lower than those in rural India, which could cause the reporting of a dead child to be even more sensitive and result in respondent hesitation. The results suggest that interviewers are sensitive to such cues; they must be trained to act on them by neutrally encouraging respondents to provide more accurate answers.

To curtail omissions, survey instruments can do the following: (a) program the survey instrument to provide alerts to interviewers when respondents who are more prone to omissions (e.g., EWs in low-income households) are administered the fertility history question; (b) include an instruction below the fertility history question that tells the interviewer to specifically probe the respondent for married-out daughters and children who died in infancy. A special instruction can appear when miscarriages or stillbirths are reported earlier by the respondent so that the interviewer can clarify the difference between a miscarriage or stillbirth and a live birth. The strongest general method to reduce the omission prevalence is by focusing on dead children (73% of all omitted births are those of dead children), especially those who died during infancy (71% of omitted dead children were infants when they died).

Our results bring attention to the role that interviewers play in reducing omissions. Our experience of several interviewer trainings across surveys in the developing world shows that interviewers, who are almost always contractual workers with different field agencies, for whom interviewing might be a part-time job, are often not equipped with interviewing skills such as sensitive probing. This increases the likelihood of missing dead children in the fertility history question. This is probably why we see the missing age-at-death variable (a sensitive question) associated with the omissions of dead children (also sensitive). When fieldwork is in progress, survey managers will not know if births are being missed or not. However, the presence of missing age-at-death information in fertility histories collected by an interviewer can signal that the interviewer is hesitant to ask about dead children and needs to be retrained in this regard. Thus systems can be set up to track this indicator and interviews by such interviewers can be scrutinized. Interviewers must also be cautioned against simply copying over the roster information (one of the shortcuts interviewers can use) into the fertility history, since that will miss nonresident alive children in addition to missing dead children.

Our results showed that interviewers with a higher workload are associated with a lower chance of omission. This perhaps reflects that interviewer longevity is associated with enjoyment of the interview process and with data collection organizations making greater efforts to retain better interviewers. However, good interviewers are also prone to fatigue, which could impact omissions; quality-control systems should monitor paradata (Couper 1998) on a rapid basis (Sharma 2019) to indicate if interviewers are speeding through the fertility section. Future researchers can listen to electronic recordings of interactions between the interviewer and respondents for better insight into causal mechanisms underlying omissions. Cognitive interviews (Willis 2005) can also be undertaken.

In conclusion, the data suggest that individuals revise their demographic histories in response to survey situations in a way that may not match the scientific questions of interest to the research community. Unless researchers understand the forces behind these revisions, they risk underestimating fertility and mortality and overestimating the number of missing girls. The results suggest a need for caution as survey programs such as the DHS move from collecting fertility histories to collecting pregnancy histories. Similar to the omission of dead children from fertility histories, pregnancy histories may also suffer from under-enumeration associated with miscarriages and pregnancy losses.

8. Acknowledgments

We thank the two anonymous reviewers and the associate editor for their comments, which greatly improved the paper. Research reported in this paper was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health, award number R01HD041455, and the Bill and Melinda Gates Foundation, award numbers INV-009903 and INV-048560.

References

Aizpurua, E., Heiden, E.O., Park, K.H., Wittrock, J., and Losch, M.E. (2018). Investigating respondent multitasking and distraction using self-reports and interviewers' observations in a dual-frame telephone survey. *Survey Methods: Insights from the Field* 11(2): 1–10. doi:10.29115/SP-2018-0025.

Arnold, F., Choe, M.K., and Roy, T.K. (1998). Son preference, the family-building process and child mortality in India. *Population Studies* 52(3): 301–315 doi:10.1080/0032472031000150486.

Asaria, M., Mazumdar, S., Chowdhury, S., Mazumdar, P., Mukhopadhyay, A., and Gupta, I. (2019). Socioeconomic inequality in life expectancy in India. *BMJ Global Health* 4(3): e001445. doi:10.1136/bmjgh-2019-001445.

Bardley, S. (2016). When quality matters: Linking the reliability of demographic and health survey data to biases in international mortality, fertility, and family planning estimates [Ph.D. thesis]. Berkeley: University of California Berkeley. Retrieved from: https://escholarship.org/uc/item/5j33z8pf.

Basu, A.M. (2004). The squabble that never ends: Religion and fertility. *Economic and Political Weekly* 39(39): 4294–4296.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). doi:10.18637/jss.v067.i01.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57(1): 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x.

Benjamini, Y. and Yekutieli, D. (2005). False discovery rate: Adjusted multiple confidence intervals for selected parameters. *Journal of the American Statistical Association* 100(469): 71–81. doi:10.1198/016214504000001907.

Billiet, J. and Loosveldt, G. (1988). Improvement of the quality of responses to factual survey questions by interviewer training. *Public Opinion Quarterly* 52(2): 190–211. doi:10.1086/269094.

Blacker, J. (1994). Some thoughts on the evidence of fertility decline in Eastern and Southern Africa. *Population and Development Review* 20(1): 200. doi:10.2307/2137639.

Borkotoky, K. and Unisa, S. (2014). Indicators to examine quality of large scale survey data: An example through district level household and facility survey. *PLoS ONE* 9(3): e90113. doi:10.1371/journal.pone.0090113.

Cannell, C.F. and Fowler, F.J. (1964). A note on interviewer effect in self-enumerative procedures. *American Sociological Review* 29(2): 270. doi:10.2307/2092130.

Choi, Y., Li, Q., and Zachary, B. (2018). Measuring fertility through mobile-phone based household surveys: Methods, data quality, and lessons learned from PMA2020 surveys. *Demographic Research* 38(55): 1663–1698. doi:10.4054/DemRes.2018.38.55.

Clark, S. (2000). Son preference and sex composition of children: Evidence from India. *Demography* 37(1): 95–108. doi:10.2307/2648099.

Couper, M.P. (1998). Ensuring survey quality in a CASIC environment. In: *JSM Proceedings*, Survey Research Methods Section. Alexandria, VA: American Statistical Association: 41–49. http://www.asasrms.org/Proceedings/papers/1998_006.pdf.

Das, M.B., Kapoor, S., and Nikitin, D. (2010). A closer look at child mortality among Adivasis in India. (World Bank Policy Research Working Paper 5231). Washington, DC: World Bank. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1565992.

DeMatteis, J., Young, L., Dahlhamer, J., Langley, R., Murphy, R., Kristen, O., and Sharma, S. (2020). Falsification in surveys. Task force report for the American Association for Public Opinion Research and the American Statistical Association. https://www.aapor.org/wp-content/uploads/2022/11/AAPOR_Data_Falsification_Task_Force_Report-updated.pdf.

Desai, S., Dubey, A., Joshi, B.L., Sen, M., Shariff, A., and Vanneman, R. (2010). *Human development in India: Challenges for a society in transition*. New Delhi, Oxford University Press.

Espeut, D. and Becker, S. (2015). The validity of birth and pregnancy histories in rural Bangladesh. *Journal of Health, Population and Nutrition* 33(1). doi:10.1186/s41043-015-0027-8.

Feeney, G. and Jianhua, Y. (1994). Below replacement fertility in China? A close look at recent evidence. *Population Studies* 48(3): 381–394. doi:10.1080/0032472031000147926.

Githens, P.B., Glass, C.A., Sloan, F.A., and Entman, S.S. (1993). Maternal recall and medical records: an examination of events during pregnancy, childbirth, and early infancy. *Birth* 20(3): 136–141. doi:10.1111/j.1523-536X.1993.tb00438.x.

Gold, K.J., Dalton, V.K., and Schwenk, T.L. (2007). Hospital care for parents after perinatal death. *Obstetrics and Gynecology* 109(5): 1156–1166. doi:10.1097/01.AOG.0000259317.55726.df.

Gold, K.J., Sen, A., and Leon, I. (2018). Whose fault is it anyway? Guilt, blame, and death attribution by mothers after stillbirth or infant death. *Illness, Crisis and Loss* 26(1): 40–57. doi:10.1177/1054137317740800.

Haws, R.A., Mashasi, I., Mrisho, M., Schellenberg, J.A., Darmstadt, G.L., and Winch, P.J. (2010). 'These are not good things for other people to know': How rural Tanzanian women's experiences of pregnancy loss and early neonatal death may impact survey data quality. *Social Science and Medicine* 71(10): 1764–1772. doi:10.1016/j.socscimed.2010.03.051.

Hayford, S.R. and Morgan, S.P. (2008). The quality of retrospective data on Cohabitation. *Demography* 45(1): 129–141. doi:10.1353/dem.2008.0005.

Helleringer, S., Liu, L., Chu, Y., Rodrigues, A., and Fisker, A.B. (2020). Biases in survey estimates of neonatal mortality: Results from a validation study in urban areas of Guinea-Bissau. *Demography* 57(5): 1705–1726. doi:10.1007/s13524-020-00911-6.

Holbrook, A.L., Johnson, T.P., Kapousouz, E., and Cho, Y.I. (2020). Exploring the antecedents and consequences of Interviewer Reading Speed (IRS) at the question level. In: Olson, K., Smyth, J.D., Dykema, J., Holbrook, A.L., Kreuter, F., and West, B.T. (eds.). *Interviewer effects from a total survey error perspective*. New York: Chapman and Hall/CRC: 237–252. doi:10.1201/9781003020219-23.

ICF (2020). Demographic and Health Survey interviewer's manual. Rockville, MD: ICF. https://dhsprogram.com/pubs/pdf/DHSM1/DHS8-Interviewer%27s-Manual-EN-25Sep2020-DHSM1.pdf.

ICF Macro (2009). Training field staff for DHS surveys. Calverton, MD: ICF Macro. https://dhsprogram.com/publications/publication-dhsm3-dhs-questionnaires-and-manuals.cfm.

Japec, L. (2007). Interviewer error and interviewer burden. In: Lepkowski, J.M., Tucker, C., Brick, J.M., de Leeuw, E.D., Japec, L., Lavrakas, P.J., Link, M.W., and Sangster, R.L. (eds.). *Advances in telephone survey methodology*. Hoboken, NJ: John Wiley and Sons: 185–211. doi:10.1002/9780470173404.ch9.

Johnson, T.P. and Van de Vijver, F.J. (2003). Social desirability in cross-cultural research. In: Harkness, J., Van de Vijver, F.J., and Mohler, P. (eds.). *Cross-cultural survey methods*. Hoboken, NJ: Wiley: 195–204.

Klasen, S. and Wink, C. (2002). A turning point in gender bias in mortality? An update on the number of missing women. *Population and Development Review* 28(2): 285–312. doi:10.1111/j.1728-4457.2002.00285.x.

Knäuper, B. (1999). The impact of age and education on response order effects in attitude measurement. *Public Opinion Quarterly* 63(3): 347. doi:10.1086/297724.

Koch, A., Blom, A.G., Stoop, I., and Kappelhof, J.W. (2009). Data collection quality assurance in cross-national surveys: The example of the ESS. *Methoden, Daten, Analysen (MDA)* 3(2): 219–247.

Krosnick, J.A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* 5(3): 213–236. doi:10.1002/acp.2350050305.

Krosnick, J.A. (1999). Survey research. *Annual Review of Psychology* 50(1): 537–567. doi:10.1146/annurev.psych.50.1.537.

Krosnick, J.A. and Alwin, D.F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly* 51(2): 201. doi:10.1086/269029.

Krosnick, J.A., Narayan, S., and Smith, W.R. (1996). Satisficing in surveys: Initial evidence. *New Directions for Evaluation* 1996(70): 29–44. doi:10.1002/ev.1033.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality and Quantity* 47(4): 2025–2047. doi:10.1007/s11135-011-9640-9.

Leone, T., Sochas, L., and Coast, E. (2021). Depends who's asking: Interviewer effects in demographic and health surveys abortion data. *Demography* 58(1): 31–50. doi:10.1215/00703370-8937468.

Lindberg, L., Kost, K., Maddow-Zimet, I., Desai, S., and Zolna, M. (2020). Abortion reporting in the United States: An assessment of three national fertility surveys. *Demography* 57(3): 899–925. doi:10.1007/s13524-020-00886-4.

Liu, L., Kalter, H.D., Chu, Y., Kazmi, N., Koffi, A.K., Amouzou, A., Joos, O., Munos, M., and Black, R.E. (2016). Understanding misclassification between neonatal deaths and stillbirths: Empirical evidence from Malawi. *PLOS ONE* 11(12): e0168743. doi:10.1371/journal.pone.0168743.

Loosveldt, G. (1997). Interaction characteristics of the difficult-to-interview respondent. *International Journal of Public Opinion Research* 9(4): 386–394. doi:10.1093/ijpor/9.4.386.

Mangione, T.W., Fowler, F.J., and Louis, T.A. (1992). Question characteristics and interviewer effects. *Journal of Official Statistics* 8: 293–293.

Marckwardt, A.M. (1973). Evaluation of an experimental short interview form designed to collect fertility data: The case of Peru. *Demography* 10(4): 639–657. doi:10.2307/2060888.

Meekers, D. (1991). The effect of imputation procedures on first birth intervals: Evidence from five African fertility surveys. *Demography* 28(2): 249–260. doi:10.2307/2061278.

Michael, M., King, L., Guo, L., McKee, M., Richardson, E., and Stuckler, D. (2013). The mystery of missing female children in the Caucasus: An analysis of sex ratios by birth order. *International Perspectives on Sexual and Reproductive Health* 39(2): 97–102. doi:10.1363/3909713.

Mohanty, I. and Gebremedhin, T.A. (2018). Maternal autonomy and birth registration in India: Who gets counted? *PLOS ONE* 13(3): e0194095. doi:10.1371/journal.pone.0194095.

Murphy, M. (2009). Where have all the children gone? Women's reports of more childlessness at older ages than when they were younger in a large-scale continuous household survey in Britain. *Population Studies* 63(2): 115–133. doi:10.1080/00324720902917238.

Murphy, S., Shevlin, M., and Elklit, A. (2014). Psychological consequences of pregnancy loss and infant death in a sample of bereaved parents. *Journal of Loss and Trauma* 19(1): 56–69. doi:10.1080/15325024.2012.735531.

Olson, K. and Peytchev, A. (2007). Effect of interviewer experience on interview pace and interviewer attitudes. *Public Opinion Quarterly* 71(2): 273–286. doi:10.1093/poq/nfm007.

Olson, K. and Smyth, J.D. (2020). What do interviewers learn? Changes in interview length and interviewer behaviors over the field period. In: Olson, K., Smyth, J.D., Dykema, J., Holbrook, A.L., Kreuter, F., and West, B.T. (eds.). *Interviewer effects from a total survey error perspective*. New York: Chapman and Hall/CRC: 279–292. doi:10.1201/9781003020219-26.

Pardeshi, G.S. (2010). Age heaping and accuracy of age data collected during a community survey in the Yavatmal district, Maharashtra. *Indian Journal of Community Medicine: Official Publication of Indian Association of Preventive and Social Medicine* 35(3): 391–395. doi:10.4103/0970-0218.69256.

Patel, R., Marbaniang, S. P., Srivastava, S., Kumar, P., and Chauhan, S. (2021). Why women choose to deliver at home in India: A study of prevalence, factors, and socio-economic inequality. *BMC Public Health* 21(1): 1785. doi:10.1186/s12889-021-11779-5.

Pullum, T.W. and Becker, S. (2014). Evidence of omission and displacement in DHS Birth Histories. https://dhsprogram.com/pubs/pdf/MR11/MR11.pdf.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

Robitaille, M.C. and Chatterjee, I. (2018). Sex-selective abortions and infant mortality in India: The role of parents' stated son preference. *The Journal of Development Studies* 54(1): 47–56. doi:10.1080/00220388.2016.1241389.

Sahu, D., Nair, S., Singh, L., Gulati, B.K., and Pandey, A. (2015). Levels, trends and predictors of infant and child mortality among scheduled tribes in rural India. *The Indian Journal of Medical Research* 141(5): 709–719. doi:10.4103/0971-5916.159593.

Schober, M.F. and Conrad, F.G. (1997). Does conversational interviewing reduce survey measurement error? *Public Opinion Quarterly* 61(4): 576–602. doi:10.1086/297818.

Sharma, S. (2019). *Paradata, interviewing quality, and interviewer effects.* [Ph.D. thesis]. Ann Arbor: University of Michigan. https://deepblue.lib.umich.edu/handle/2027.42/150047.

Singh, M., Shekhar, C., and Shri, N. (2023). Patterns in age at first marriage and its determinants in India: A historical perspective of last 30 years (1992–2021). *SSM-Population Health* 22: 101363. doi:10.1016/j.ssmph.2023.101363.

United Nations (1984). The contribution of the World Fertility Survey data to our understanding of fertility levels and trends in selected developing countries. *Population Bulletin of the United Nations* 16: 62–76.

Willis, G. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage Publications. doi:10.4135/9781412983655.

Wu, L.L., Martin, S.P., and Long, D.A. (2001). Comparing data quality of fertility and first sexual intercourse histories. *The Journal of Human Resources* 36(3): 520. doi:10.2307/3069629.

Youden, W.J. (1950). Index for rating diagnostic tests. *Cancer* 3(1): 32–35. doi:10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3.

Appendix A

Figure A-1: Screenshot of the fertility history question in the IHDS questionnaire

20. Fertility History (continued)

Eligible Woman No: 1 EW1No

Now, I would like to talk to you about your live births, whether still alive or not, starting with the first birth you had.
RECORD TWINS ON SEPARATE LINES, BUT CONNECT WITH A BRACKET. WRITE 99 IN 20.2 IF NOT IN HOUSEHOLD ROSTER

20.1	20.2	20.3	20.4	20.5		20.6		20.7	20.8	
BIRTH ID	HOUSEHOLD **ROSTER ID** NA=99	**What name was given to your (first / next) baby?** [START WITH FIRST BIRTH]	**Is [NAME]** a Boy=1 or Girl?=2	**What was the month and year when [NAME] was born?**		IF BIRTH DATE IS NOT KNOWN: **About how old (is / would have been)?**		**Where is [NAME] now?** With Respondent=1 Elsewhere=2 Dead=3	IF DEAD **How old was [NAME] when (he / she) died? <1 month=5555**	
				MONTH	YEAR	YEARS	MONTHS		YEARS	MONTHS
BH1	BH2	BH3NM	BH4	BH5a	BH5b	BH6a	BH6b	BH7	BH8a	BH8b
01										
02										
03										
04										

Appendix B: Summary of the omission model fit to the full sample

Table A-1: Log odds of omission in the wave 2 fertility history of a child reported in wave 1; standard errors (SE) and B-H–adjusted p-values

	Log odds	SE	B-H p-values
Intercept	–5.78	0.19	**< 0.01**
Child-level variables			
1. Dead in wave 1 (ref: alive in wave 1)	5.18	0.11	**< 0.01**
2. Female (ref: male)	1.12	0.09	**< 0.01**
3. Age (actual or purported) at wave 2 (scaled)	1.04	0.05	**< 0.01**
Mother-level variables			
4. Age (at wave 2, scaled)	–0.11	0.04	**0.01**
5. Years of education (scaled)	–0.20	0.05	**< 0.01**
6. Any difficulty conveying IW purpose (ref: no)	0.26	0.11	**0.03**
7. Did respondent have difficulty understanding questions?	0.01	0.10	0.95
8. Was respondent confident? (ref: usually)			
Sometimes	0.13	0.07	0.11
Rarely	0.45	0.11	**< 0.01**
Household (HH) variables			
9. Rural (ref: urban)	-0.20	0.06	**< 0.01**
10. Religion-caste (ref: Hindu, upper caste)			
Adivasi	0.14	0.10	0.21
Dalit	-0.25	0.08	**< 0.01**
Muslim	-0.18	0.10	**0.08**
Other backward castes (OBC)	-0.13	0.08	0.11
Other religions	-0.12	0.21	0.61
11. Asset ownership (ref: Q5, richest)			
Quintile 1 (poorest)	0.33	0.11	**< 0.01**
Quintile 2	0.40	0.11	**< 0.01**
Quintile 3	0.23	0.10	**0.02**
Quintile 4	0.11	0.09	0.25
12. Multiple eligible women in HH	0.02	0.06	0.77
Interviewer-level variables			
13. Interviewer workload (scaled)	–0.36	0.08	**< 0.01**
14. Interview sequence (scaled)	0.09	0.04	**0.03**
15. Missing any age-at-death in wave 2	0.15	0.12	0.25
Interactions			
16. Child reported dead in wave 1 x female child	–1.09	0.11	**< 0.01**
17. Mother’s education (years, scaled) x child reported dead in wave 1	0.26	0.06	**< 0.01**
18. Child reported dead in wave 1 x child's age (scaled; wave 2)	–1.06	0.05	**< 0.01**
19. Interviewer’s workload (scaled) x child reported dead in wave 1	0.18	0.06	**< 0.01**

Note: This model was fitted using 72,678 cases.

Appendix C: Summary of the alive-omitted model

Table A-2: Log odds of omission in the wave 2 fertility history of a child reported alive in wave 1, standard errors (SE) and B-H–adjusted p-values

	Log odds	SE	B-H p-values
Intercept	–5.59	0.19	**< 0.01**
Child-level variables			
1. Female (ref: male)	0.86	0.11	**< 0.01**
2. Age (actual or purported) at wave 2 (scaled)	0.99	0.10	**< 0.01**
3. Female x age (scaled) at wave 2	0.32	0.10	**< 0.01**
Mother-level variables			
4. Age (at wave 2, scaled)	–0.27	0.08	**< 0.01**
5. Years of education (scaled)	–0.11	0.06	0.12
6. Any difficulty conveying IW purpose (ref: no)	0.28	0.16	0.14
7. Did R have difficulty understanding questions?	-0.06	0.16	0.72
8. Was R confident? (ref: usually)			
Sometimes	0.19	0.11	0.14
Rarely	0.58	0.15	**< 0.01**
Household (HH) variables			
9. Rural (ref: urban)	–0.11	0.10	0.38
10. Religion-caste (ref: Hindu, upper caste)			
Adivasi	0.18	0.15	0.32
Dalit	–0.24	0.13	0.12
Muslim	–0.08	0.15	0.67
Other backward castes (OBC)	–0.22	0.11	0.12
Other religions	–0.33	0.35	0.40
11. Asset ownership (ref: Q5, richest)			
Quintile 1 (poorest)	0.55	0.17	**< 0.01**
Quintile 2	0.33	0.16	**0.10**
Quintile 3	0.14	0.14	0.38
Quintile 4	0.03	0.13	0.81
12. Multiple eligible women in HH	–0.14	0.10	0.20
Interviewer-level variables			
13. Interviewer workload (scaled)	–0.35	0.08	**< 0.01**
14. Interview sequence (scaled)	0.06	0.06	0.40

Note: This model was fitted using 66,314 cases.

Appendix D: Summary of the dead-omitted model

Table A-3: Log odds of omission in the wave 2 fertility history of a child reported dead in wave 1, standard errors (SE) and B-H–adjusted p-values.

	Log odds	SE	B-H p-values
Intercept	–1.28	0.23	**< 0.01**
Child-level variables			
1. Infant death	0.59	0.11	**< 0.01**
2. Female (ref: male)	0.06	0.06	0.40
3. Age (actual or purported) at wave 2 (scaled)	–0.10	0.05	0.09
Mother-level variables			
4. Age (at wave 2, scaled)	-0.04	0.05	0.52
5. Years of education (scaled)	0.01	0.04	0.93
6. Any miscarriage/stillbirth	0.15	0.08	0.10
7. Any difficulty conveying IW purpose (ref: no)	0.21	0.15	0.21
8. Did R have difficulty understanding questions?	0.01	0.13	0.96
9. Was R confident? (ref: usually)			
Sometimes	0.12	0.10	0.30
Rarely	0.39	0.16	**0.05**
Household (HH) variables			
10. Rural (ref: urban)	–0.24	0.09	**0.03**
11. Religion-caste (ref: Hindu, upper caste)			
Adivasi	0.21	0.14	0.20
Dalit	–0.23	0.11	**0.09**
Muslim	–0.24	0.13	0.11
Other backward castes (OBC)	–0.07	0.11	0.58
Other religions	0.08	0.29	0.85
12. Asset ownership (ref: Q5, richest)			
Quintile 1 (poorest)	0.30	0.16	0.11
Quintile 2	0.50	0.15	**< 0.01**
Quintile 3	0.34	0.14	**0.05**
Quintile 4	0.22	0.14	0.19
13. Multiple eligible women in HH	0.17	0.08	**0.09**
Interviewer-level variables			
14. Interviewer workload (scaled)	–0.32	0.09	**< 0.01**
15. Interview sequence (scaled)	0.09	0.05	0.11
16. Any missing age-at-death in wave 2	0.43	0.17	**0.05**
Interaction			
17. Infant death x any missing age-at-death in wave 2	–0.29	0.14	**0.09**

Note: This model was fitted using 6,364 cases.

Appendix E: Predictive Power of Variable Subsets

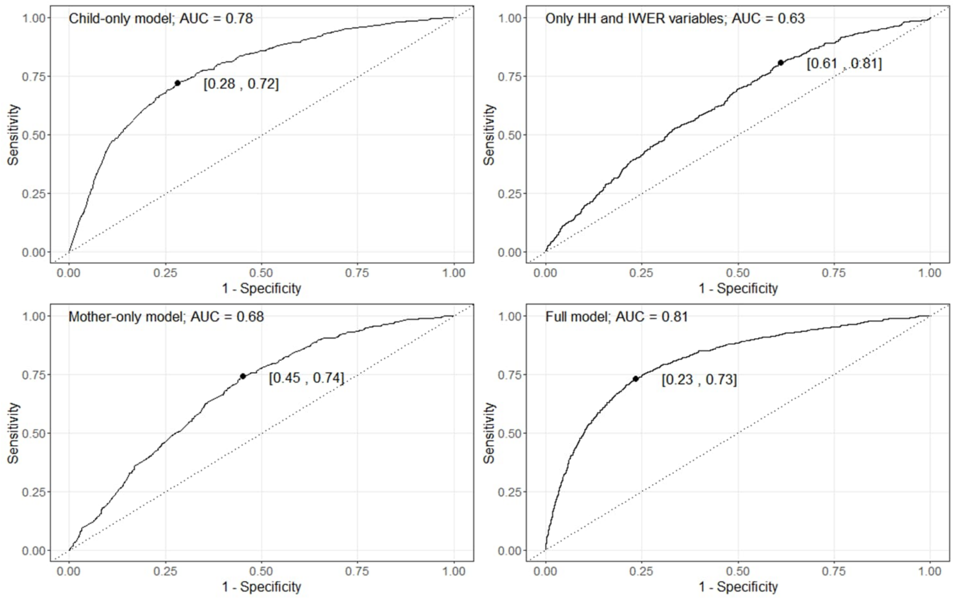
To get a sense of the relative impact of different sets of predictor variables, we fit subset models to both the alive-omitted and dead-omitted models that include: (a) only the child variables, (b) only the mother variables, and (c) only the household and interviewer variables. We obtain predicted probabilities of omissions for all cases in our dataset under each of these subset models as well as for the full models.

We then use different thresholds of predicted probabilities to classify a case as omitted or not omitted, thereby getting measures of true positive rates (TPR, or sensitivity) and false positive rates (FPR, or 1 – specificity). This helps us arrive at the optimal threshold using Youden's method (Youden 1950), which balances the need to increase TPR and reduce FPR. The AUC curves for model subsets of the alive-omitted and dead-omitted models are displayed in Figures A-2 and A-3, respectively.

A comparison of the full models in the figures shows that the AUC for the alive-omitted model (0.81) is quite good and much better than that of the dead-omitted model (AUC of 0.67). That is, it is much more difficult to predict which child will be omitted in the next wave among children reported dead in the current wave compared to predicting omitted children among alive children. The small percentage of omissions among alive children (1.2%) is more sharply identifiable compared to the large percentage of omissions among dead children (31% omissions).

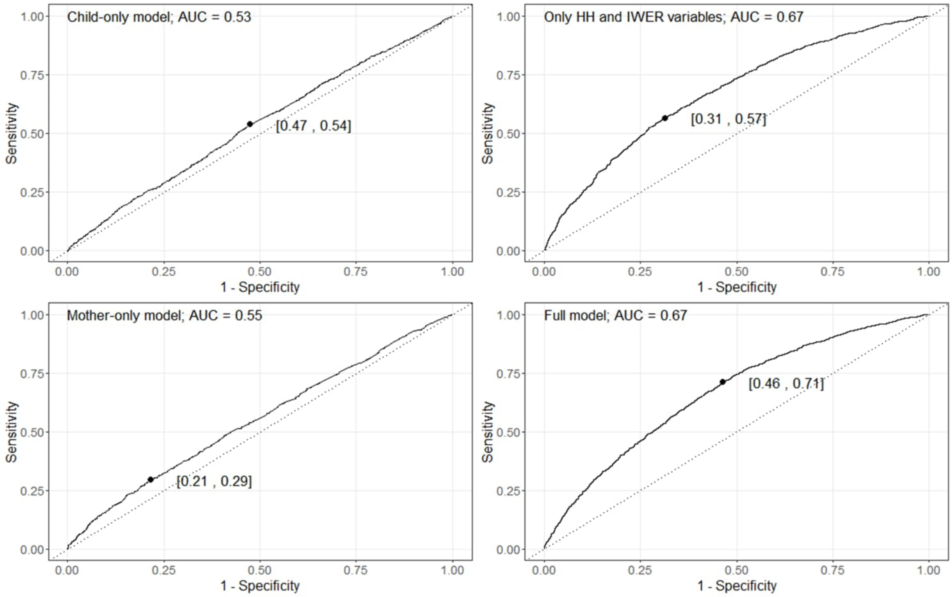
The AUC for the mother and child subset models are better when predicting omissions among alive children compared to predicting omissions among dead children, while the household and interviewer variables perform approximately the same in both situations.

Figure A-2: AUC curves for the alive-omitted models



Notes: The horizontal axis represents 1 – specificity (or the false positive rate), and the vertical axis represents sensitivity (or the true positive rate). The diagonal solid line represents a purely by chance random classifier. The more the departure of the curve from this line, the better our predictions. The dot on the curve is the point at which the optimal threshold probability was used.

Figure A-3: AUC curves for the dead-omitted models



Notes: The horizontal axis represents 1 – specificity (or the false positive rate), and the vertical axis represents sensitivity (or the true positive rate). The diagonal solid line represents a purely by chance random classifier. The more the departure of the curve from this line, the better our predictions. The dot on the curve is the point at which the optimal threshold probability was used. The interaction of the infant variable with the any missing age-at-death variable was included in the household and interviewer model.