Measuring open defecation in India using survey questions: Evidence from a randomized survey experiment

Sangita Vyas, Nikhil Srivastav, Divya Mary, Neeta Goel, Sujatha Srinivasan, Ajaykumar Tannirkulam, Radu Ban, Dean Spears, Diane Coffey

Affiliations
Economics, University of Texas at Austin, Austin, TX, USA (S Vyas MPA, D Spears PhD);
Population Research Center, University of Texas at Austin, Austin, TX, USA (S Vyas MPA, D Spears PhD, D Coffey PhD); **LBJ School of Public Affairs, University of Texas at Austin, Austin, TX, USA** (N Srivastav BA); **IFMR Lead, Chennai, India** (D Mary MS, S Srinivasan MBA, A Tannirkulam PhD); **International Initiative for Impact Evaluation (3ie), New Delhi, India** (N Goel PhD); **Bill & Melinda Gates Foundation** (R Ban PhD); **Indian Statistical Institute, Delhi, India** (D Spears PhD, D Coffey PhD).

Correspondence to:
Sangita Vyas, Economics, University of Texas at Austin, Austin, TX 78712, USA;
sangita.vyas@utexas.edu

Abstract

Rural India is home to more than half of the world's open defecation. India's most recent Demographic and Health Survey (DHS) estimates that 54% of households in rural India defecated in the open in 2015-2016. This measure is based on a household-level question that asks about the behavior of everyone in the household in one question. Yet, studies in India find substantial open defecation among individuals living in households with latrines, suggesting that measures based on household-level questions likely underestimate true open defecation in rural India. We conducted a randomized survey experiment in the rural parts of four Indian states in 2018. We randomly assigned households to receive one of two types of survey questions measuring sanitation behavior. 1,215 households were asked about latrine use or open defecation individually for every household member. 1,216 households were asked the question used in India's DHS, which asks, at the household level, what type of facility members of the household usually use. Using two methods for comparing the rate of open defecation measured by each question, the individual-level question found 20 to 21 (95% CI 16 to 25 for both estimates) percentage points more open defecation than the household-level question, among all households, and 28 to 29 (95% CI 22 to 35 for both estimates) percentage points more open defecation among households that received material or financial assistance to construct their latrines. This study is the first to provide evidence that individual-level questions find more open defecation than household-level questions. Because reducing open defecation in India will be essential to meeting the Sustainable Development Goals, and because exposure to open defecation has serious consequences for child mortality, health, and human capital development, it is essential to monitor its progress as accurately as possible.

Keywords: India, latrine use measurement, sanitation, survey methods, survey experiment

Introduction

Rural India is home to more than half of the world's open defecation.[1] Because the persistence of open defecation threatens gains in child health, the Sustainable Development Goals (SDGs) call for its elimination by 2030. Progress towards eliminating open defecation in rural India will be essential to meeting this goal. India's most recent Demographic and Health Survey (DHS), conducted between January 2015 and December 2016, estimates that 54% of households in rural India defecated in the open, down from 75% in the 2005-2006 DHS.[2,3] This measure is based on a household-level question that asks about the behavior of everyone in the household in the same question.

Recent evidence from studies in rural India suggests, however, that it is common for individuals living in households with latrines to nevertheless defecate in the open. In rural parts of five states of north India, Coffey et al. found that 21% of individuals defecated in the open, despite owning a latrine.[4] In rural Tamil Nadu, Yogananth and Bhatnagar report that 54% of respondents defecated in the open despite having a household latrine.[5] In Odisha, Barnard et al. found that less than half of members of households with latrines reported using their latrines at all times.[6]

Experimental studies of sanitation interventions have found similar results. Clasen et al. report on a sanitation intervention in Odisha and note open defecation among individuals living in households with latrines as a reason for not observing impacts on child health outcomes.[7] Patil et al. conducted a sanitation intervention in Madhya Pradesh and experienced a similar problem: modest increases in latrine coverage, and even more modest reductions in open defecation.[8] These findings suggest that open defecation among latrine-owning households is substantial. Since it is probable that latrine use is the socially desirable response to questions on sanitation behavior, measures based on household-level questions, such as those from the DHS, will likely underestimate true open defecation in rural India, particularly among households with latrines.

Because open defecation is an individual behavior, an individual-level survey question may be able to more accurately measure it compared to a household-level question, particularly among households with latrines. We designed this study to experimentally test this hypothesis in rural India. In this study, we aimed to investigate whether a balanced question about latrine use or open defecation for every member of a household finds different levels of open defecation compared to a household-level question.

This is the first study to experimentally vary survey methodology to improve upon the measurement of open defecation currently being used. Jenkins et al. study sanitation survey methods, and report on an index they develop for quantifying household excreta disposal.[9] Their study focuses on developing and piloting a new tool rather than comparing different measures. Sinha et al. compare answers to survey questions on latrine use behavior to measures of actual

behavior generated from passive latrine use monitors that were set up in the latrines of respondents, and find poor to moderate agreement between the two measures.[10] Our study contributes to this literature by comparing estimates of open defecation obtained from questions that can be administered in a large household survey, and highlighting potential sources of error in open defecation measurement.

Four years ago, the Government of India launched the Swachh Bharat Mission (SBM), a national sanitation campaign, which aims to eliminate open defecation in India by 2019. Many latrines have been constructed in rural India as a result of this campaign. Yet, the effect the SBM has had on reducing open defecation is still unknown. Because large reductions in open defecation in India are essential to meeting the SDGs, and because exposure to open defecation has serious consequences for child mortality, health, and human capital development, it is essential to monitor its progress as accurately as possible.

Methods

The study received ethical approval for research involving human subjects from the Institute for Financial Management and Research's Institutional Review Board in India and is registered in the Registry for International Development Impact Evaluations (RIDIE), number 5b55458ca54d1.

Sample: mostly latrine-owning households in rural parts of four states

This study uses as its sampling frame the study areas of 3ie's Promoting Latrine Use in Rural India Thematic Window. This Window has funded four independent research teams to conduct randomized control trials of distinct behavioral campaigns to promote the use of pit latrines in rural parts of Bihar, Gujarat, Karnataka, and Odisha. The study areas are spread across India, representing different contexts and varying levels of rural open defecation.

Because these trials focus on behavioral strategies rather than latrine construction, they are being carried out in villages that had high levels of coverage of pit latrines at baseline, relative to other rural parts of the same states. The households that comprise the sampling frame for this study are those that were identified as having a functional latrine in a census conducted by the research teams in the villages in which they were working. In all states except for Odisha, only households that had been excluded from the research teams' samples could be selected for this study. In Odisha, however, the sample selected for this study overlaps with the research teams' sample. We aimed to survey households that own latrines because we expect that an important source of misreporting of open defecation comes from individuals who do not use the functional latrines that their households own, and that other members of their households use.

Figure 1 describes the sample selection. The villages visited in each state were randomly selected from the full set of villages included in the 3ie research teams' studies. The full set of villages were selected by the research teams in collaboration with the implementation agencies they were working with. The research team led by Oxford Policy Management worked with World Vision in Bihar, the team led by London School of Hygiene and Tropical Medicine worked with Coastal Salinity Prevention Cell (CSPC) in Gujarat, Eawag worked with Wateraid in Karnataka, and Emory University worked with the Rural Welfare Institute (RWI) in Odisha. Data for our study were collected in 22 to 25 villages in each of the four study areas. In most areas, we sampled more villages than we actually visited in order to facilitate coordination with the research teams. 95 villages were visited in total.

Up to 40 households in each village were randomly assigned to receive the household or individual questions. In some villages, fewer than 40 households were assigned because fewer than 40 households met the eligibility criteria. The survey team visited as many assigned households as it could in these villages, given time constraints, and availability of household members. On average, the survey team interviewed 25 households per village. Data collection took place between March and July 2018.

In each state, data collection took place after the 3ie research teams had conducted their censuses and baselines, but before they had started their interventions. Since in all states, the households visited in this study were also visited when the research teams conducted their censuses, response bias may be a concern. This would not, however, impact the internal validity of this study since randomization generates equal response bias, in expectation, across treatment arms.

Random variation in latrine use questions at the household level

We randomly assigned the type of latrine use question administered in the survey at the household level. Roughly half of the households were assigned individual-level questions on latrine use. The other half were assigned a household-level question. One of the authors who was not involved in data collection carried out the randomization using a random number generator in Stata. Because of the nature of the study, it was not possible to blind the respondents or surveyors to the type of survey question administered in the survey. However, in the interest of data quality, respondents were not explicitly told that the primary purpose of the survey, which took approximately 25 minutes to complete, was to measure open defecation. Additionally, surveyors did not know which survey question had been assigned to a household until starting the survey with the household.

The individual-level questions asked for every household member age five or older whether the individual defecated in the open or used the latrine. The preface to this series of questions was: "I have seen that some people defecate in the open, and some people use the latrine. Now I want to

ask about where you and your family members defecate." Then, the surveyor asked the following question for each individual in the household, and coded the answer in a household roster: "The last time [*name of household member*] defecated, did [*name of household member*] defecate in the open or use the latrine?" The answer options included latrine, open, and somewhere else. Surveyors used the last option, which meant that the household member defecated in a bedpan, cloth, or other place, in less than 0.5% of cases. Because the priming statement and the behavior question include both open defecation and latrine use, they are balanced between the two different behaviors and could reduce social desirability bias. The surveyor asked household members who were participating in the interview directly about their behavior, and asked the main respondent, in most cases an adult female member of the household, to report on the behavior of their family members who were not participating in the interview.

The rest of the households were assigned the household-level question used in India's DHS: "What kind of toilet facility do members of your household usually use?"[11] The answer codes were also the same as those used in India's DHS: flush to piped sewer system, flush to septic tank, flush to pit latrine, flush to somewhere else, flush to don't know where, ventilated improved pit or biogas latrine, pit latrine with slab, pit latrine without slab or open pit, twin pit or composting toilet, dry toilet, and no facility or uses open space or field. We also included an individual-level question on mobile ownership or preferring vegetarian food versus non-vegetarian food in the surveys that asked the household-level question so that both types of surveys would take approximately the same amount of time to complete.

There are three main factors that differ between the two types of latrine use survey modules: the level of aggregation, the reference period, and the presence of a priming statement. Therefore, the differences in reported open defecation that we observe reflect the fact that the two sets of questions vary on all of these factors combined.

Statistical analyses

The outcome of interest is reported open defecation. For the household-level questions, we created a dummy variable that is equal to one if the response was "no facility or uses open space or field," and zero otherwise. The unit of observation for households assigned the individual-level question is the individual, while the unit of observation for those assigned the household-level question is the household. Therefore, in order to directly compare and test the significance of differences in reported open defecation between the two question types, we construct estimates that use the same unit of observation. We impute individual-level open defecation from responses to the household-level questions, and household-level open defecation rates from responses to the individual-level questions. To construct individual-level open defecation using the household-level questions, we assign the answer from the household question to each individual in the household. For instance, if a household consists of five members, and the

answer to the household-level question is "no facility/uses open space or field," then we assign all five household members to open defecation. Similarly, to construct household-level open defecation using the individual-level questions, we average the responses among individuals in the household, and assign this average as the household value. For example, if a household has five members, and three defecate in the open while two use the latrine, then the imputed household-level open defecation is 0.6. Our main analysis tests differences in measured open defecation by question type. We show pooled results, as well as results by study area.

We also conduct sub-group analyses. First, we investigate whether the difference in reported open defecation by question type depends on whether the latrine was constructed privately, or with assistance from the government or an NGO. In practice, assistance to construct latrines often comes from the government, but sometimes NGOs get involved in facilitating the implementation of the government program. As part of the SBM, the Government of India assists rural households to construct latrines either by providing financial assistance directly to households so they can construct their own latrines, or by local government officials constructing latrines for households. In the discussion that follows, we will describe a household as having "received help" if it received financial assistance or a partially or completely constructed latrine from the government or an NGO.

The Indian government promotes and constructs latrines with pits that are approximately 60 cubic feet.[12] However, many rural Indians aspire to construct latrines with pits that are much larger, so that they can avoid emptying the pit, a task that is associated with ritual pollution.[13–15,19] Compared to latrines constructed privately, those constructed with government help are less likely to be used due to concerns over purity and pit emptying. Since a large fraction of rural households are likely to receive latrines with help from the government as a result of the SBM, it is important to explore how much open defecation different types of latrine use questions measure, based on having received help to construct the latrine.

Second, we investigate whether the difference in measured open defecation between the two question types is statistically different for males compared to females. Sex differences are an important aspect to explore because observational studies have found consistently higher open defecation among latrine owners for males compared to females.[4] This observation could reflect greater demand for latrine use among females due to, for instance, greater psychosocial stress experienced when defecating in the open,[16] or it could be because of cultural norms that keep females in their reproductive years inside the home.

We cluster standard errors in all analyses at the village level. Statistical analyses were conducted using Stata (version 11).

Results

Table 1 shows summary statistics for households assigned the two types of latrine use questions. The total sample consisted of 2,431 households, which were approximately equally divided across question type in each of the study areas. The means presented in Table 1 show that there were no significant differences on measures relevant for latrine use between households assigned different types of latrine use questions. Households in both groups had approximately the same number of household members, fraction female, fraction Hindu, educational attainment of the household head, and asset ownership of 13 assets, including mobile phone, electricity, radio, television, fan, mosquito net, bicycle, motorcycle, car, chair, gas stove, pressure cooker, and shoes for everyone in the family. It is important for the validity of the results that the sample is balanced on religious composition, since studies have documented an association between household religion and latrine use.[17,18]

Table 1 also shows that, as the study design intended, most households in both groups had a latrine. Conditional on having a latrine, 64% of households had gotten help from the government or an NGO to build the latrine, and the average pit size was 180 cubic feet. It is important that the sample is balanced on these two characteristics because, compared to latrines constructed privately, those constructed with government help are less likely to be used because of concerns over purity and pit emptying.[13–15,19] Finally, among households with latrines, approximately 80% of them appeared to the surveyor to be in use upon observation.

Figure 2 presents the main results of the study; it shows means and 95% confidence intervals from the individual-level (shown as dashed red bars) and the household-level (shown as solid blue bars) questions, for the full dataset and for different sub-samples. Observations are individuals for the individual-level estimates, and households for the household-level estimates. The first set of estimates shown in the figure uses the full sample. The second, third, fourth, and fifth sets of estimates break the sample up by project area. In the full sample, and in each project area, the individual questions find more open defecation. The sixth set of estimates uses only households with latrines and finds similar results as the full sample. The size of the difference in reported open defecation between question types is large and consistent. No matter how the data are broken up, the individual-level, balanced latrine use questions find significantly higher rates of open defecation than the household-level question.

Table 2 shows actual and imputed open defecation rates, measured at the individual and household levels. Columns 1 and 3 are the estimates shown in Figure 2. Columns 2 and 4 show imputed open defecation rates, at the individual and household level, respectively. Imputed values are calculated based on the method described in the Statistical Analyses section. Column 5 shows the difference in the measured rate of open defecation between the individual and household questions when observations are individuals, and column 6 shows the same difference

when observations are households. In the full sample, the individual-level, balanced questions find 21 (95% CI 16 to 25) percentage points more open defecation than the household-level question when observations are individuals, and 20 (95% CI 16 to 25) percentage points more open defecation when observations are households. Notably, the individual-level questions measure consistently higher levels of open defecation in the full sample and in all subsamples, irrespective of how the difference is calculated. All differences are significant at the one percent level.

The seventh set of estimates shown in Figure 2 show reported open defecation from individual- and household-level questions among households that received help to construct their latrines. Comparing these estimates to the sixth set of estimates in the figure, which include all households with latrines, suggests that the household-level question underestimates open defecation by more among those that received help to construct their latrines, compared to those that did not. The seventh row in Table 2 shows that, using different methods for computing differences, individual-level questions measure 28 to 29 (95% CI 22 to 35 for both estimates) percentage points more open defecation than the household-level question, among those that received help to construct their latrines, compared to those that did not.

The first two columns in Table 3 test whether the difference in measured open defecation between the two question types is statistically different among households that received help to construct their latrines compared to households that did not receive help to construct their latrines. The first column in Table 3 uses individuals as observations, and the second column uses households. The coefficients in the third row represent the difference-in-differences estimate. The difference in measured open defecation between the two question types is 18 (95% CI 11 to 25) to 19 (95% CI 11 to 26) percentage points larger among households that received help to construct their latrines, compared to households that did not, depending on the method used to compute the difference. Among households that did not receive help to construct their latrines, the individual-level questions find 10 (95% CI 5 to 15) percentage points more open defecation than the household-level question.

The last column in Table 3 investigates whether the difference in measured open defecation between the two question types is statistically different among males compared to females. Using the full sample, individual questions find 4 (95% CI 2 to 7) percentage points more open defecation than the household question among males compared to females. Breaking the sample up by sex, the individual questions find 23 (95% CI 18 to 28) percentage points more open defecation than the household question among males, and 19 (95% CI 14 to 24) percentage points more among females. The household question underestimates open defecation by more among males compared to females.

Discussion

Our findings show that in our sample, individual-level, balanced questions find 20 to 21 (95% confidence intervals 16 to 25 for both estimates) percentage points more open defecation than the household-level question. This is both a statistically significant and practically important difference. This study presents compelling evidence that India's DHS, which provides the most recent nationally representative estimates of open defecation for rural India, and other surveys that ask household-level questions, greatly underestimate open defecation among households with latrines.

We also found that the difference in reported open defecation between the two question types is significantly greater for households that received help to construct their latrines compared to households that did not. Among households that received help to construct their latrines, the individual-level questions find 28 to 29 (95% confidence intervals 22 to 35 for both estimates) percentage points more open defecation than the household-level question. This suggests that as more and more households receive government assistance for a latrine through the Swachh Bharat Mission, household-level questions will become even less accurate at estimating open defecation.

The larger difference in measured open defecation between the two question types among households that received help compared to those that did not is likely arising from higher rates of open defecation among households that received help to construct their latrines. There are several reasons that could explain why households receiving assistance may be less likely to use their latrines. First, these households are likely to have lower demand for latrine use, compared to households that built latrines on their own. Second, households that received help have latrines with pits that are on average 150 cubic feet smaller than the pits of latrines in households that did not receive help. Because of concerns over ritual purity, rural Indians are less likely to use latrines with pits that need to be emptied manually every few years, like the latrines that are promoted and constructed by the government.[13-15,19] Whether only one, both, or other factors are leading to more open defecation among households that received help, the individual-level questions are better able to capture this open defecation than the household-level question.

We also find a statistically significant difference in reported open defecation between the two question types for males compared to females. The difference between the individual-level, balanced questions and the household-level question is 4 (95% CI 2 to 7) percentage points more for males compared to females. This supports evidence that, conditional on latrine ownership, males are more likely to defecate in the open compared to females. Individual-level questions understate the difference in open defecation between the two sexes by less than the household-level questions. Compared to the difference in reported open defecation by receiving help to construct the latrine, however, differences by sex are not as large.

Measuring open defecation at the individual level is feasible. Our survey team's experience suggests that adding the balanced, individual-level questions on use to a survey that already contains a household roster increases survey time by about 2 minutes, on average. Of course, the amount of time required to ask the individual-level questions depends on the number of individuals in the household.

The household-level question asked in the DHS also collects information on the types of latrines that households own, data that is still of great interest to researchers and practitioners. Therefore, individual questions on use, combined with a separate question on the types of latrines that households own, would satisfy both goals: evaluating latrine infrastructure, and measuring open defecation as accurately as possible.

A limitation of our study is that the samples from these project areas are not representative of the rural parts of the states they are in, nor are they collectively representative of rural India. The households in this study are much more likely to have a latrine than the average rural Indian household, and therefore, the individuals in this study are more likely to use a latrine. For this reason, the estimates presented here should not be considered as estimates of open defecation for any of the states, or the country as a whole. Rather, they show a large and significant difference in reported open defecation based on the type of question asked.

Measuring open defecation at the individual-level in a large household survey is doable and will provide a more accurate estimate of open defecation in rural India. Since reducing open defecation in India is important for meeting the SDGs, and since open defecation is an important factor contributing to poor health among children in India, it is important to measure its progress as accurately as possible.

References

1   WHO and Unicef. *Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines*. Geneva: World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), 2017.

2   International Institute for Population Sciences (IIPS) and ICF. National Family Health Survey (NFHS-4), 2015-16 [Dataset]. IAHR74DT.DTA. India. Mumbai: IIPS [Producers]. ICF [Distributor], 2017.

3   International Institute for Population Sciences (IIPS) and Macro International. National Family Health Survey (NFHS-3), 2005–06 [Dataset]. IAHR52DT.DTA. India. Mumbai: IIPS [Producers]. ICF [Distributor], 2007.

4   Coffey, D, A Gupta, P Hathi, et al. Revealed preference for open defecation. *Economic & Political Weekly* 2014, 49(38), 43-55.

5   Yogananth, N, & T Bhatnagar. Prevalence of open defecation among households with toilets and associated factors in rural south India: an analytical cross-sectional study. *Transactions of The Royal Society of Tropical Medicine and Hygiene* 2018, 112(7), 349-360.

6   Barnard, S, P Routray, F Majorin, et al. Impact of Indian Total Sanitation Campaign on latrine coverage and use: a cross-sectional study in Orissa three years following programme implementation. *PloS One* 2013, *8*(8), e71438.

7   Clasen, T, S Boisson, P Routray, et al. Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection, and child malnutrition in Odisha, India: a cluster-randomised trial. *The Lancet Global Health* 2014, 2(11), e645-e653.

8   Patil, SR, BF Arnold, AL Salvatore, et al. The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLoS Medicine* 2014, 11(8), e1001709.

9   Jenkins, MW, MC Freeman, & P Routray. Measuring the safety of excreta disposal behavior in India with the new Safe San Index: Reliability, validity and utility. *International Journal of Environmental Research and Public Health* 2014, 11(8), 8319-8346.

10   Sinha, A, CL Nagel, E Thomas, WP Schmidt, B Torondel, S Boisson, & TF Clasen. Assessing latrine use in rural India: a cross-sectional study comparing reported use and passive latrine use monitors. *The American Journal of Tropical Medicine and Hygiene* 2016, 95(3), 720-727.

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

12 Ministry of Drinking Water and Sanitation. Guidelines for Swachh Bharat Mission (Gramin). Government of India, New Delhi, India, 2017.

13 Coffey, D, A Gupta, P Hathi, et al. Understanding open defecation in rural India: Untouchability, pollution, and latrine pits. *Economic & Political Weekly* 2017, 52(1), 59-66.

14 O'Reilly, K, R Dhanju, & E Louis. Subjected to sanitation: Caste relations and sanitation adoption in rural Tamil Nadu. *The Journal of Development Studies* 2017, 53(11), 1915-1928.

15 Routray, P., WP Schmidt, S Boisson, T Clasen, & MW Jenkins. Socio-cultural and behavioural factors constraining latrine adoption in rural coastal Odisha: an exploratory qualitative study. *BMC Public Health* 2015, 15(1), 880.

16 Sahoo, KC, KR Hulland, BA Caruso, et al. Sanitation-related psychosocial stress: a grounded theory study of women across the life-course in Odisha, India. *Social Science & Medicine* 2015, 139, 80-89.

17 Vyas, S, & D Spears. Sanitation and Religion in South Asia: What Accounts for Differences across Countries? *The Journal of Development Studies* 2018, 54(11), 1-17.

18 Geruso, M, & D Spears. Neighborhood sanitation and infant mortality. *American Economic Journal: Applied Economics* 2018, 10(2), 125-62.

19 Coffey, D, & D Spears. *Where India goes: Abandoned toilets, stunted development and the costs of caste.* Harper Collins, India, 2017.

Figure 1: Sample selection

total villages in
sampling frame =
366

randomly selected
villages =
117

visited villages = 95

households randomly
assigned individual
questionnaire
n = 1614

households randomly
assigned household
questionnaire
n = 1600

completed individual
interviews
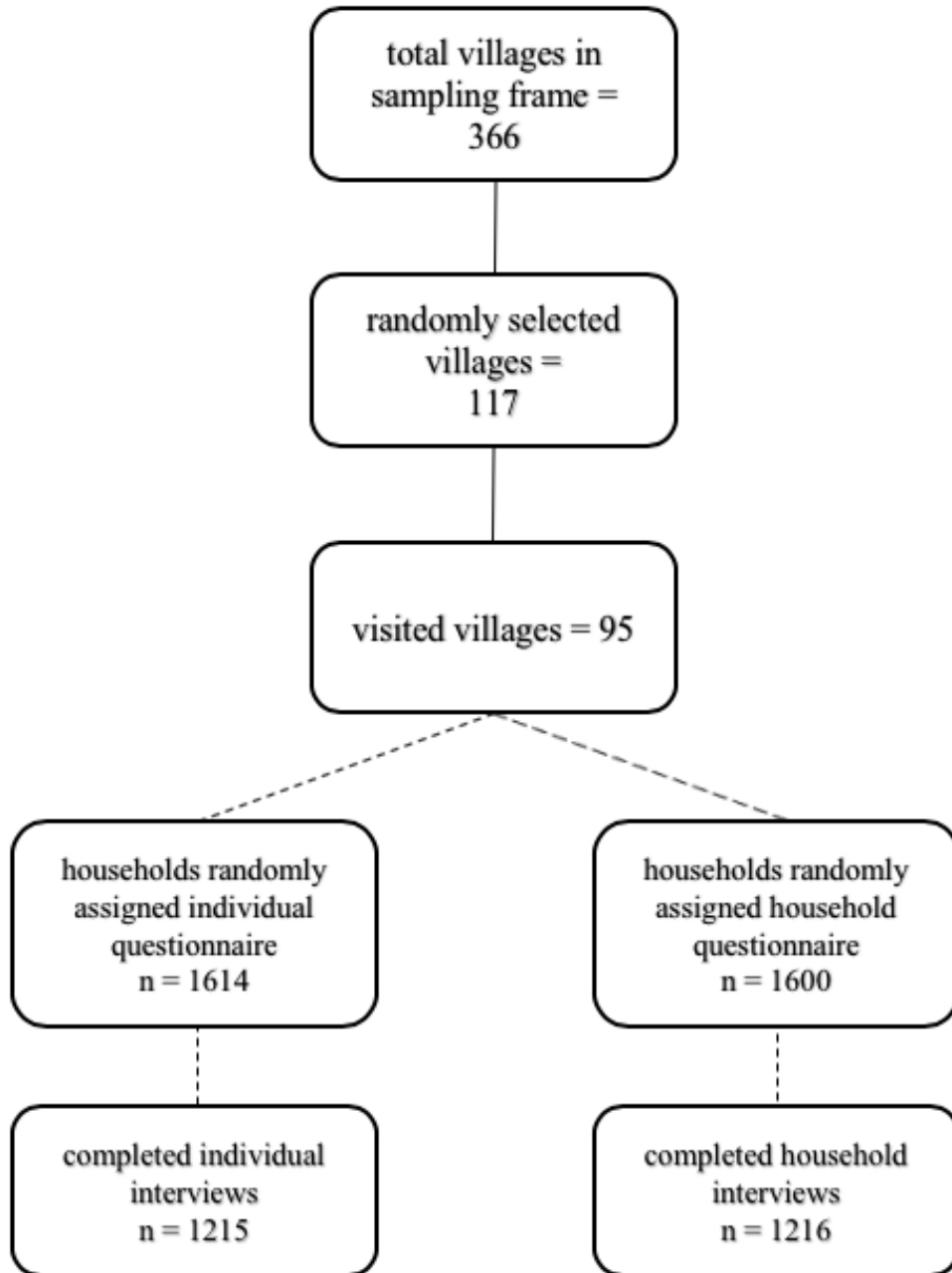n = 1215

completed household
interviews
n = 1216

Figure 2: Individual-level, balanced latrine use questions find significantly higher rates of open defecation than the household-level question
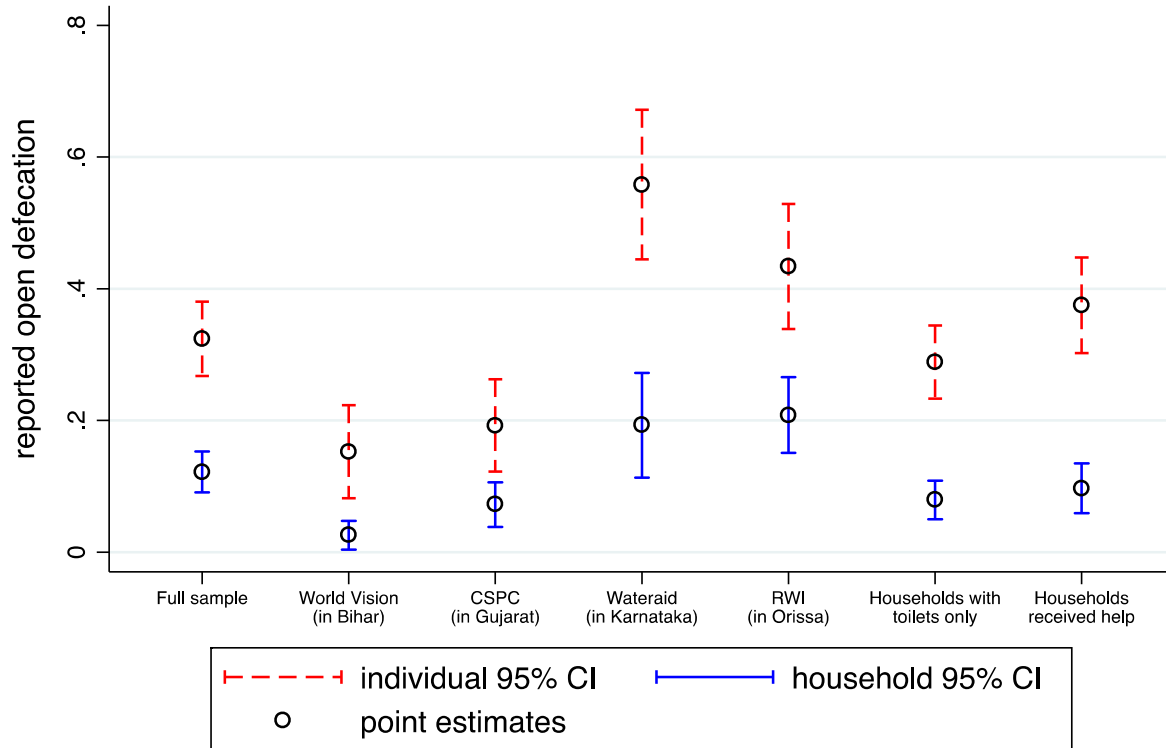
Figure shows means and confidence intervals. Red, dashed lines indicate responses to individual-level, balanced latrine use questions, and blue solid lines indicate responses to the household-level question. Unit of observation is individuals for individual-level questions and households for household-level questions. Confidence intervals are computed using standard errors clustered by village.

Table 1: Randomization balance: No significant differences in observed means between households assigned different latrine use questions

	individual	household	difference
	(1)	(2)	(3)
Number of households	1,215	1,216	-1
Number of households by project			
World Vision (in Bihar)	316	313	3
Coastal Salinity Prevention Cell (in Gujarat)	309	319	-10
Wateraid (in Karnataka)	297	296	1
Rural Welfare Institute (in Odisha)	293	288	5
Household members	5.685	5.604	0.0803
	(0.0950)	(0.101)	(-0.117 to 0.278)
Female	0.491	0.489	0.00190
	(0.00485)	(0.00490)	(-0.0115 to 0.0153)
Hindu	0.967	0.960	0.00737
	(0.00882)	(0.0102)	(-0.00383 to 0.0186)
Household head completed at least	0.288	0.319	-0.0310
eight years of schooling	(0.0192)	(0.0199)	(-0.0675 to 0.00552)
Count of assets (max 13)	8.202	8.234	-0.0327
	(0.111)	(0.117)	(-0.196 to 0.130)
Has latrine	0.943	0.946	-0.00251
	(0.00795)	(0.00781)	(-0.0215 to 0.0165)
Got help from government or NGO	0.625	0.656	-0.0309
to build toilet (given has toilet)	(0.0304)	(0.0274)	(-0.0692 to 0.00747)
Pit size (cubic feet, given has toilet)	179.0	180.7	1.666
	(14.33)	(14.27)	(-25.94 to 22.61)
Toilet looks used (given has toilet)	0.805	0.794	0.0103
	(0.0234)	(0.0233)	(-0.0243 to 0.0450)

Standard errors clustered by village under means in columns 1 and 2. 95% CI under differences in column 3.
**p<0.01, * p<0.05.

Table 2: Individual-level, balanced latrine use questions find significantly higher rates of open defecation than household level questions

unit of observation:	individuals		households			
question type:	individual	household (imputed)	household	individual (imputed)	difference (1)-(2)	difference (4)-(3)
	(1)	(2)	(3)	(4)	(5)	(6)
1. Full sample						
Estimates	0.324	0.115	0.122	0.326	0.209**	0.204**
	(0.0288)	(0.0150)	(0.0158)	(0.0289)	(0.163 to 0.254)	(0.160 to 0.248)
n (individuals or households)	13,070	13,070	2,431	2,431		
2. World Vision sample (in Bihar)						
Estimates	0.153	0.0184	0.0256	0.154	0.134**	0.128**
	(0.0361)	(0.00796)	(0.0112)	(0.0321)	(0.0659 to 0.202)	(0.0719 to 0.184)
n (individuals or households)	3,675	3,675	629	629		
3. Coastal Salinity Prevention Cell sample (in Gujarat)						
Estimates	0.192	0.0848	0.0721	0.169	0.108**	0.0973**
	(0.0358)	(0.0212)	(0.0174)	(0.0299)	(0.0355 to 0.180)	(0.0417 to 0.153)
n (individuals or households)	3,340	3,340	628	628		
4. Wateraid sample (in Karnataka)						
Estimates	0.558	0.184	0.193	0.550	0.374**	0.358**
	(0.0580)	(0.0365)	(0.0406)	(0.0575)	(0.262 to 0.486)	(0.247 to 0.468)
n (individuals or households)	3,112	3,112	593	593		
5. Rural Welfare Institute sample (in Odisha)						
Estimates	0.434	0.204	0.208	0.450	0.230**	0.242**
	(0.0484)	(0.0284)	(0.0294)	(0.0463)	(0.150 to 0.309)	(0.170 to 0.313)
n (individuals or households)	2,943	2,943	581	581		
6. Households with toilets only						
Estimates	0.289	0.0734	0.0791	0.291	0.215**	0.211**
	(0.0284)	(0.0137)	(0.0150)	(0.0287)	(0.166 to 0.265)	(0.162 to 0.261)
n (individuals or households)	12,366	12,366	2,296	2,296		
7. Households that received help						
Estimates	0.375	0.0895	0.0968	0.378	0.285**	0.281**
	(0.0370)	(0.0181)	(0.0193)	(0.0368)	(0.220 to 0.351)	(0.217 to 0.346)
n (individuals or households)	7,958	7,958	1,470	1,470		

Standard errors clustered by village under means in columns 1 through 4. 95% CI under differences in columns 5 and 6. **p<0.01, * p<0.05. Household imputed refers to estimates of individual open defecation imputed from answers to the household-level question. Each individual in the household is given the same answer as the household-level answer. Individual imputed refers to estimates of household open defecation imputed from answers to the individual level questions. The household estimates are constructed by averaging open defecation among individuals in the household.

Table 3: Sub-group analyses

unit of observation:	individuals	households	individuals
	(1)	(2)	(3)
Individual-level question	0.100**	0.0992**	0.229***
	(0.0548 to 0.146)	(0.0566 to 0.142)	(0.182 to 0.275)
Received help for construction	0.0470*	0.0514**	
	(0.00962 to 0.0844)	(0.0134 to 0.0893)	
Individual-level question X Received help	0.185**	0.182**	
	(0.113 to 0.257)	(0.114 to 0.251)	
Female			0.00134
			(-0.0118 to 0.0144)
Individual-level X Female			-0.0409***
			(-0.0646 to -0.0172)
Constant	0.0425**	0.0455**	0.115***
	(0.0180 to 0.0670)	(0.0201 to 0.0708)	(0.0846 to 0.145)
n (individuals or households)	12,366	2,296	13,070

95% CI under coefficients, calculated using standard errors clustered by village. **p<0.01, * p<0.05.