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

DEMOGRAPHIC RESEARCH

VOLUME 47, ARTICLE 4, PAGES 73–110
PUBLISHED 12 JULY 2022
https://www.demographic-research.org/Volumes/Vol47/4/
DOI: 10.4054/DemRes.2022.47.4

Research Article

Measuring contraceptive use in India: Implications of recent fieldwork design and implementation of the National Family Health Survey

Kaushalendra Kumar

Abhishek Singh

Amy O. Tsui

Contents

Demographic Research: Volume 47, Article 4
Research Article

Measuring contraceptive use in India: Implications of recent fieldwork design and implementation of the National Family Health Survey

Kaushalendra Kumar[1]

Abhishek Singh[1]

Amy O. Tsui[2]

Abstract

BACKGROUND
India's National Family Health Surveys (NFHS) have provided critical population-level data to inform public policy and research. Although fertility declined, NFHS-4 (2015–2016) reported lower modern contraceptive and female sterilization use compared with NFHS-3 (2005–2006).

OBJECTIVE
This study assesses selected survey design and interviewer factors' influences on respondent reporting of modern contraceptive and female sterilization use.

METHODS
With data on 582,144 married childbearing-aged females, the analysis pursues multivariable logistic models of both outcomes using survey covariates, assesses interviewer deviance residuals, and estimates multi-level cross-classified random intercept models for state, cluster and interviewer effects.

RESULTS
Adjusted odds ratios (AORs) for reporting modern use in NFHS-4 versus NFHS-3 were 1.21 (1.17–1.26) and 1.66 (1.59–1.74) for sterilization. The AOR for each interview month after survey launch was 1.16 (1.15–1.17) for modern use and 1.18 (1.16–1.19) for sterilization. The AOR for respondents interviewed in the first versus second survey phase was 1.35 (1.30–1.40) for modern methods and 1.12 (1.07–1.17) for female sterilization. Interviewer deviance residuals for both contraceptive outcomes were larger in NFHS-4 than NFHS-3. Eliminating problematic interviews raised modern use 2.0% points and sterilization 1.3% points. Larger state, community cluster and interviewer effects were observed for NFHS-4 versus NFHS-3.

[1] International Institute for Population Sciences, Mumbai, India.
[2] Johns Hopkins University, Baltimore, Maryland, USA. Email: atsui1@jhu.edu.

CONCLUSIONS
The five-fold expansion of NFHS-4's sample likely challenged pre-existing survey protocols and may have lowered modern method use by up to 6% points and female sterilization by 2% points.

CONTRIBUTION
The roles of survey fieldwork and interviewers, as sources of measurement error, are important to consider when interpreting change observed in cross-sectional estimates.

1. Background

The National Family Health Survey (NFHS) program in India has been a critical source of population-level data to inform public policy and research for nearly three decades since the first round, NFHS-1 (1992–1993), which was followed by NFHS-2 (2002–2003), NFHS-3 (2005–2006), and NFHS-4 (2015–2016). NFHS data have enabled assessing trends in and determinants of fertility, infant and child mortality, and gender equity, as well as the utilization of contraception, immunization, and other health services. NFHS has also measured improvements in female literacy and child nutritional status, finding more women schooled and fewer children stunted, underweight, or anemic. The government's global commitments are evaluated with NFHS data; for example, Bora and Saikia (2018) gauge district-level trends toward meeting UN Sustainable Development Goals (SDGs) for newborn and under-five child mortality. For contraceptive access, New et al. (2017) examine multiple data sources (NFHS, the Annual Health Surveys, and district-level surveys) to identify state-level gaps in reaching SDG targets of 75% satisfied demand. Similarly, NFHS measurements have been used to estimate progress toward meeting family planning needs for 120 million additional women by 2020 as per the global Family Planning 2020 goal (Stover and Sonneveldt 2017). Government policymakers and program officials have come to expect each NFHS round to provide complete and accurate estimates of key demographic and social indicators. As Pullum et al. (2018: 65) note regarding survey data quality, "It is essential to have data of the highest possible standard, and that requires continuously improved detection of potential data quality issues, and a better understanding of how those issues arise and can be controlled."

An unexpected finding in one development indicator monitored by the NFHS was an apparent decrease from 2005–2006 to 2015–2016 in contraceptive use, the proportion of married women of reproductive age using any method of contraception (Roy, Porwal, and Acharya 2021; IIPS and ICF 2017; Pradhan and Dwivedi 2019). After rising from 40.7% to 48.2% to 56.3% over NFHS-1, NFHS-2, and NFHS-3, respectively, contraceptive use was estimated to be 53.5% in NFHS-4, or a drop of 2.8 percentage points (pps). Use of a modern contraceptive method also appeared to shift downward by 0.8 pps from 48.5% to 47.7% and female sterilization use by 1.3 pps from 37.3% to 36.0%. Pill and condom use showed slightly higher prevalence in NFHS-4 while traditional method use was lower by 2.0 pps (5.8% versus 7.8%). In contrast, the decade between NFHS-3 and NFHS-4 saw the total fertility rate (TFR) drop from 2.7 to 2.2 births per woman; wanted TFR was slightly lower at 1.8 births in NFHS-4 compared to 1.9 births in NFHS-3 (IIPS and ICF 2017; IIPS and Macro International 2007). This shift toward less use of contraception is discordant, if not implausible, with the lower demand for fertility, which augurs instead for an increase in contraceptive use.

Social, economic, and demographic forces over the decade clearly play an important role in the explanation, largely through the motivations and composition of contraceptive users. Our study, though, is interested in the influence of enacted changes in survey design between the two NFHS rounds. Foremost among the change is an expansion of reporting domains. In NFHS-3, estimates of key population indicators were sought at the national and state levels. In NFHS-4, this was expanded to include district-level estimates, with the consequence of the overall sample size increasing dramatically from 124,385 to 699,686 women of childbearing age 15 to 49 years, or more than five-fold. The attendant logistics for fieldwork to implement a significantly larger survey may have generated training coordination and supervisory challenges to monitoring data quality and warrant examination.

A number of survey factors, such as sample, coverage, measurement, and nonresponse, are known sources of potential error (Groves et al. 2009). We focus on two types of nonsampling errors – field implementation and measurement error – that may have influenced the contraceptive use estimate. We do not address the possibility of sampling frames, nonresponse errors, or data processing errors in part because the NFHS has been technically and continuously overseen by the International Institute for Population Sciences (IIPS) under the sponsorship of the Ministry of Health and Family Welfare. Any such errors are likely to be systemic and present in both rounds. In addition, response rates have been high in NFHS: 95% of eligible women in round 3 and 97% in round 4. If this study's results support the role of survey design factors implicated in the measurement of the current use of modern contraception and of female sterilization, and if low use levels continue to be observed in the fifth round recently fielded in 2019–2020,[3] the NFHS will be an even more essential data source for identifying possible reasons.

Our paper will test three hypotheses, the first that survey fieldwork design factors are independently associated with a woman's self-report of contraceptive use, the second that interviewer effects are larger in NFHS-4 than NFHS-3, and third that multiple levels

[3] These data are not yet publicly accessible.

of the effects from state, cluster, and interviewer exist, again larger for the later round. If these hypotheses are supported, it will be important to interpret NFHS-based trends with these caveats in mind. NFHS-5 replicated the survey procedures of NFHS-4 with little change.

We begin with a brief overview of the measurement of contraceptive use through large-scale surveys, followed by a short description of NFHS design and implementation in India. We then detail the analytic sample, variables, measures, and methods used to test our hypotheses of the associations of individual women's reported contraceptive use with fieldwork design factors, with interviewer effects and variance across multiple survey levels. We conclude with a summary of the key findings and discuss their implications for the ensuing NFHS rounds.

Measurement of contraceptive use. A measurement error arises when the obtained assessment of an item deviates from its actual value as a result of interaction between the questionnaire, interviewer, and respondent. There is a large volume of literature documenting fieldwork and fieldworker effects in household surveys (Blom and Korbmacher 2013; Davis et al. 2010; Groves 2004; Groves et al. 2009; MacQuarrie et al. 2018; Singh, Kumar, and Arnold 2022; Tourangeau and Yan 2007). Studies have shown that there is a tendency for respondents to give and for interviewers to record culturally accepted answers (Davis et al. 2010; Liu and Wang 2016; Yang and Yu 2008). Fieldworker effects also arise when interviewers do not read the questions verbatim, as intended, or when they add other information that may confuse or mislead the respondent (Kasprzyk 2005). Sensitive questions are particularly more prone to fieldwork effects (Bignami-Van Assche, Reniers, and Weinreb 2003; Leone, Sochas, and Coast 2021; Tourangeau and Yan 2007; West and Blom 2017). Singh, Kumar, and Arnold (2022) find much larger interviewer effects in NFHS-4 than in NFHS-3 on sensitive questions such as women justifying a woman's refusal to have sex with her husband, women justifying wife beating, women's experience of physical and sexual violence, and whether the woman's father ever beat her mother. Notably, Leone, Sochas, and Coast (2021), using data from NFHS conducted in India in 1998–1999, do not find any interviewer effects in responses to abortion-related questions.

Survey-based measurement of contraceptive practice. There is a long history of measurement of individual contraceptive practice through large-scale household surveys, which for non-Western regions traces back to multinational programs such as the World Fertility Surveys (WFS) (Sprehe 1974), the Demographic and Health Surveys (DHS) (Corsi et al. 2012; Fisher and Way 1988), the Contraceptive Prevalence Surveys (CPS) (Lewis 1983), the Reproductive Health Surveys (Morris 2000), and more recently the Performance Monitoring and Accountability 2020 surveys (Zimmerman et al. 2017). These survey programs, all collecting primarily cross-sectional data, undergird an overwhelming majority of the global research regarding levels, trends, and factors related to population-based contraceptive behaviors (Fabic, Choi, and Bird 2012). This body of research has benefited from largely standardized survey questions asked of a common sample population: female respondents in childbearing years, generally defined as 15 to 49. Women self-report their use of a pregnancy avoidance method and type, if any, at the time of the face-to-face interview based on a structured questionnaire administered by a female interviewer. Anderson and Cleland (1984), comparing contraceptive prevalence estimates between closely timed WFS and CPS surveys in the same countries, find relatively consistent estimates despite minor differences in survey procedures. (With the advent of the male questionnaire in the DHS, eligible men aged 15 to 54 years were also asked about contraceptive use although often linked to use at last sex.)

The female respondents' answers to two basic questions compose the data for what has become the most widely used population measure of current contraceptive use, the contraceptive prevalence rate (CPR). The two questions are (1) "Are you currently doing something or using any method to delay or avoid becoming pregnant?" and if yes, (2) "Which method are you using?" Slight variations have occurred in wording for different surveys over time, such as the inclusion "or your partner" in the first question. More relevant, however, for measurement reliability is whether questions about method-specific knowledge, awareness, and possible previous use of one or more methods precedes the question about current use, as these can prime the interviewee's readiness to respond. Choi et al. (2019) find substantial underreporting of both current use and female sterilizations among Indian women in Rajasthan who report ever being sterilized.

Beyond question wording, a number of other survey instrument–related factors play a role in the validity, reliability, and reproducibility of self-reported measurement of contraceptive use. For one, the referent time period for "currently" is left to the respondent to define. Another factor is the mode of administration (e.g., face-to-face, computer-assisted, or self-administered), which has been studied for data quality implications (Greenleaf et al. 2020). Third, validation of contraceptive use in low-income settings poses its own challenges and can require clinical testing of blood samples for the presence of contraceptive hormones (Achilles et al. 2018). In high-income countries, researchers can rely on medical prescription or electronic client records to compare against individual patient responses (e.g., Ichikawa et al. 2015). Moreover, with respect to the type of method, an individual woman's simultaneous use of more than one contraceptive method (e.g., dual-method use) can complicate assessment. Most survey programs assign the most effective method as the woman's response, usually disregarding others mentioned. Finally, repeated interviews with panel samples do not necessarily mean accurate reporting. A short-term follow-up study of female clients of health facilities in six urban sub-Saharan African settings found their method reported during an exit interview was often inconsistent with the one they recalled when interviewed four months later (Tsui et al. 2021).

There have been measurement innovations over time to better capture the dynamics of contraceptive practice, including daily diaries (Huber et al. 2013) and a retrospective calendar embedded into survey questionnaires (Strickler et al. 1997; Goldman, Moreno, and Westoff 1989; Bradley, Schwandt, and Khan 2009; Smith, Edwards, and Free 2018; Callahan and Becker 2012). The calendar records the respondent's recalled episodes of contraceptive use, type of method, and reasons for stopping over a period of about five to six years. These duration data, however, are subject to the same potential biases of social desirability or inaccurate recall as self-reported current use. Stuart and Grimes (2009) cite the former as the major threat to measurement reliability. Strickler et al. (1997) summarize the issue aptly, acknowledging that individual-level contraceptive self-reports from calendar instruments will vary more than estimates aggregated at the population level.

Despite their apparent deficits, self-reported contraceptive use is relied upon to obtain the numerator in estimating CPR. The denominator is usually defined as all women ages 15 to 49, although marital status and/or sexual experience are sometimes applied to approximate the actual population at risk of pregnancy. Women who are currently pregnant, however, are retained in the denominator because the gestation length is short.

The CPR can be further conditioned on the use of modern types of contraception methods (mCPR), which can include any of the following: female or male sterilization; hormonal implants, injectables, or pills; intrauterine devices; male or female condoms; diaphragms; foam; jelly; and fertility-awareness methods. The modern category excludes rhythm or periodic abstinence, withdrawal, and folk methods. Use of menstrual regulation or abortion is excluded as a contraceptive method. The mCPR is then the percentage of childbearing-aged women reporting current use of a modern method of contraception.

NFHS design and field implementation. The first NFHS in 1992–1993 covered all states except Sikkim and interviewed about 90,000 women who have ever been married. NFHS-2 was conducted in all states, also interviewing the same sample composition and size. Its content expanded over NFHS-1 to include questions about reproductive health, women's autonomy, domestic violence, women's and child nutrition, anemia, and salt iodization. NFHS-3 in 2005–2006 continued the survey design and content of prior rounds while adding community-based HIV testing as well as a male interview. NFHS-4 in 2015–2016 expanded its survey domains, beyond nation and state or union territory, to estimate indicators at the district level, previously obtained through the District-Level Household Survey (IIPS 2001). Both NFHS-3 and NFHS-4 differed in scale and complexity from earlier rounds (see Singh, Kumar, and Arnold 2022) with each occurring in two phases. The NFHS-4 survey content expanded beyond NFHS-3 by adding the collection of women's blood pressure and blood glucose. Also for the first time, the questionnaire was administered as a computer-assisted personal interview; as a result, the field teams did not include a field editor. (The main differences between NFHS-3 and NFHS-4 are summarized in Appendix Table A-1.)

NFHS-3's first phase occurred from November 2005 to May 2006 and the second phase from April to August 2006, with an average duration of fieldwork in a state of 4.6 months. It involved a sample of 3,850 primary sampling units (PSUs), 264 teams, and 916 interviewers. In contrast, NFHS-4 was conducted over a two-year period, January 2015 to December 2016, with an average duration of fieldwork in a state of 7.8 months. NFHS-4's first phase was from January to December 2015 and the second phase from January to December 2016. It involved a sample of 28,521 PSUs, 875 field teams, and 2,734 interviewers (IIPS and ICF 2017).

The implementing role of survey agencies is important to consider when discussing the quality of NFHS data. Agencies submitted competing bids to be engaged in NFHS work, and those selected were assigned to one or more specific states or part of a state (see NFHS-3 and NFHS-4 final reports for details). In NFHS-3, 18 survey agencies – 13 private and 5 population research centers (PRCs), established by the Ministry of Health and Family Welfare – were involved in fieldwork of the smaller-scaled NFHS-3 (IIPS and Macro International 2007), while in NFHS-4 only 14 survey agencies – 11 private and 3 PRCs – were engaged (IIPS and ICF 2017). Many important tasks, such as the hiring of field interviewers, training of field interviewers, ensuring adherence to survey protocols, monitoring and supervising fieldwork in real time, and administering timely payment of salary and other remuneration, rested with the survey agencies. Their execution likely varied from one survey agency to another. Due to the complex bidding process adopted to hire field agencies in NFHS, the cost of the survey per household that they proposed varied substantially. It is possible that field agencies may at times have reduced interviewers' salaries to increase their profits.

Table 1: Descriptive statistics for selected contraceptive outcomes and selected covariates of interest for currently married women 15 to 49, NFHS-3 and NFHS-4

Cohort measure	NFHS-3: 2005–2006		NFHS-4: 2015–2016	
	% / mean	95% CI	% / mean	95% CI
N	85,960		496,184	
Outcomes				
Currently using modern contraception (%)	48.8	(48.1–49.4)	47.7	(47.4–48.0)
Currently using female sterilization (%)	37.4	(36.7–38.1)	36.0	(35.8–36.3)
Survey design				
Survey months elapsed since interview started (mean)	1.7	(1.6–1.7)	2.4	(<2.4–2.4)
Fieldwork phase 1 (%)	63.7	(63.1–64.3)	65.1	(64.8–65.4)
Fieldwork phase 2 (%)	36.3	(35.7–36.9)	34.9	(34.6–35.2)
Respondent background				
Education (with primary or above schooling) (%)	53.3	(52.4–54.1)	67.1	(66.8–67.3)
Residence in urban area (%)	30.9	(30.2–31.5)	33.4	(33.1–33.8)
Household in poorest wealth quintile (%)	27.6	(26.7–28.4)	16.1	(15.9–16.3)
Household in highest wealth quintile (%)	14.6	(13.9–15.3)	24.1	(23.7–24.5)
Belong to scheduled caste/tribe (%)	26.8	(25.8–27.9)	29.4	(29.0–29.9)
Belong to Hindu religion (%)	81.8	(80.7–83.0)	81.5	(81.1–81.9)
Year of birth (mean)	1973.9	(1973.8–1974)	1982.0	(1982–1982.1)
Age in years (mean)		(31.2–31.4)	32.8	(32.8–32.8)
Number of sons ever born (mean)	1.4	(1.3–1.4)	1.1	(<1.1–1.1)
Number of daughters ever born (mean)	1.5	(<1.5–1.5)	1.3	(1.2–1.3)
Desires more sons than daughters (%)	25.5	(24.9–26.1)	21.3	(21.0–21.5)

In NFHS-3 only one training event for field staff teams was organized per state. For NFHS-4 IIPS advised a training event be held for every 13 teams. With some 800 field teams (each comprised of seven members), at least 15 populous states had to hold multiple training events, the schedule and frequency of which were the responsibility of the survey agencies. Organizing these events, which were to provide field staff with standardized and uniform training, at times occurred simultaneously and strained staff resources, particularly in the early days of fieldwork. In large states, such as Maharashtra, Bihar, and Uttar Pradesh, the number of teams ranged from 44 to 48 to 90, respectively. Uttar Pradesh was segmented into three efforts (east, west, and central) and conducted by three agencies. Singh, Kumar, and Arnold (2022), assessing and comparing interviewer effects on responses to selected sensitive questions, find larger ones in NFHS-4. They attribute the effects to the larger workload of field interviewers, gauged by the increase in their average monthly interview count and duration of fieldwork. The average number of interviews conducted by an interviewer in NFHS-3 was 136 compared to 256 in NFHS-4. The large-scale simultaneous and multiple-state launch in phase 1 of NFHS-4 fieldwork likely placed pressure also on central training capacities and protocols, with the eventual consequence of interviewers remaining inexperienced and initially less confident in conducting interviews and asking and probing for complete responses regarding contraceptive use.

2. Data and methods

Data sources. The data analyzed for this paper are derived from NFHS-3 and NFHS-4. Women not currently married are excluded (36,460 in NFHS-3 and 200,059 in NFHS-4). Cases with missing values on the covariates of interest are also excluded (1,965 or 2.2% in NFHS-3 and 3,443 or 0.7% in NFHS-4). The analytic sample is thus confined to women 15 to 49 years currently married or in a union at the time of the survey with complete data – 85,960 for NFHS-3 and 496,184 for NFHS-4. We apply the survey sample weights as constructed and accounted for complex sample selection design for each round when presenting results from our descriptive and multivariate analyses.

Measures. We examine two contraceptive use outcomes that are likely sensitive to interviewer experience and respondent comprehension, the first being the reported current use of a modern contraceptive method and the second being the reported use of female sterilization. These are based on the two standard questions mentioned earlier, which did not change over the two rounds. Answering yes to whether she is currently using requires the respondent recognize an ongoing behavior related to pregnancy avoidance. She must also be able to identify and willingly report the method to the interviewer. Although female sterilization accounts for nearly three-quarters of modern method use in India, neither NFHS-3 nor NFHS-4 included a direct question on whether the woman had been sterilized. Median age at sterilization was 25.5 years for women in NFHS-3 and 25.7 years in NFHS-4 (IIPS and ICF 2017; IIPS and Macro International 2007). If respondents' recall and interpretation of what constitutes current use of contraception were faulty, there can be an undercount of users of sterilization and possibly other methods, such as traditional methods.

We concentrate on three survey design and field implementation factors: (1) the NFHS round (3 or 4), (2) whether the woman was interviewed in fieldwork phase 1 or 2, and (3) the fieldwork teams' experience at the time of her interview. We classify the female respondent as interviewed in round 3 or 4 and if her interview took place in phase 1 or 2. We expect the round-specific variable to capture measurement error and other unique survey conditions not otherwise directly observed or explained by the model's other covariates. For distinguishing fieldwork phases, in NFHS-3 12 states/union territories were covered in phase 1 and 17 in phase 2. In NFHS-4, 17 states and territories participated in phase 1 and included populous states such as Madhya Pradesh, Bihar, Karnataka, Tamil Nadu, and one-third of the Uttar Pradesh sample. Phase 2 included Rajasthan, the other two-thirds of the Uttar Pradesh sample, Odisha, Gujarat, and another 13 states or union territories. Because corrections to fieldwork practices and acquired interviewer experience grew by phase 2, we hypothesize this survey design factor will be associated with improved contraceptive use reporting. For the third design factor, we calculate for each respondent the duration (in months) since the interviews in her state first began. If the survey began in January and she was interviewed in March, her value is two months. Higher values for this variable are expected to reflect more experience on the part of the field interviewer.

Nine respondent-level background variables are selected as covariates for adjustment; these have been commonly used in other analyses to capture potential variation in contraceptive practice among Indian women due to socioeconomic, demographic, reproductive, and gender preference influences (e.g., Pradhan and Dwivedi 2019; Pal and Makepeace 2003; Chacko 2001). Socioeconomic variables include the female respondent's education (measured dichotomously as any/no schooling), urban residence (at the time of the survey) versus rural, and household wealth category (poorest quintile, richest quintile, or middle three-fifths as the reference). The household asset score was standardized using NFHS-3 factor weights and applied to households in NFHS-4. Women were also classified according to their household head's caste, specifically scheduled caste or tribe (yes/no) and religious affiliation (Hindu or other). Demographic and reproductive variables include the woman's year of birth and numbers of sons and daughters who have ever been born, which jointly capture parity at the time of the survey, all measured continuously. The woman's preference for sons versus daughters is based on her expressed son preference and measured dichotomously as her desire to have more sons than daughters (yes/no). After being asked to reflect on her ideal number of children, she is asked, "How many of these children would you like to be boys, how many would you like to be girls, and for how many would it not matter if it's a boy or a girl?" Unfortunately, it is not possible to measure a woman's use of abortion accurately, which can influence her contraceptive use, as there were too few reports of terminated pregnancies in the questionnaire's calendar. It is worth noting, though, that abortion was legalized in 1971, that research (e.g., Singh et al. 2018; Singh, Remez, and Tartaglione 2010) suggests it is highly underestimated in surveys, and that sex-selective abortion of females persists (Jha et al. 2011).

Analytic approach. We first estimate multivariable logistic regressions of the reported modern contraceptive method and female sterilization use by survey months elapsed since the interview was initiated in the respondent's state and adjusted for urban/rural residence using

$$log[p/(1-p)] = \beta_0 + \beta_1 M + \beta_2 U,$$

where p is the expected probability of using contraception and M and U are the survey months elapsed since the interviews started in the state and an urban residence dummy, respectively. These state-specific patterns for NFHS-3 and NFHS-4 are graphically shown in Figures 1a and 1b (the odds ratios are provided in Appendix Table A-2a and A-2b). This exploratory analysis establishes whether the reporting of current use of modern contraception and sterilization increased with the duration of fieldwork in the states.

Next, descriptive statistics on the two NFHS female samples for the outcomes and covariates of interest are presented in Table 1. We then estimate another set of multivariable binary logistic regression models for modern contraceptive and sterilization method use at the time of the survey (Table 2) as a function of survey design factors, adjusted for respondent background covariates and state fixed effects. The log odds of modern contraceptive method use can be mathematically expressed as

$$log[p/(1-p)] = \beta_0 + \beta_1 N + \beta_2 M + \beta_3 F + \rho C + \theta S,$$

where p is the expected probability of using a modern contraceptive method, N is the survey round of NFHS (3 or 4), M is the survey months elapsed since the interview started in the state, F is the fieldwork phase (1 or 2), and C is the vector of background variables (socioeconomic, demographic, reproductive, and gender preferences). S is the state fixed effect that captures the NFHS round-specific (time-invariant) unobserved heterogeneity across states. This analysis helps assess the relative importance of the survey design and implementation factors and the background variables (shown in left panel of Table 2).

We also estimate from the regression models the average marginal effects of the different factors, or the predicted probabilities if all women have the attribute of interest, holding the other covariates' values at their mean values (right panel of Table 2). The multivariable regression models include state fixed effects (not shown); these are correlated with the market research agency. The standard errors of all estimates have been adjusted for the surveys' complex sampling design using the Taylor series linearization method.

Table 2: Results of multivariable logit regression and average marginal effects for two contraceptive outcomes on survey design factors and selected background covariates from the combined sample of NFHS-3 and NFHS-4 currently married female respondents ages 15–49

		Current use of modern contraception	Current use of sterilization	Current use of modern contraception		Current use of sterilization	
Covariate	Value	OR (95% CI)	OR (95% CI)	AME	SE	AME	SE
Married women (*N*)		582,144		582,144			
Survey design/implementation							
NFHS round	3 (Ref)			0.445	0.003	0.293	0.003
	4	1.21 (1.17,1.26)	1.66 (1.59,1.74)	0.485	0.001	0.375	0.001
Survey months elapsed since interview started	C	1.16 (1.15,1.17)	1.18 (1.16,1.19)	0.030	0.001	0.027	0.001
Fieldwork phase	First (Ref)			0.457	0.002	0.356	0.002
	Second	1.35 (1.30,1.40)	1.12 (1.07,1.17)	0.518	0.003	0.375	0.003
Respondent background							
Educated	None=0 (Ref)			0.458	0.002	0.366	0.002
	Any=1	1.17 (1.14,1.20)	0.97 (0.94,0.99)	0.489	0.001	0.360	0.001
Poorest	None=0 (Ref)			0.493	0.001	0.371	0.001
	Yes=1	0.65 (0.63,0.66)	0.72 (0.69,0.74)	0.404	0.003	0.316	0.003
Richest	None=0 (Ref)			0.475	0.001	0.370	0.001
	Yes=1	1.09 (1.06,1.12)	0.83 (0.80,0.85)	0.492	0.003	0.339	0.002
Urban	None=0 (Ref)			0.479	0.001	0.374	0.001
	Yes=1	0.99 (0.96,1.02)	0.82 (0.79,0.85)	0.478	0.002	0.341	0.002
Scheduled caste/tribes	None=0 (Ref)			0.482	0.001	0.361	0.001
	Yes=1	0.95 (0.93,0.98)	1.02 (0.99,1.05)	0.471	0.002	0.365	0.002
Hindu religion	None=0 (Ref)			0.399	0.003	0.259	0.003
	Yes=1	1.62 (1.57,1.67)	2.19 (2.10,2.28)	0.496	0.001	0.386	0.001
Year of birth	C	0.96 (<0.96,>0.96)	0.93 (<0.93,>0.94)	–0.008	0.000	–0.012	0.000
Number of sons ever born	C	1.19 (1.18,1.20)	1.19 (1.18,1.20)	0.036	0.001	0.029	0.001
Number of daughters ever born	C	1.70 (1.68,1.72)	1.80 (1.78,1.83)	0.108	0.001	0.100	0.001
Desires more sons than daughters	None=0 (Ref)			0.484	0.001	0.367	0.001
	Yes=1	0.88 (0.86,0.90)	0.89 (0.87,0.92)	0.458	0.002	0.348	0.002

Notes: AME = average marginal effects. C = continuous units. Model standard errors adjusted for complex survey design. State-level fixed effects included.

Second, to investigate the potential effects of interviewers we conduct two analyses, the first of which assesses if there is significant variation across interviewers in recording contraceptive use by examining their deviance residuals. Deviance residuals combine the magnitude of the difference from the outcome's mean with the number of cases on which the interviewer-specific prevalence is based (see Pullum et al. 2018 and Pullum 2019; McCullagh and Nelder 1989). Following Pullum et al. (2018), we treat interviewers as units of analysis, where each interviewer conducts n interviews and records contraceptive use n_1 and nonuse n_0 number of times, such that $n = n_1 + n_0$. The generalized linear model function is

$$g(n_1) = b_0 + b_1 T, family(binomial, n)\ link(logit),$$

where $g(.)$ is the linearizing link function, b_0 is constant, and T is a dummy variable for the interviewer's state-specific team. We estimate this logit regression for outcome y with an interviewer ID variable code to obtain the total deviance D. Post-estimation we predict the deviance residual for each interviewer. We assess the percent of interviewers with deviance residuals statistically significant (Table 3) and then calculate D by summing deviance residuals across interviewers. D follows a chi-square distribution with $K - 1$ degrees of freedom (K being the total number of interviewers). Deviance residuals should be small; large values suggest significant interviewer effects. These can be compared between NFHS-3 and NFHS-4, as presented in Table 4 (top panel).

Table 3: Percent of interviewers whose deviance residual is significant at alpha values <0.05 and 0.01

	NFHS-3: 2005–2006 (*n* = 632)		NFHS-4: 2015–2016 (*n* = 2,614)	
Outcome	%	*N*	%	*N*
Current use of modern contraception				
p ≤ 0.05	2.8	632	25.8	2,614
p ≤ 0.01	0.5	632	15.9	2,614
Current use of female sterilization				
p ≤ 0.05	3.0	632	20.4	2,614
p ≤ 0.01	0.3	632	13.2	2,614

Notes: In NFHS-3, 284 of the total 916 interviewers could not be included in this analysis due to inaccurate ID numbers. In NFHS-4 the number of interviewers excluded was 120 out of 2,734 total.

If significant variation across interviewers is observed, the next step is to identify problematic interviewers. Interviewers whose deviance residual differed significantly from the overall mean using $p \leq 0.05$ or $p \leq 0.01$ as critical region cutoffs are classified as problematic. We then remove the respondents interviewed by problematic interviewers and re-estimate the two outcomes to compare them to the original values (bottom panel of Table 4).

Table 4: Modern contraceptive methods and sterilization use reported and estimated after removing respondents of problematic interviewers

	NFHS-3: 2005–2006		NFHS-4: 2015–2016	
Actual	%	95% CI	%	95% CI
Currently using modern contraception	*N* = 85,960		*N* = 496,184	
	48.8	(48.1–49.4)	47.7	(47.4–48.0)
Currently using female sterilization	*N* = 85,960		*N* = 496,184	
	37.4	(36.7–38.1)	36.0	(35.8–36.3)
Estimated after removing respondents of problematic interviewers				
Currently using modern contraception	*N* = 83,951		*N* = 364,858	
	48.9	(48.2–49.5)	49.9	(49.6–50.2)
Currently using female sterilization	*N* = 84,443		*N* = 395,087	
	37.3	(36.6–38.0)	37.3	(37.0–37.6)

Note: Interviewers are identified as problematic if their deviance residual was significant at $p \leq 0.05$ levels.

Because the deviance analysis is a fixed-effects approach, we also assess interviewer effects separate from those of survey strata (state and community cluster) by estimating cross-classified random intercept models that are more standard in the survey methodological literature for analyses of interviewer effects (West and Blom 2017). The models are estimated for both contraceptive use outcomes; we obtain the intraclass correlation coefficients (ICCs) for the state, sample cluster or PSU, and interviewer and residual (woman) levels (presented in Table 5). We model the contraceptive use outcomes (y) for the i^{th} respondent interviewed by the j^{th} interviewer in the k^{th} PSU in the l^{th} state as

$$\text{Logit}\,(\Pr(y_{ijkl}=1|C_i, \varsigma_j, \varsigma_k, \varsigma_l)) = \beta_0 + \rho C_i + \varsigma_j + \varsigma_k + \varsigma_l,$$

where C is a vector of individual-level socioeconomic, demographic, and residence-related characteristics. While ς_j and ς_k represent the cross-classified interviewer and PSU-level random intercepts, ς_l represents the state-level random intercept.

We cross-classify the random intercept models at the interviewer level and the PSU level because interviewers are not nested within the PSUs in surveys from the NFHS. This approach helps account for the interviewer effects that might arise due to assignment of some interviewers to PSUs in which respondents may have a different propensity to report contraceptive use outcomes (Leone, Sochas, and Coast 2021). The ICC in this case indicates the percent of the variance in the contraceptive use outcomes that is explained by variation due to interviewers, PSUs, and states. The ICC allows us to compare the interviewer effects with those of state and community, as well as between the two surveys. It is important to capture the measurement errors that interviewers may introduce in a survey through their skills; beliefs; and social, economic, and demographic characteristics.

Table 5: Variance and ICC values from multilevel cross-classified logistic models of state, sample cluster, interviewer, and woman level effects on two contraceptive use outcomes. Currently married women 15–49 in NFHS-3 and NFHS-4

			Current use of modern contraception		Current use of female sterilization	
Survey		Level	Null/bivariate model	Random intercept/multivariable model	Null/bivariate model	Random intercept/multivariable model
NFHS-3: 2005–2006	Variance	State	0.46	0.51	0.67	1.05
		PSU	0.31	0.37	0.37	0.58
		Interviewer	0.00	0.00	0.00	0.00
		Residual	3.29	3.29	3.29	3.29
	ICC	***State***	***11.3%***	***12.2%***	***15.5%***	***21.4%***
		PSU	***7.7%***	***8.8%***	***8.6%***	***11.7%***
		Interviewer	***0.0%***	***0.0%***	***0.0%***	***0.0%***
NFHS-4: 2015–2016	Variance	State	0.60	1.01	1.15	2.00
		PSU	0.51	0.55	0.60	0.87
		Interviewer	0.34	0.45	0.29	0.50
		Residual	3.29	3.29	3.29	3.29
	ICC	***State***	***12.6%***	***19.0%***	***21.6%***	***30.1%***
		PSU	***10.8%***	***10.4%***	***11.3%***	***13.0%***
		Interviewer	***7.3%***	***8.5%***	***5.5%***	***7.5%***

Notes: ICC = intraclass correlation coefficient, which measures the percentage share of the total variance contributed by different levels: women, interviewer, PSU, and state. Random intercept/multivariable models are adjusted for variables controlled in Table 2.

3. Results

Figures 1a and 1b confirm the expected variation in fieldwork experience in relation to the reporting of current modern contraceptive and female sterilization use across the states and union territories among currently married women. A pattern of nominal differences is visible for NFHS-3 in state-specific odds ratios for reported contraceptive use by survey months elapsed since the start of interviews, adjusted only for urban/rural residence. A much more dramatic pattern of differences is readily apparent for states in NFHS-4. The state trend lines over the months show that as field teams gained more experience, the odds of reporting modern contraceptive and female sterilization use increased relative to the month when the survey began in each state. The trend lines for NFHS-4 modern contraceptive use rise strikingly quickly in several states (Haryana, Uttar Pradesh East, Madhya Pradesh West, Bihar, Karnataka, Tamil Nadu, Odisha, Himachal Pradesh, Uttar Pradesh Central, and Jammu and Kashmir). For example, in

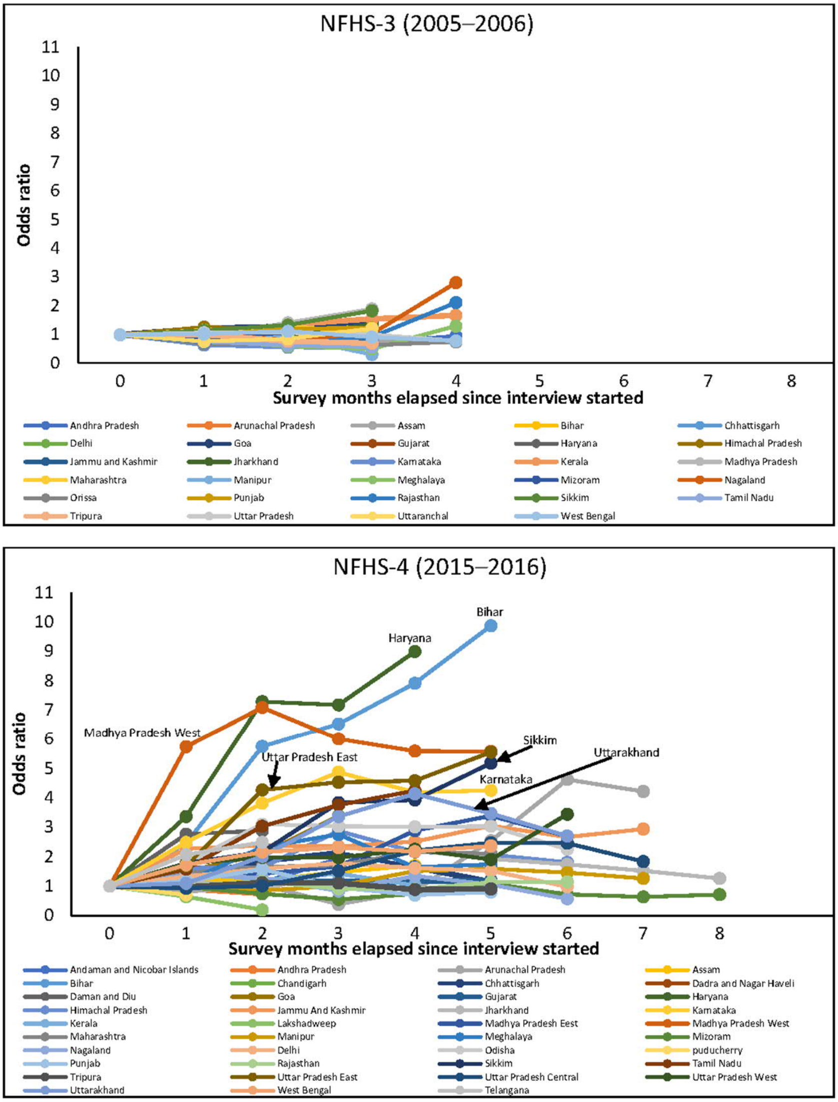
Bihar state, 48 field teams launched interviewing in phase 1; with low reporting of contraceptive use in the first few months, the adjusted odds compared to the first month rose to 9.87 by six months (Appendix Table A-2a). With this exploratory confirmation, we proceed to investigate the relative influence of survey design factors and interviewer effects.

Descriptive statistics for the two survey samples are provided in Table 1 and reveal changes in the percentages of currently married women reporting use of modern contraception and female sterilization. The prevalence of modern contraception use is 48.8% (48.1%–49.4%) in NFHS-3, which declined to 47.7% (47.4%–48.0%) in NFHS-4, with no overlap in confidence intervals. Current use of female sterilization is 37.4% (36.7%–38.1%) in NFHS-3 and 36.0% (35.8%–36.3%) in NFHS-4, also with no overlap in confidence intervals. For both outcome measures, less use is measured in NFHS-4.

The average length of time since a fieldwork start to a woman's interview is longer in NFHS-4 at 2.4 months than NFHS-3 at 1.7 months. Nearly two-thirds of interviews occurred in phase 1.

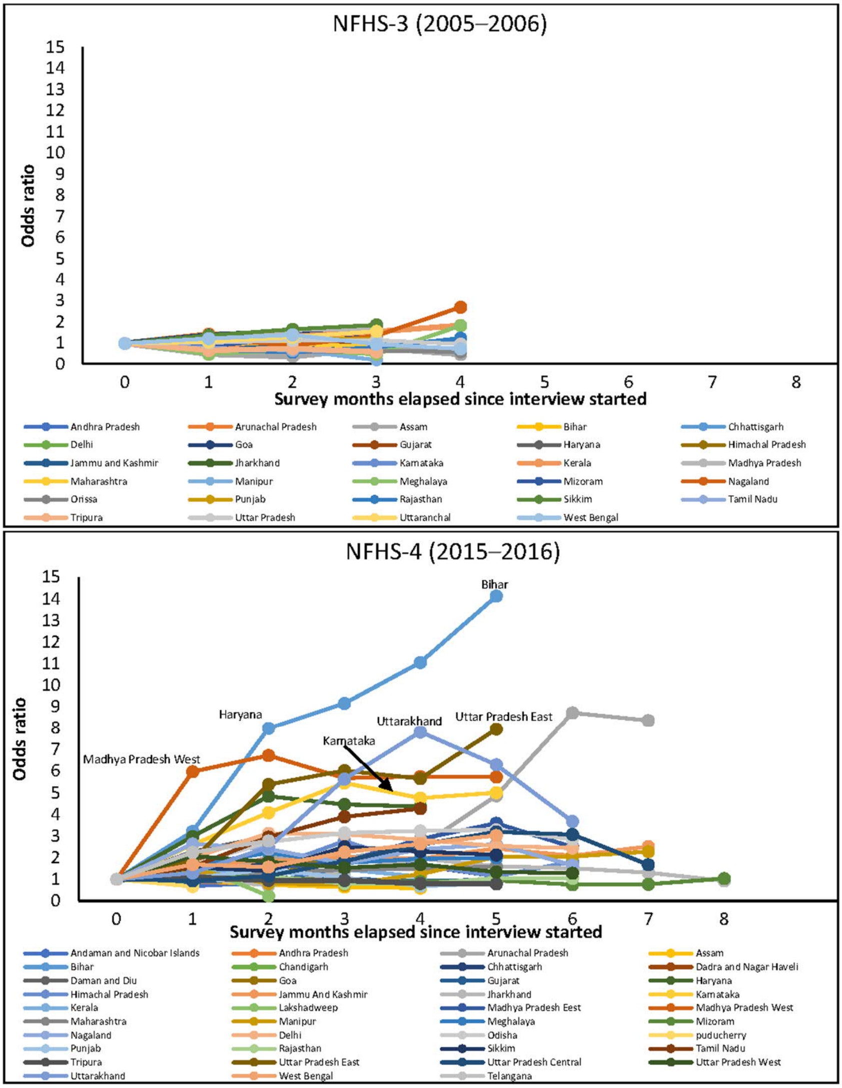
Although our focus is on the role of survey design factors, it is worth noting the change in sample composition over the decade. The percent of currently married women with at least primary schooling rises from 53.3% in NFHS-3 to 67.1% in NFHS-4. Urban residence increases slightly from 30.9% to 33.4%. Based on the NFHS-3 standardized asset index, the percent of women living in households in the poorest wealth quintile declined from 27.6% to 16.1% and in the wealthiest quintile rose from 14.6% to 24.1%. The percent of women belonging to households in the scheduled caste or tribe changed slightly from 26.8% to 29.4% between the two rounds, and those in households belonging to the Hindu religion remains same at 82%. The average age of the female sample increases slightly from 31.3 to 32.8 years between rounds. The average number of sons ever born at survey declines from 1.4 to 1.1 as does the number of daughters, from 1.5 to 1.3. Son preference declines from 25.5% of respondents in NFHS-3 to 21.3% in NFHS-4.

Figure 1a: Adjusted odds ratios of modern contraceptive use among currently married women by months since interview started in a state, NFHS-3 and NFHS-4 (reference is month 0 of the survey)



Notes: Odds ratios are adjusted for urban/rural residence. Analysis includes only those months in which 50 or more interviews were completed in each state/territory in each round.

Figure 1b: Adjusted odds ratios of sterilization use among currently married women by months since interview started in a state, NFHS-3 and NFHS-4 (reference is month 0 of the survey)



Notes: Odds ratios are adjusted for urban/rural residence. Analysis includes only those months in which 50 or more interviews were completed in each state/territory in each round.

We estimate multivariable logistic regression models in Table 2 for each of the two contraceptive use outcomes with the combined sample of currently married women from the two rounds and calculate the average marginal effects for the covariates. We interpret only the results for the survey factors and not the background covariates, which serve as controls. The estimated adjusted odds ratio (AOR) of reporting modern use was 1.21 (1.17 to 1.26), which is higher for NFHS-4 compared to NFHS-3 respondents. Likewise, the AOR for reporting use of female sterilization in NFHS-4 was even larger at 1.66 (1.59 to 1.74) than for NFHS-3. Each month following a survey launch increased the reporting of modern contraceptive use, with an AOR of 1.16 (1.15 to 1.17), and female sterilization, with an AOR of 1.18 (1.16 to 1.19). Furthermore, fieldwork during the second survey phase, compared to the first, had a higher odds of respondents reporting modern use, with AOR of 1.35 (1.30 to 1.40), and of female sterilization, with an AOR of 1.12 (1.07 to 1.17).

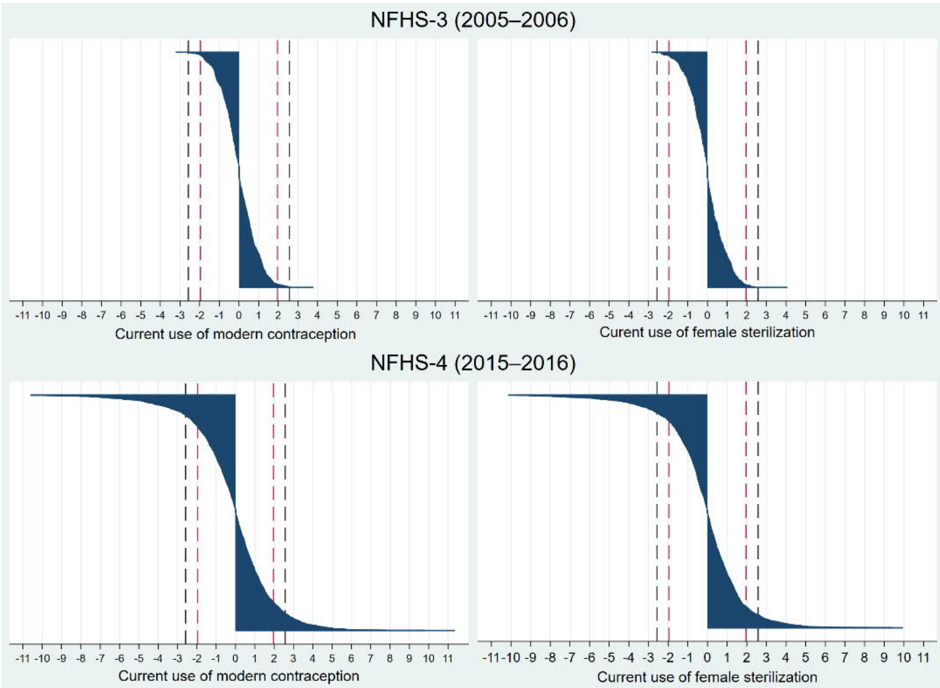
It is easier to interpret and compare the effect sizes of these AORs by calculating the average marginal effects, which are the average changes in the probability of either modern contraceptive use or female sterilization use given a unit increase in the covariate of interest, while all other model covariates are held at their mean values. For example, the predicted probability of modern contraception use is higher if all currently married women experienced NFHS-4 survey conditions compared to NFHS-3 (0.485 versus 0.445). Similarly, the predicted probability for female sterilization use is 0.375 if all women experienced NFHS-4 survey conditions compared to 0.293 if all experienced NFHS-3 conditions. These differences of 4.0 pps in modern contraception use and 8.3 pps in female sterilization use are sizable and can be compared to the observed differences of 1.1 pps and 1.4 pps, respectively. The round-related results differ from our expectations and are discussed further below.

In the case of phase effects, where the second phase has more experienced interviewers, the predicted probability of modern use if all respondents had been interviewed in the second phase is 0.518 compared to 0.457 if all were in the first phase, and for female sterilization 0.375 versus 0.356. Net of all other model covariates, interviews conducted each additional month after the survey began are associated with a 3.0 pps increase in the predicted mCPR and 2.7 pp increase in female sterilization. Both time since survey launch and phase 2 participation (i.e., being interviewed by field investigators with greater experience) are consistently associated with higher predicted probabilities of use for the two outcomes. The models indicate that predicted modern contraception use could be higher by 6.1 pps in NFHS-4 based on the fieldwork phase and predicted female sterilization use higher by 2.0 pps, not lower, as was observed. The magnitudes of the survey design factors' differences are nontrivial, particularly since two-thirds of the sample were interviewed in phase 1 and may thus play a role in the unexpectedly low measured levels of modern contraceptive use in NFHS-4.

As noted earlier, the direction of association for round 4 compared to round 3 effects for the two outcomes was opposite to our expectations. We conducted sensitivity analyses by testing the sequential inclusion of each respondent background covariate and observing any shifts in the AORs for the three survey design factors (results shown in Appendix Tables 3a and 3b). When only the design factors are present, with or without state fixed effects, the odds ratio for the NFHS round is less than 1.0, meaning the reporting of either contraceptive outcome has lower odds in NFHS-4 compared to NFHS-3. With the inclusion of the woman's birth cohort (year of birth), the odds ratio for the NFHS-4 round rises above 1.0, specifically to 1.47 (95% CI = 1.42–1.52) for modern method use and 1.82 (95% CI = 1.75–1.89) for sterilization use. The odds ratio for birth year is consistently below 1.0, at 0.91 and 0.93, respectively, for the two outcomes, indicating lower odds for younger women reporting contraceptive use. The NFHS round is the only survey design covariate in the models whose AOR changes from negative to positive odds with the inclusion of the woman's age.

In terms of assessing the role of interviewer effects in the two surveys, we examine deviance residuals from the overall means of the two outcomes across interviewers. Table 3 and Figure 2 present the percent of interviewers whose deviance residual is significant at $p \leq 0.05$ and $p \leq 0.01$ levels. These conventional values are used to readily and visually demarcate critical regions. Two vertical red lines are plotted in Figure 2: The one on the left identifies negative deviance residuals and the one on the right positive deviance residuals that are large enough to be significant at the 0.05 level. Two black lines are also plotted to demarcate the negative and positive deviance residuals large enough to be significant at the 0.01 level. These p-values help identify residual deviances that are too large to be attributed to random variation. In NFHS-3, 284 of the 916 total interviewers and 120 of the total 2,734 in NFHS-4 were excluded because their ID numbers were not 1, 2, or 3. Only a maximum of three female interviewers could be on a team. Of the 632 remaining interviewers, a small proportion had large negative deviance residuals, and a few had large positive residuals for modern contraceptive use at the 0.05 and 0.01 levels (2.8% and 0.5%, respectively) and fewer for sterilization (3.0% and 0.3%). By contrast, in NFHS-4, a large percent of the 2,614 interviewers had large negative or positive deviance residuals at the 0.05 and 0.01 levels (25.8% and 15.9%, respectively) for current modern use and the percentages at the two p-levels remain sizable for sterilization use (20.4% and 13.2%). Further examination found that the problematic interviewers were primarily from the same 'outlier' states seen in Figures 1a and 1b. The larger NFHS-4 values than those of NFHS-3 for the two outcomes in Table 3 support the hypothesis of greater interviewer effects in the later survey.

Figure 2: Deviance residuals from the overall mean across interviewers for modern contraceptive use and sterilization use outcomes: NFHS-3 2005–2006 and NFHS-4 2015–2016



Note: Deviance residuals are significant at p ≤ 0.05 outside the red dashed lines and p ≤ 0.01 outside the black dashed lines.

A large proportion of negative deviance residuals, as observed, can lead to a lower estimate of modern contraceptive prevalence. NFHS-4 has more and larger significant deviance residuals than NFHS-3. At the same time, NFHS-3 has the same or more positive deviance than negative deviance residuals while NFHS-4 has more negative than positive deviance residuals. We ascertain the effect of problematic interviewers by removing their interviewed respondents from the samples and re-estimating modern contraceptive and sterilization use (bottom panel of Table 4). There is negligible change in the two outcomes in NFHS-3 after removing problematic interviewers, but there are significant increases in NFHS-4 of 2.2 pps in modern contraceptive prevalence from 47.7% (47.4% to 48.0%) to 49.9% (49.6% to 50.2%) and 1.3 pps in sterilization use from 36.0% (35.8% to 36.3%) to 37.3% (37.0% to 37.6%).

To investigate further the interviewer effects compared to those of other survey strata in the two NFHS rounds, we estimate a four-level cross-classified random intercept logit model for each of the two outcomes. This analysis tests the presence and relative size of interviewer effects in the two survey rounds and distinguished from state, community (PSU), and residual/woman level effects. A key challenge in estimating interviewer effects is the need to distinguish them from those at the community and other levels. Because of India's population heterogeneity, natural variation in social organization, beliefs, and customs will exist across and within states and communities. The results of the multilevel multivariable models are presented in Table 5. The woman's background covariates used in the models in Table 2 are also used to adjust the estimated effects in Table 5's models. We assess the percent of variance explained by heterogeneity at the state, PSU or cluster, interviewer, and residual (or woman) levels, expecting to find them to be larger in NFHS-4 compared to NFHS-3. If detected, these results will suggest that the larger scale and complexity of NFHS-4 likely had a direct impact on the measurement of contraceptive use through variation in fieldwork implementation. Results for the null or bivariate model are shown first, followed by those for the random intercept or multivariable model, with NFHS-3 estimates in the top panel and NFHS-4 ones in the bottom panel. The ICC values measure the percentage of the total variance contributed by the each of the four levels.

To explain Table 5, we interpret the findings for NFHS-3. The null model for modern contraceptive use shows 0.46 is the variance contributed by the state compared to 0.51 in the multivariable model. Likewise, 0.31 is the variance contributed by the PSU, and 3.29 is the default value at the individual- (woman-) level variance, which are unchanged in the multivariable model. In terms of the ICC, the null model shows the state contributes 11.3%, the PSU 7.7%, interviewers 0%, and individual women 81% (= 100.0 – [11.3 + 7.7]) of the total variation.

The random intercept models for both outcomes generally show larger ICC values than the null model, justifying the inclusion of the covariates. We observe from the ICCs in the random intercept model that state heterogeneity is an important source of influence across the two outcomes and is greater in NFHS-4 compared to NFHS-3. For modern contraceptive use it accounts for 19.0% of the variance in modern contraceptive use in NFHS-4 and 12.2% in NFHS-3. For sterilization use, state heterogeneity accounts for 30.1% of the variance in NFHS-4 and for 21.4% in NFHS-3. PSU-level heterogeneity accounts for 10.4% of the variance in the NFHS-4 modern use model compared to 8.8% in the NFHS-3. For the female sterilization use models, it accounts for 13.0% of the variance in NFHS-4 compared to 11.7% in NFHS-3. Compared to state and PSU levels, interviewer-level heterogeneity has less influence. While it is negligible in NFHS-3 across both outcomes, it accounts for 8.5% and 7.5% of the variance in modern contraceptive and female sterilization use, respectively, in NFHS-4. State-level variance can be partially attributed to the performance of survey agencies uniquely assigned to and responsible for implementation in the different states. Their training, logistics, and operational effectiveness will also be reflected in the variance observed at the PSU levels. Some variation at the state and PSU levels will be due to inherent natural differences in local beliefs and norms about contraceptive use across these geographies, but by themselves these do not account for the high ICC values observed for the two levels.

4. Conclusions

In the decade between NFHS-3 and NFHS-4, there have been important changes in Indian women's education, their households' economic well-being, and family formation, continuing trends first monitored in 1992–1993 with NFHS-1. Contraceptive practices remain centered around the woman's lifetime decision to use permanent contraception, although spacing methods complement her birth planning. Because contraceptive use has risen to cover more than half the married female population by NFHS-3, the dip observed in modern contraceptive and sterilization method uses in round 4 is likely more than a statistical curiosity. The small reductions are empirically significant and inconsistent with the observed fertility reduction in actual and wanted levels.

Our analyses have focused on testing three hypotheses about the potential influences of survey design and implementation differences between NFHS-3 and NFHS-4, specifically survey round year, fieldwork phase, and interviewer experience. Exploratory analysis established higher odds of contraceptive and sterilization use in several populous states reported in later months after the survey began relative to the first month. With the survey round factor, we found the predicted probability of modern use would be 4.0 pps higher (48.5% compared to 44.5%) if all respondents experienced the round-specific conditions of NFHS-4 as opposed to NHFS-3, net of woman's age and other covariates, and 8.3 pps higher (37.5% versus 29.3%) for female sterilization. These differences are large in magnitude but opposite in direction to the observed changes. Further analysis showed the NFHS-4 round had lower adjusted odds of reported contraceptive use in models without individual background covariates but reversed to a positive influence once the woman's age was included. Younger women were recorded to be less likely to use contraception, whether out of choice or interaction with the field interviewer.

We then found that if all respondents were interviewed in phase 2, when interviewers were more experienced, the predicted probability of modern contraception use would be 6.1 pps higher and 2.0 pps higher for female sterilization. In addition, we found support for the third hypothesized factor – the influence of time since the survey was launched in the woman's state of residence, likely reflecting gained interviewer experience. We estimated an increase of about 3.0 pps in the monthly probability of modern contraceptive and female sterilization use. The findings from our analyses thus support the association of these three selected survey design factors with measured contraceptive use, net of variation in individual respondent characteristics.

We found support for our second hypothesis of larger interviewer effects in NFHS-4 than NFHS-3 by calculating deviance residuals from the overall means across interviewers for both outcomes. A large proportion of negative deviance residuals can lower the estimate of contraceptive use. NFHS-4 has more negative and larger significant deviance residuals than NFHS-3. Problematic interviewers were identified and tended to be from the same outlier states observed in the exploratory analyses. Eliminating their interviews and re-estimating the outcome indicators led to a 2.0 pps increase in mCPR and a 1.3 pps increase in female sterilization use.

Last, we found through multilevel cross-classified random intercept logit models support for our third hypothesis of larger variations at the interviewer, state, and PSU levels in NFHS-4 than NFHS-3. State-level clustered effects are in part attributable to the assigned field agencies and were followed in size by clustering within interviewers, signaling that the number of interviews assigned to each interviewer and their own characteristics could have affected the quality of contraceptive measurement. As shown in Appendix Table 1, the number of PSUs covered in NFHS-4 compared with NFHS-3 increased more than seven times, the number of interviewing teams and the number of interviewers more than three times, the average number of interviews conducted by an interviewer approximately doubled, and the average duration of the survey in a state increased from 4.6 to 7.8 months (see also Singh, Kumar, and Arnold 2022). The considerable increase in the NFHS-4 workload and field duration could have led to higher fatigue among the field interviewers, field supervisors, field agencies, and the coordinating agency, weakening monitoring and supervision of fieldwork in NFHS-4. The replacement of the field editing with the computer-assisted personal interview may also have inhibited timely detection of unusual response patterns.

The shift in sample size from previous NFHS rounds to a round five times larger in NFHS-4 likely impacted attention on the data gathered on contraceptive use by interviewers. This study's findings will be informed by replicating the analysis with data from 2019–2021 NFHS-5 when it is publicly released. NFHS-5 shared the same coverage objectives as NFHS-4 (i.e., to estimate indicators at the national, state, and district levels). Its sample size and field procedures were similar to NFHS-4, and the same questions to measure contraceptive use were employed.

Perhaps most pertinent for future research of Indian contraceptive use dynamics is the issue of recall bias, whether the female respondents accurately recalled and reported their sterilization and other contraceptive method experiences. Choi et al.'s (2019) study finding sterilization underreporting in Rajasthan notes the need for an explicit survey question asking the respondent if she had ever been sterilized. This is planned for NFHS-6.

Research has shown that mental anchoring (i.e., using recent behaviors as a cognitive heuristic for current behaviors) can influence recall in surveys (Godlonton, Hernandez, and Murphy 2018). Older women who have been sterilized many years earlier may not report themselves as currently using contraception or perceive their use of traditional methods as doing something to avoid or delay a pregnancy. Applied psychologists have in recent decades investigated the cognitive and communicative processes underlying survey responses (Schwarz 2007), including the roles of interpretation and memory for estimation and event reporting.

This study has taken advantage of an opportunity to investigate the role of survey design and field implementation factors with two large-scale national surveys that have been technically framed and consistently overseen by one organization. The two survey rounds share similar survey measurement and procedures and differ by the substantial expansion in the most recent round's sample size. It has focused on two types of nonsampling errors – field implementation and measurement error – to estimate their effects in assessing levels of modern contraceptive use. In addition to the strength of the large sample power, the analysis also offers insights into the influences of state, sample cluster, and interviewer-level variation on contraceptive reporting with deviance residual analysis and advanced multilevel modeling. At the same time, the analysis is limited by not exploring other potential sources of survey bias or causes behind women's contraceptive decision-making. Most relevant is that the study was not able to directly assess women's use of pregnancy termination as an alternative to contraceptive use. The absence of reliable reporting of abortion use leaves open the question that modern contraceptive use could be declining with the increased use of elective terminations.

However, it is the standardized, and therein consistency, of NFHS-derived measures that make its scale, scope, and temporality features of paramount importance for programmatic guidance. NFHS will continue to serve as a data resource that generates and meets expectations of quality and precision. It is essential that the survey accurately measures contraceptive behaviors so that any observed significant change is not the result of field implementation or measurement error. In the context of the Family Planning 2020 progress estimates, irregular fieldwork factors could have underestimated the number of new modern contraceptives by as much as 12 million women based on the estimated prevalence from this study. Across all four NFHS rounds, more than 1 million female respondents have volunteered information about their sexual and reproductive lives, as well as their roles as wives, mothers, and workers. This data endowment, and the ensuing knowledge generated, place a significant responsibility on, and accountability by, the communities of scholars, policymakers, and practitioners who are the direct beneficiaries.

5. Acknowledgments

This study was supported in part by grants from the Bill & Melinda Gates Foundation under grant number OPP117_01 to the Bill & Melinda Gates Institute for Population and Reproductive Health at the Johns Hopkins Bloomberg School of Public Health and grant number OPP1179208 to the International Institute for Population Sciences. The views expressed in the study do not necessarily reflect those of the foundation. The authors thank the anonymous reviewers for their helpful comments.

References

Achilles, S., Mhlanga, F., Musara, P., Poloyac, S., Chirenje, Z., and Hillier, S. (2018). Misreporting of contraceptive hormone use in clinical research participants. *Contraception* 97(4): 346–353. doi:10.1016/j.contraception.2017.09.013.

Anderson, J.E. and Cleland, J. (1984). The World Fertility Survey and Contraceptive Prevalence Surveys: A comparison of substantive results. *Studies in Family Planning* 15(1): 1–13. doi:10.2307/1965479.

Bignami-Van Assche, S., Reniers, G., and Weinreb, A. (2003). An assessment of the KDICP and MDICP data quality: Interviewer effects, question reliability and sample attrition. *Demographic Research* S1(2): 31–76. doi:10.4054/DemRes.2003.S1.2.

Blom, A. and Korbmacher, J. (2013). Measuring interviewer characteristics pertinent to social surveys: A conceptual framework. *Survey Methods: Insights from the Field*. https://surveyinsights.org/?p=817.

Bora, J. and Saikia, N. (2018). Neonatal and under-five mortality rate in Indian districts with reference to Sustainable Development Goal 3: An analysis of the National Family Health Survey of India (NFHS), 2015–2016. *PLOS One* 13(7): e0201125. doi:10.1371/journal.pone.0201125.

Bradley, S.E.K., Schwandt, H., and Khan, S. (2009). Levels, trends and reasons for contraceptive discontinuation. (DHS Analytical Studies 20). Calverton: ICF Macro.

Callahan, R. and Becker, S. (2012). The reliability of calendar data for reporting contraceptive use: Evidence from rural Bangladesh. *Studies in Family Planning* 43(3): 213–222. doi:10.1111/j.1728-4465.2012.00319.x.

Chacko, E. (2001). Women's use of contraception in rural India: A village-level study. *Health and Place* 7(3): 197–208. doi:10.1016/S1353-8292(01)00009-0.

Choi, Y.J., Khanna, A, Zimmerman, L., Radloff, S., Zachary, B., and Ahmed, D. (2019). Reporting sterilization as a current contraceptive method among sterilized women: Lessons learned from a population with high sterilization rates, Rajasthan, India. *Contraception* 99: 131–136. doi:10.1016/j.contraception.2018.10.006.

Corsi, D., Neuman, M., Finlay, J., and Subramanian, S. (2012). *International Journal of Epidemiology* 41(6): 1602–1613. doi:10.1093/ije/dys184.

Davis, R.E., Couper, M.P., Janz, N.K., Caldwell, C.H., and Resnicow, K. (2010). Interviewer effects in public health surveys. *Health Education Research* 25(1): 14–26. doi:10.1093/her/cyp046.

Fabic, M., Choi, Y. and Bird, S. (2012). A systematic review of Demographic and Health Surveys: Data availability and utilization for research. *Bulletin of the World Health Organization* 90: 604–612. doi:10.2471/BLT.11.095513.

Fisher, A. and Way, A. (1988). The Demographic and Health Surveys Program: An overview. *International Family Planning Perspectives* 14(1): 15–19. doi:10.2307/2947652.

Goldman, N., Moreno, L., and Westoff, C. (1989). Collection of survey data on contraception: An evaluation of an experiment in Peru. *Studies in Family Planning* 20(3): 147–157. doi:10.2307/1966569.

Greenleaf, A.R., Gadiaga, A., Guiella, G., Turke, S., Battle, N., Ahmed, S., and Moreau, C. (2020). Comparability of modern contraceptive use estimates between a face-to-face survey and a cellphone survey among women in Burkina Faso. *PLoS ONE* 15(5): e0231819. doi:10.1371/journal.pone.0231819.

Godlonton, S., Hernandez, M., and Murphy, M. (2018). Anchoring bias in recall data: Evidence from Central America. *American Journal of Agricultural Economics* 100(2): 479–501. doi:10.1093/ajae/aax080.

Groves, R. (2004). The interviewer as a source of survey measurement error. In: Groves, R. (ed.). *Survey errors and survey costs* (2nd ed.) New York: Wiley-Interscience: 357–406. doi:10.1002/0471725277.ch8.

Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer, E., and Tourangeau, R. (2009). *Survey methodology (*2nd ed.*).* Hoboken: John Wiley and Sons.

Huber, L.R.B., Broel, E., Mitchelides, A., Dmochowski, J., Dulin, M., and Scholes, D. (2013). Comparison of prospective daily diaries and retrospective recall to measure oral contraceptive adherence. *Contraception* 88(4): 492–497. doi:10.1016/j.contraception.2013.02.007.

Ichikawa, L., Hubbard, R., Operskalski, B., LaCroix, A., Ott, S., and Scholes, D. (2015). A comparison of self-reported oral contraceptive use and automated pharmacy data in perimenopausal and early postmenopausal women. *Annals of Epidemiology* 25(1): 55–59. doi:10.1016/j.annepidem.2014.10.009.

International Institute for Population Sciences (2001). Reproductive and Child Health Project, Rapid Household Survey (Phase I and II), 1998–1999. Mumbai. http://rchiips.org/pdf/rch1/National_Report_RCH-1.pdf.

IIPS and ICF (2017). National Family Health Survey (NFHS-4), 2015–16: India. Mumbai: International Institute Population Sciences (IIPS).

IIPS and Macro International (2007). National Family Health Survey (NFHS-3), 2005–06: India: Volume I. Mumbai: IIPS.

Jha, P., Kesler, M., Kumar, R., Ram, F., Ram, U., Aleksandrowicz, L., Bassani, D., Chandra, S., and Banthia, J. (2011). Trends in selective abortions of girls in India: Analysis of national representative birth histories from 1990 to 2005 and census data from 1991 to 2011. *The Lancet* 377(9781): 1921–1928. doi:10.1016/S0140-6736(11)60649-1.

Kasprzyk, D. (2005). Measurement errors in household surveys: Sources and measurement. In: *Household sample surveys in developing and transition countries*. New York: United Nations: 171–198.

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.

Lewis, G. (1983). The Contraceptive Prevalence Survey Project: Content and status. *Population Index* 49(2): 189–198. doi:10.2307/2737212.

Liu, M. and Wang, Y. (2016). Interviewer gender effect on acquiescent response style in 11 Asian countries and societies. *Field Methods* 28(4): 327–344. doi:10.1177/1525822X15623755.

MacQuarrie, K., Winfrey, W., Meijer-Irons, J., and Morse, A. (2018). Consistency of reporting of terminated pregnancies in DHS calendars. (DHS Methodological Reports 25). Rockville: ICF.

McCullagh, P. and Nelder, J. (1989). *Generalized linear models* (2nd edition). London: Chapman and Hall. doi:10.1007/978-1-4899-3242-6.

Morris, L. (2000). History and current status of reproductive health surveys at CDC. *American Journal of Preventive Medicine* 19(1Suppl): 31–34. doi:10.1016/S0749-3797(00)00176-8.

New, J., Cahill, N., Stover, J., Gupta, Y., and Alkema, L. (2017). Levels and trends in contraceptive prevalence, unmet need, and demand for family planning for 29 states and union territories in India: A modelling study using the Family Planning Estimation Tool. *Lancet Global Health* 5(3): e350–e358. doi:10.1016/S2214-109X(17)30033-5.

Pal, S. and Makepeace, G. (2003). Current contraceptive use in India: Has the role of women's education been overemphasized? *European Journal of Development Research* 15: 146–169. doi:10.1080/09578810312331287415.

Pradhan, M. and Dwivedi, L. (2019). Changes in contraceptive use and method mix in India: 1992–93 to 2015–16. *Sexual and Reproductive Healthcare* 19: 56–63. doi:10.1016/j.srhc.2018.12.006.

Pullum, T. (2019). Strategies to assess the quality of DHS data. (DHS Methodological Reports 26). Rockville: ICF.

Pullum, T., Juan, C., Khan, N., and Staveteig, S. (2018). The effect of interviewer characteristics on data quality in DHS surveys. (DHS Methodological Reports 24). Rockville: ICF.

Roy, T.K., Porwal, A., and Acharya, R. (2021). Is contraceptive prevalence declining in India? An appraisal based on NFHS-4. *Economic and Political Weekly* 56(41): 32–37.

Schwarz, N. (2007). Cognitive aspects of survey methodology. *Applied Cognitive Psychology* 21: 277–287. doi:10.1002/acp.1340.

Singh, A., Kumar, K., and Arnold, F. (2022). How interviewers affect responses to sensitive questions on the justification for wife beating and refusal to have conjugal sex, and domestic violence in India. *Studies in Family Planning* 53(2): 259–279. doi:10.1111/sifp.12193.

Singh, S., Remez, L., and Tartaglione, A. (eds.) (2010). *Methodologies for estimating abortion incidence and abortion-related morbidity: A review*. New York: Guttmacher Institute; and Paris: International Union for the Scientific Study of Population.

Singh, S., Shekhar, C., Acharya, R., Moore, A., Stillman, M., Pradhan, M., Frost, J., Sahoo, H., Algarajan, M., Hussain, R., Sundaram, A., Vlassoff, M., Kalyanwala, S., and Browne, A. (2018). The incidence of abortion and unintended pregnancy in India, 2015. *Lancet Global Health* 6: e111–120. doi:10.1016/S2214-109X(17)30453-9.

Smith, C., Edwards, P., and Free, C. (2018). Assessing the validity and reliability of self-report data on contraception use in the Mobile Technology for Improved Family Planning (MOTIF) randomised controlled trial. *Reproductive Health* 15: 50. doi:10.1186/s12978-018-0494-7.

Sprehe, T. (1974). The World Fertility Survey: An international program of fertility research. *Studies in Family Planning* 5(2): 35–41. doi:10.2307/1965382.

Stover, J. and Sonneveldt, E. (2017). Progress toward the Goals of FP2020. *Studies in Family Planning* 48(1): 83–88. doi:10.1111/sifp.12014.

Strickler, J., Magnani, R., McCann, H., and Brown, L. (1997). The reliability of reporting contraceptive behavior in DHS calendar data: Evidence from Morocco. *Studies in Family Planning* 28(1): 44–63. doi:10.2307/2137970.

Stuart, G. and Grimes, D. (2009). Social desirability bias in family planning studies: A neglected problem. *Contraception* 80(2): 108–112. doi:10.1016/j.contraception.2009.02.009.

Tourangeau, R. and Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin* 133(5): 859–883. doi:10.1037/0033-2909.133.5.859.

Tsui, A., Cardona, C., Srivatsan, V., OlaOlorun, F., Omoluabi, E., Akilimali, P., Gichangi, P., Thiongo, M., PMA Agile Team, Radloff, S., and Anglewicz, P. (2021). Is client reporting on contraceptive use always accurate? Measuring consistency and change with a multi-country study. *Studies in Family Planning* 52(3): 361–382. doi:10.1111/sifp.12172.

West, B. and Blom, A. (2017). Explaining interviewer effects: A research synthesis. *Journal of Survey Methodology and Statistics* 5(2): 175–211.

Yang, M. and Yu, R. (2008). The interviewer effect when there is an education gap with the respondent: Evidence from a survey on biotechnology in Taiwan. *Social Science Research*: 37(4): 1321–1331. doi:10.1016/j.ssresearch.2008.05.008.

Zimmerman, L., Olson, H., PMA Principal Investigators Group, Tsui, A., and Radloff, S. (2017). PMA2020: Rapid turn-around survey data to monitor family planning service and practice in ten countries. *Studies in Family Planning* 48(3): 293–303. doi:10.1111/sifp.12031.

Appendix

Table A-1: Selected survey design features of NFHS-3 and NFHS-4

Feature	NFHS-3	NFHS-4
Survey years	2005–2006	2015–2016
Fieldwork dates	November 2005–August 2006	January 2015–December 2016
Phase 1	November 2005–May 2006	January 2015–December 2015
Phase 2	June 2006–August 2006	January 2016–December 2016
Average duration (mos)	4.6	7.8
# of states	29	
# of union territories	6	
# of districts	640	
# of field agencies	18	14
Private	13	11
Population research centers	5	3
# of primary sampling units	3,850	28,521
# of field teams*	264	875
# of interviewers	916	2,734
# of female	632	2,596
Sample size		
# of females age 15 to 49	124,385	699,686
Response rate (%)	95	97
Average total # of interviews/interviewer	136	256
Other differences	HIV testing; men's interview	No field editor on field team; CAPI mode; added biomarkers (blood pressure, blood glucose)

Notes: * Comprised of one field supervisor, one field editor, three female interviewers, one male interviewer, two health investigators; one training event to be organized for every 13 teams (approximately 100 participants). CAPI = computer-assisted personal interview.

Table A-2a: State-specific odds ratios for woman reporting modern contraceptive use by months since interviewing began in state, adjusted for urban/rural residence: NFHS-3 and NFHS-4

State/territory	NFHS-3 (2005–2006) Months since interviewing began in the state					State/territory	NFHS-4 (2015–2016) Months since interviewing began in the state								
	1	2	3	4	5		1	2	3	4	5	6	7	8	9
Jammu and Kashmir	1.00	0.74	0.80	0.71		Andaman and Nicobar Islands	1.00	0.88	1.12	1.94					
Himachal Pradesh	1.00	1.26	1.19			Andhra Pradesh	1.00	0.91	0.75						
Punjab	1.00	0.98	1.21	1.24		Arunachal Pradesh	1.00	0.81	1.04	0.39	0.84	2.54	4.62	4.22	
Uttaranchal	1.00	0.76	0.86	1.22		Assam	1.00	1.24	1.19	1.47	1.70				
Haryana	1.00	0.78	0.71			Bihar	1.00	2.46	5.75	6.53	7.91	9.87			
Delhi	1.00	1.13	1.06	1.53		Chandigarh	1.00	1.15							

Table A-2a: (Continued)

State/territory	NFHS-3 (2005–2006) Months since interviewing began in the state				
	1	2	3	4	5
Rajasthan	1.00	0.74	0.68	0.92	2.11
Uttar Pradesh	1.00	1.07	1.07	1.00	0.78
Bihar	1.00	1.07	1.11	1.35	
Sikkim	1.00	1.15	1.34	1.83	
Arunachal Pradesh	1.00	0.79	0.72	0.96	
Nagaland	1.00	1.03	0.81	1.03	2.81
Manipur	1.00	0.87	0.92	0.31	
Mizoram	1.00	0.83	1.12		
Tripura	1.00	1.00	0.77	0.70	
Meghalaya	1.00	0.81	0.59	0.51	1.30
Assam	1.00	0.72	0.72	0.80	0.75
West Bengal	1.00	1.02	1.12	0.91	0.78
Jharkhand	1.00	0.98	0.96	1.36	
Orissa	1.00	0.69	0.66	0.63	0.75
Chhattisgarh	1.00	0.79	0.73	0.79	
Madhya Pradesh	1.00	0.93	1.41	1.89	
Gujarat	1.00	1.18	0.96	1.24	
Maharashtra	1.00	1.15	1.30	1.03	
Andhra Pradesh	1.00	1.12	1.00	0.73	0.95
Karnataka	1.00	0.64	0.56	0.54	
Goa	1.00	1.22	1.29	1.47	
Kerala	1.00	0.77	1.35	1.55	1.67
Tamil Nadu	1.00	0.76	0.64	0.57	

State/territory	NFHS-4 (2015–2016) Months since interviewing began in the state								
	1	2	3	4	5	6	7	8	9
Chhattisgarh	1.00	1.29	1.85	2.13	1.65	1.15			
Dadra and Nagar Haveli	1.00	1.67	1.42						
Daman and Diu	1.00	2.77	2.89						
Goa	1.00	1.50	2.08	3.39					
Gujarat	1.00	0.95	1.14	1.15	1.16	1.20			
Haryana	1.00	3.36	7.29	7.17	8.99				
Himachal Pradesh	1.00	1.62	1.64	2.88	2.17	2.07	1.81		
Jammu And Kashmir	1.00	2.27	2.38	2.34	2.52	3.05	2.67	2.94	
Jharkhand	1.00	1.20	0.89	1.53	2.21	1.93	1.74	1.52	1.26
Karnataka	1.00	2.48	3.82	4.87	4.19	4.25			
Kerala	1.00	1.51	1.20	1.43	1.03				
Lakshadweep	1.00	0.64	0.20						
Madhya Pradesh East	1.00	1.42	1.42	1.67	2.88	3.39	2.69		
Madhya Pradesh West	1.00	5.74	7.09	6.01	5.59	5.58			
Maharashtra	1.00	1.44	1.65	1.72	2.16	2.11			
Manipur	1.00	0.95	0.85	1.01	1.51	1.59	1.46	1.26	
Meghalaya	1.00	1.14	2.32	2.76	1.64	1.74			
Mizoram	1.00	0.97	0.74	0.55	0.74	1.09	0.72	0.63	0.71
Nagaland	1.00	1.38	1.29	0.83	1.35	1.05	0.57		
Delhi	1.00	1.33	1.59	1.78	1.61	1.52	0.98		
Odisha	1.00	1.77	3.11	3.03	3.02	3.03	2.26		
Puducherry	1.00	0.70							
Punjab	1.00	1.16	1.57	0.96	0.70	0.80			
Rajasthan	1.00	1.02	1.16	0.91	0.90	1.13	1.14		
Sikkim	1.00	1.78	2.16	3.83	3.93	5.18			
Tamil Nadu	1.00	1.58	3.03	3.76	4.24				
Tripura	1.00	0.98	1.12	1.10	0.89	0.90			
Uttar Pradesh East	1.00	1.69	4.27	4.53	4.58	5.55			
Uttar Pradesh Central	1.00	0.93	1.01	1.52	2.21	2.47	2.46	1.83	
Uttar Pradesh West	1.00	1.78	1.97	1.97	2.25	1.90	3.44		
Uttarakhand	1.00	1.10	2.01	3.36	4.14	3.47	2.70		
West Bengal	1.00	1.72	2.16	2.32	2.18	2.37			

Table A-2b: State-specific odds ratios for woman reporting female sterilization use by months since interviewing began in state, adjusted for urban/rural residence: NFHS-3 and NFHS-4

NFHS-3 (2005–2006)

	Months since interviewing began in the state				
State/territory	1	2	3	4	5
Jammu and Kashmir	1.00	0.93	0.79	0.59	
Himachal Pradesh	1.00	0.92	1.04		
Punjab	1.00	1.07	1.21	1.19	
Uttaranchal	1.00	1.08	1.30	1.57	
Haryana	1.00	1.08	0.92		
Delhi	1.00	1.34	1.11	1.25	
Rajasthan	1.00	0.69	0.58	0.82	1.24
Uttar Pradesh	1.00	1.03	1.14	1.14	0.96
Bihar	1.00	1.14	1.15	1.58	
Sikkim	1.00	1.37	1.66	1.86	
Arunachal Pradesh	1.00	1.43	1.41	1.21	
Nagaland	1.00	1.20	0.89	1.36	2.71
Manipur	1.00	0.62	0.69	0.20	
Mizoram	1.00	0.87	1.34		
Tripura	1.00	0.67	0.75	0.59	
Meghalaya	1.00	0.49	0.68	0.48	1.83
Assam	1.00	0.48	0.36	0.79	0.47
West Bengal	1.00	1.23	1.40	0.95	0.74
Jharkhand	1.00	1.30	1.23	1.84	
Orissa	1.00	0.69	0.59	0.63	0.61
Chhattisgarh	1.00	0.67	0.69	0.76	
Madhya Pradesh	1.00	1.05	1.62	1.69	
Gujarat	1.00	1.14	0.90	1.11	
Maharashtra	1.00	1.01	0.98	0.83	

NFHS-4 (2015–2016)

	Months since interviewing began in the state								
State/territory	1	2	3	4	5	6	7	8	9
Andaman and Nicobar Islands	1.00	0.69	0.84	1.57					
Andhra Pradesh	1.00	0.94	0.81						
Arunachal Pradesh	1.00	0.99	2.03	1.98	2.52	4.87	8.72	8.37	
Assam	1.00	0.97	0.74	0.65	0.59				
Bihar	1.00	3.22	7.99	9.16	11.05	14.12			
Chandigarh	1.00	0.86							
Chhattisgarh	1.00	1.30	1.68	2.05	1.67	1.18			
Dadra and Nagar Haveli	1.00	1.88	1.86						
Daman and Diu	1.00	2.28	3.00						
Goa	1.00	1.03	1.03	2.36					
Gujarat	1.00	0.92	1.04	0.98	0.89	0.94			
Haryana	1.00	2.98	4.84	4.47	4.37				
Himachal Pradesh	1.00	1.64	1.50	2.77	1.67	1.15	1.82		
Jammu And Kashmir	1.00	1.80	1.83	2.03	1.96	2.43	2.09	2.52	
Jharkhand	1.00	1.08	0.72	1.36	1.97	1.61	1.51	1.31	0.92
Karnataka	1.00	2.60	4.09	5.46	4.76	5.01			
Kerala	1.00	1.51	1.24	1.47	1.12				
Lakshadweep	1.00	1.69	0.23						
Madhya Pradesh East	1.00	1.52	1.54	1.79	2.86	3.59	2.50		
Madhya Pradesh West	1.00	5.98	6.73	5.68	5.76	5.73			
Maharashtra	1.00	1.19	1.28	1.46	1.49	1.60			
Manipur	1.00	1.43	0.82	0.74	1.23	2.05	2.06	2.28	
Meghalaya	1.00	1.64	2.23	1.76	1.95	1.99			
Mizoram	1.00	1.13	1.11	0.92	0.90	0.96	0.76	0.77	1.03

Table A-2b: (Continued)

	NFHS-3 (2005–2006)						NFHS-4 (2015–2016)									
	Months since interviewing began in the state							Months since interviewing began in the state								
State/ territory	1	2	3	4	5	State/territory		1	2	3	4	5	6	7	8	9
Andhra Pradesh	1.00	1.33	1.15	0.92	1.09	Nagaland		1.00	2.64	2.41	1.73	2.39	2.63	1.66		
Karnataka	1.00	0.65	0.63	0.52		Delhi		1.00	2.00	3.12	3.09	2.82	2.55	2.44		
Goa	1.00	1.38	1.57	1.39		Odisha		1.00	1.67	2.73	3.15	3.24	3.29	2.85		
Kerala	1.00	0.79	1.32	1.56	1.84	Puducherry		1.00	0.67							
Tamil Nadu	1.00	0.78	0.66	0.60		Punjab		1.00	1.05	1.47	1.07	0.68	0.83			
						Rajasthan		1.00	0.97	1.13	0.81	0.78	1.06	1.04		
						Sikkim		1.00	1.55	1.39	2.50	2.29	2.13			
						Tamil Nadu		1.00	1.57	2.96	3.89	4.27				
						Tripura		1.00	1.10	0.95	0.95	0.81	0.79			
						Uttar Pradesh East		1.00	1.90	5.39	6.03	5.66	7.96			
						Uttar Pradesh Central		1.00	0.93	1.15	1.84	2.60	3.19	3.08	1.68	
						Uttar Pradesh West		1.00	2.08	1.81	1.53	1.69	1.35	1.28		
						Uttarakhand		1.00	1.30	2.48	5.63	7.82	6.30	3.67		
						West Bengal		1.00	1.68	1.57	2.25	2.64	3.03			

Table A-3a: Results of multivariable logit regression of woman's reported modern contraceptive use on selected survey design factors, adjusted for her background covariates: Combined sample of NFHS-3 and NFHS-4 currently married female respondents age 15–49

Covariate		Adjusted odds ratios/95% CI					
	Value	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Currently married women (*N*)	582,144						
Survey design/implementation							
NFHS round	3 (Ref)						
	4	0.91 (0.88,0.94)	0.85 (0.83,0.88)	1.47 (1.42,1.52)	1.21 (1.17,1.26)		
Survey months elapsed since interview started		1.07 (1.06,1.08)	1.14 (1.13,1.15)	1.15 (1.14,1.16)	1.16 (1.15,1.17)		
Fieldwork phase	First (Ref)						
	Second	0.87 (0.86,0.89)	1.28 (1.24,1.32)	1.35 (1.3,1.39)	1.35 (1.3,1.4)		
State dummy/fixed effect		No	Yes	Yes	Yes	Yes	Yes
Respondent background							
Year of birth	C			0.93 (0.93,0.93)	0.96 (0.96,0.96)	0.94 (0.94,0.94)	0.97 (0.97,0.97)
Educated	None = 0 (Ref)						
	Any = 1				1.17 (1.14,1.2)		1.15 (1.12,1.17)
Poorest	None = 0 (Ref)						
	Yes = 1				0.65 (0.63,0.66)		0.6 (0.59,0.62)
Richest	None = 0 (Ref)						
	Yes = 1				1.09 (1.06,1.12)		1.14 (1.1,1.17)
Urban	None = 0 (Ref)						
	Yes = 1				0.99 (0.96,1.02)		1.01 (0.98,1.03)
Scheduled caste/tribes	None = 0 (Ref)						
	Yes = 1				0.95 (0.93,0.98)		0.96 (0.94,0.99)
Hindu religion	None = 0 (Ref)						
	Yes = 1				1.62 (1.57,1.67)		1.61 (1.56,1.66)
Number of sons ever born	C				1.19 (1.18,1.2)		1.2 (1.19,1.21)
Number of daughters ever born	C				1.7 (1.68,1.72)		1.71 (1.69,1.73)
Desires more sons than daughters	None = 0 (Ref)						
	Yes = 1				0.88 (0.86,0.9)		0.87 (0.85,0.89)

Table A-3b: Results of multivariable logit regression of woman's reported current use of female sterilization on selected survey design factors, adjusted for her background covariates: Combined sample of NFHS-3 and NFHS-4 currently married female respondents age 15–49

Covariate		Adjusted odds ratios/95% CI					
	Value	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Currently married women (*N*)	582,144						
Survey design/implementation							
NFHS round	3 (Ref)						
	4	0.92 (0.89,0.95)	0.84 (0.81,0.87)	1.82 (1.75,1.89)	1.66 (1.59,1.74)		
Survey months elapsed since interview started	C	1.03 (1.02,1.04)	1.13 (1.12,1.14)	1.15 (1.14,1.16)	1.18 (1.16,1.19)		
Fieldwork phase	First (Ref)						
	Second	0.76 (0.74,0.78)	1.04 (1.01,1.08)	1.1 (1.06,1.15)	1.12 (1.07,1.17)		
State dummy/fixed effect			Yes	Yes	Yes	Yes	Yes
Respondent background							
Year of birth	C			0.91 (0.91,0.91)	0.93 (0.93,0.94)	0.92 (0.92,0.92)	0.94 (0.94,0.95)
Educated	None = 0 (Ref)						
	Any = 1				0.97 (0.94,0.99)		0.94 (0.92,0.97)
Poorest	None = 0 (Ref)						
	Yes = 1				0.72 (0.69,0.74)		0.64 (0.62,0.66)
Richest	None = 0 (Ref)						
	Yes = 1				0.83 (0.8,0.85)		0.9 (0.87,0.92)
Urban	None = 0 (Ref)						
	Yes = 1				0.82 (0.79,0.85)		0.82 (0.8,0.85)
Scheduled caste/tribes	None = 0 (Ref)						
	Yes = 1				1.02 (0.99,1.05)		1.04 (1.01,1.07)
Hindu religion	None = 0(Ref)						
	Yes = 1				2.19 (2.1,2.28)		2.18 (2.09,2.27)
Number of sons ever born	C				1.19 (1.18,1.2)		1.2 (1.19,1.21)
Number of daughters ever born	C				1.8 (1.78,1.83)		1.83 (1.8,1.85)
Desires more sons than daughters	None = 0 (Ref)						
	Yes = 1				0.89 (0.87,0.92)		0.88 (0.86,0.9)