Increasing average exposure to open defecation in India, 2001-2011

Dean Spears[1,*]

1 Spears: Centre for Development Economics, Delhi School of Economics, Delhi University, Delhi, India

∗ **E-mail: dean@riceinstitute.org**

Abstract

Although most people in India defecate in the open without using a latrine, and most people in the world who defecate in the open live in India, the fraction of households in India who defecate in the open decreased between the 2001 and 2011 rounds of the Indian census. However, an active and growing recent literature indicates that exposure to a high *density* of open defecation – measured as people defecating in the open per square kilometer – may be an especially important threat to early-life health and human capital accumulation. We use 2001 and 2011 Indian census data to show that exposure to open defecation has not decreased by this measure, on average, and according to some computations has increased. Most people in India live in a district where the density of open defecation increased from 2001 to 2011. For the 98.6% of the population which lives outside of the union territory of Delhi, average exposure to open defecation increased over this decade. Almost half of people in India live in a state where statewide open defecation density increased; these states are concentrated along the largely rural, impoverished "Hindi Heartland" belt of north India. A Kitagawa decomposition shows that an important contributing factor was the net increasing concentration of the Indian population into places with high open defecation density. Our results have implications for where the attention of sanitation policy-makers may be most needed: about 38% of all people worldwide who defecate in the open live in an Indian district where the density of open defecation increased from 2001 to 2011.

Introduction

Open defecation, without using a toilet or latrine, is an important challenge for early-life health and human capital accumulation in developing countries. This is especially true in India, where most people defecate in the open, where most of the people worldwide who defecate in the open live [1], and where population density is high. The Indian government's 2001 and 2011 census data report that the fraction of households defecating in the open in India declined slowly over this decade, from about 63 percent to about 53 percent. The Unicef-WHO Joint Monitoring Programme estimates that the number of people defecating in the open in India peaked at 664 million in 1996 and has since declined to 605 million in 2001. If the count of persons in India defecating in the open continues to decline at this linearized rate, India would eventually eliminate open defecation in about 150 years.[1] In the meanwhile, many more generations will suffer the lasting effects of early-life exposure to fecal germs.

This paper asks whether even this description of slow average decline may be too optimistic. An active and growing recent literature suggests that the sanitation exposure that matters most – or, at least one key exposure which matters greatly – is density of open defecation in a child's local area: the number of people defecating in the open per square kilometer [2–6]. Moreover, an existing literature anticipates an interaction between population density and sanitation: sanitation should matter more where people live more densely together [7–9]. This would be consistent with open defecation causing illness by releasing fecal germs into the environment, where people are more likely to encounter them if they live closer together.

This paper assumes that open defecation density is an important measure of exposure to poor sanitation. We use the two most recent rounds of the Indian census to ask whether average exposure improved or worsened between 2001 and 2011. Even if the number of people defecating in the open decreases, average exposure to open defecation density can increase if the population is increasingly arranged such that people live near to people who defecate in the open. If population density increases more quickly than the fraction of the people defecating in the open declines, and especially if the population becomes increasingly concentrated in places with poor sanitation, then average exposure to open defecation would increase.

[1] 664 - 605 = 59 million over 15 years or 3.93 million reduced per year; 605 million remaining over 3.93 million per year equals 154 years. Of course, there is no reason to expect the decline to be linear generically, or linear in number of persons; this computation merely illustrates that the decline is slow.

Materials and Methods

Data: Indian Census of 2001 and 2011

This paper reports analysis of published aggregate census data from the 2001 and 2011 Census of India, conducted by the Government of India. These data are publicly available free of charge at censusindia.gov.in. Data are published at two levels of disaggregation: states and districts, where districts are administrative partitions of states. We replicate our analysis at the state and district levels of disaggregation. During the time period we study, there were 34 states in India; there were 594 districts in the 2001 census and 641 districts in the 2011 census.

In particular, for each census year and for each district and state, we use the following aggregate variables:

- total population (number of persons),
- population density (persons per square kilometer),
- count of households classified as defecating in the open (no toilet or latrine),
- total count of households.

From these data, we compute the fraction of households defecating in the open, f, as the ratio of the count of households defecating in the open to the total count of households. There are two limitations of this measure of open defecation. First, the census data is primarily a count of toilet or latrine ownership, not use; unfortunately, there are many latrines in India that are not used at all [10, 11]. Second, latrine ownership is recorded at the household level, where latrine use or open defecation is in fact an individual level behavior: many Indian households own latrines that some household members use and other members do not.

Statistical Analysis

The central variable that our analysis constructs is open defecation density, x (for "exposure"), the number of people who defecate in the open per square kilometer in a given area:

$$\begin{aligned} \textit{open defecation density} &= \frac{\textit{number of open defecators}}{\textit{area}} \\ &= \frac{\textit{number of open defecators}}{\textit{number of people}} \times \frac{\textit{number of people}}{\textit{area}} \\ &= fd = x, \end{aligned} \tag{1}$$

where x is open defecation density, f is the fraction of the population that defecates in the open, and d is the population density in people per square kilometer. Note that we are constrained by available data to substitute the fraction of *households* who report defecating in the open for f in our analysis.[2]

We construct our measure of average exposure to open defecation density, $\bar{x}$, as a weighted average of x_i for each place i, here a state or a district. Thus we compute

$$\bar{x} = \sum_i w_i x_i, \tag{2}$$

where w_i is the share of the Indian population in place i:

$$w_i = \frac{p_i}{\sum_j p_j}, \tag{3}$$

where p_i is the population of place i.

Finally, in order to decompose the change in average exposure to open defecation density between the two census rounds into intra-state changes in open defecation density x_i, and changes in the distribution of the population across states w_i, we perform a Kitagawa [12] decomposition of the total change. Therefore, for this decomposition case, i indexes states. A Kitagawa decomposition separates the change in a weighted average over categories into the change due to change in the weighting across categories and

[2] This is equivalent to assuming that households in India that do and do not defecate in the open are the same size, on average. According to our computations from India's 2005 National Family Health Survey (NFHS-3, India's implementation of the Demographic and Health Surveys), households that defecate in the open have 3.5% more members than households that do not, on average. This small average gap makes quantitatively very little difference if inserted into our main computations.

the change due to the average change within categories [13]:

$$
\begin{aligned}
\bar{x}_{2011} - \bar{x}_{2001} &= \sum_i w_{i,2011} x_{i,2011} - \sum_i w_{i,2001} x_{i,2001} \\
&= \sum_i (w_{i,2011} - w_{i,2001}) \left(\tfrac{x_{i,2011} + x_{i,2001}}{2} \right) + \sum_i (x_{i,2011} - x_{i,2001}) \left(\tfrac{w_{i,2011} + w_{i,2001}}{2} \right) \\
&= (\Delta \textit{ due to } w_i \textit{ redistribution among states}) - (\Delta \textit{ due to changes in } x_i \textit{ within states})
\end{aligned} \tag{4}
$$

This technique is useful here because it allows us to assess the extent to which the change in the average Indian person's exposure to state-level open defecation density is due to net changes in where people live (and what they are therefore exposed to), rather than changes over time in place specific open defecation densities.

All statistical analysis was performed using Stata 12.1 software. Ethical approval was not sought for this secondary analysis of publicly available, published aggregate data.

Results

Average exposure to open defecation density

Did the average exposure to open defecation increase or decrease between the 2001 and 2011 census rounds? Table 1 presents estimates of average exposure to open defecation density, $\bar{x}$, for 2001 and 2011. Panels A and B disaggregate India into states – assigning each person the open defecation density of her state – while Panels C and D disaggregate further into districts. As is well known, the percent of households defecating in the open decreased between the two census rounds. Only 1% of people live in a state where the fraction of households without a latrine increased, and only 5.6% live in a district where the fraction of households without a latrine increased.

As discussed above, however, evidence suggests that what matters for health and human capital is exposure to open defecation density, rather than latrine ownership *per se*. Approximately as many people live in Indian places where open defecation density increased from 2001 to 2011 as live in districts where it declined. States in India can be very large, even relative to other countries; district-level disaggregation is therefore preferable if open defecation is not uniformly distributed within states, which it is not. However, although district-level results more closely match the local experience of the average person in India, we include state-level results because of the importance of states in Indian policy-making. As Panel C shows, most Indians – 52.9% – live in districts where the density of open defecation *increased* between 2001 and 2011. This is a main result of this paper.

The percent of people exposed to state-level increases in open defecation density is smaller than the percent of people exposed to district-level increases in open defecation density because states are much larger than districts, which obscures heterogeneity in people's local environments. There were 59 districts located in states where state-level open defecation density decreased, but where district-level average open defecation density increased. The largest number of these are in Maharashtra: 10 districts out of 35 in Maharashtra. This is because the state-wide average is pulled down by a large decrease in open defecation density in the large, densely populated city of Mumbai. This change in Mumbai clearly has little effect on the disease exposure of people living in a distant district of Marahrashtra, however.

Districts which did not exist in 2001 are excluded from panels C and D, and therefore from our main result in panel C. These represent a small number of districts, resulting from mergers and bifurcations. How might this exclusion change our conclusion that most Indians live in a district where density of open defecation increased from 2001 to 2011? In the most extreme scenario, assume that *none* of these districts experienced an increase in open defecation density; even in this case, 50.9% of all Indians would live in a district where open defecation density increased from 2001 to 2011. Therefore, our main result is robust to any possible configuration of these missing districts. Alternatively, we can predict missing values by fitting a linear probability regression model to predict a binary indicator for increasing open defecation density among the districts with data.[3] Substituting these predictions for the missing values, we compute an average of 52.8%, essentially replicating the result in Panel C.

Population density is very asymmetrically distributed, and weighted mean exposure to open defecation $\bar{x}$ can therefore be very sensitive to outliers. Panels B and D compute $\bar{x}$ omitting a small part of the sample where the reduction in open defecation density was exceptionally high, as the next subsection will show. In particular, the state-level results in Panel B omit the 1.4% of the Indian population who live in Delhi, which enjoyed a decline in open defecation density of 863 open defecators per square kilometer. Outside of Delhi, among the remaining 98.6% of the population, average exposure to open defecation increased from 299.6 to 306.9 open defecators per square kilometer. In the district-level analysis, a similar result is found: omitting the extreme outliers of Delhi, Mumbai (Mumbai and Mumbai Suburban

[3]We regress an indicator for increasing open defecation density on state fixed effects, 2011 population, 2011 fraction defecating in the open, 2011 population density, and 2011 open defection density. Each of these variables is statistically significant at at least the two-sided 5% level. We use regression estimates to predict missing values, constraining them to be on the interval $[0, 1]$.

districts), and Chennai, in the remainder of the Indian population, mean exposure to open defecation increased. Many, many people live in Indian districts where the decline in the open defecation percent, f, has not kept pace with the increase in population density, d.

Distribution of exposure to open defecation density

Having considered the average change in exposure to open defecation, we can turn to the distribution. Table 2 uses the state-level data to report the average levels and change in exposure to open defecation density within each state. For many Indians, open defecation density considerably increased: in Bihar, where 8.6 percent of the Indian population lived in 2011, 140 more people per square kilometer defecated in the open in 2011 than in 2001. In Uttar Pradesh – where 16.5 of Indians live – the number of open defecators per square kilometer increased by more than 60 over this decade. Overall, 48 percent of people living in India in 2011 lived in states where the statewide average density of open defecation increased from 2001 to 2011.

Figure 1 plots these state-level changes on a map. Each state is shaded according to the percent of its 2001 open defecation density that its 2001 to 2011 change in open defecation density represents. As the map makes clear, states where the density of open defecation has increased form a contiguous belt across the impoverished, largely rural, central belt of north India. These states are sometimes referred to as the "Hindi Heartland" states, to reflect their shared linguistic, social, and cultural characteristics.

Empirical cumulative density functions permit a compact visualization of the change in exposure to poor sanitation between the two census rounds. Figure 2 again replicates the well-known fact that the open defecation percent declined between the two census rounds. At either the state or district level of disaggregation, the 2001 distribution stochastically dominates the 2011 distribution, meaning that a higher fraction of households defecated in the open in 2001 throughout the distribution.

Figure 3 instead plots the distribution of open defecation density, weighted by the number of people in each state or district, and therefore exposed to that state's or district's sanitation externalities. At the state level, in Panel A, the two distributions are very similar, and they cross several times: there is no state-level evidence that state-level exposure to open defecation density has decreased between the two censuses.

Panel B plots these distributions at the district level. It is difficult to distinguish the two curves because in panel B the figure is dominated by the very small fraction of Indian people who live in districts with a very large decrease in open defecation density. These are highly urbanized districts with high population densities. Following Table 1, panel C presents a trimmed sample, omitting Delhi, Mumbai, and Chennai. Here, again, the 2001 and 2011 distributions of exposure to open defecation are similar.

Finally, Figure 4 plots the distributions of the state-level and district-level intra-place change in density of open defecation. As before, in the full sample of Panel A, these distributions are dominated by the few highly dense places with very large absolute declines. However, the trimmed sample in Panel B shows that, outside of these few cities, about as many people live in places where open defecation density declined as in places where open defecation density increased.

Decomposition of state-level change

There are two possible components of the total change in average exposure to open defecation density, $\bar{x}_{2011} - \bar{x}_{2001}$. First, open defecation density may have changed within places. For example, open defecation density increased in Bihar state but decreased in Mumbai district. Second, either because of net migration or differential fertility and mortality, the distribution of the population among states, $\{w_i\}$, may have changed.

To illustrate the second possibility, consider a country with two districts:

- District A: 10 people who defecate in the open, 20 people who do not defecate in the open, 1 km^2 of land.
- District B: 5 people who defecate in the open, 5 people who do not defecate in the open, 1 km^2 of land.

As a whole, the country has 15 people who defecate in the open living on 2 km^2, for an overall open defecation density, x_{both} of 7.5. However, disaggregating at the district level, people in the country are exposed to an average open defecation density of

$$\bar{x} = \frac{10 \times (10 + 20) + 5 \times (5 + 5)}{40} = 8.75.$$

Now imagine an extra person is born in District A who does not defecate in the open. This birth will:

- Decrease f_{both} and f_A, the fraction of people defecating in the open, because the number of people defecating in the open has stayed the same while the population size has increased.

- Have no change on x_{both}, the country-wide aggregate open defecation density of 7.5, because neither the total count of people defecating in the open nor the area of the country changed.

- Have no change on x_A or x_B, the within-district open defecation densities, for the same reason, and

- Increase w_A and decrease w_B, because a larger fraction of the country's population now lives in District A.

Therefore, because $x_A > x_B$, $\bar{x}$ will increase, even though the fraction of people defecating in the open has decreased and no district's open defecation density has increased. The exposure of the *average person* has increased because the distribution of the population into places is now such that more people are living in places where people were already exposed to high density of open defecation.

Table 3 decomposes the state level change in average exposure to open defecation into changes in x_i and changes in w_i. The results indicate that changes in w_i are an important component of the average change. Panel A reports that, holding constat either the 2001 or 2011 state-level open defecation densities, moving from the 2001 distribution of the population into states to the 2011 distribution of the population into states increases the average exposure to open defecation density. The Kitagawa decomposition in Panel B finds that changes in the distribution of people into states is responsible for the change in exposure to open defecation density being 4 people per square kilometer greater than it otherwise would be, if the same fraction if people had instead lived in each state in 2011 as did in 2001. Thus, a larger fraction of the Indian population lived in places with higher open defecation densities in 2011 than in 2001.

Discussion

According to the Government of India's census data, most people in India live in a district where their exposure to density of open defecation increased from the 2001 census to the 2011 census. This is especially true outside of a few highly densely populated and relatively highly developed metropolises. If exposure to open defecation density is indeed a relevant risk factor for early life health and human capital accumulation – as an active and growing literature indicates – then there is little evidence that sanitation in India has been improving; indeed the average sanitation exposure in India arguably worsened from

2001 to 2011.[4]

Because India is so large, and because India is such a large fraction of the world's open defecation problem, what is true of India is informative about human exposure to sanitation globally. The Unicef-WHO Joint Monitoring Programme estimates that approximately 60% of all people who defecate in the open live in India [1]. Our computations indicate that 64% of people in India who defecate in the open live in a district where open defecation density increased over this period. Taken together, these figures imply that 38% of all people who defecate in the open worldwide live in an *Indian district* where open defecation density increased from 2001 to 2011. Of course, the motivating idea behind this paper is that open defecation has bad effects on *everyone* who lives nearby, so sanitation policy has an important role to play even in the lives of people who already use save sanitation. Our results suggest that the world's open defecation problem – and its bad consequences – may be increasingly concentrated in densely populated north India.

Acknowledgments

The author is grateful to Balaganesh M. for excellent research assistance assembling the district-level dataset and to Diane Coffey and Payal Hathi for comments and suggestions.

References

1. Joint Monitoring Programme for Water Supply and Sanitation (2012) Progress on Drinking Water and Sanitation: 2012 Update. WHO and UNICEF.

2. Spears D (2013) How much international variation in child height can sanitation explain? Policy Research Working Paper 6351, World Bank.

3. Spears D (2012) Effects of rural sanitation on infant mortality and human capital: Evidence from a local governance incentive in India. Research program in development studies, Princeton.

4. Hathi P, Haque S, Pant L, Coffey D, Spears D (2014) Do toilets spill over?: Population density and the effect of sanitation on early-life health. working paper, r.i.c.e.

[4] This would be true on average, for India as a whole; sanitation did improve in *some* districts by more than in others, and those districts saw relatively greater improvements in infant mortality and child height [3].

5. Coffey D (2014) Sanitation externalities, disease, and children's anemia. working paper presented at annual meetings of population association of america, Princeton University.

6. Vyas S, Gupta A, von Medeazza G, Spears D (2014) Open defecation, population density, and child height in Madhya Pradesh, India: An ecological analysis of new data on 22,000 children. working paper, r.i.c.e.

7. Black M, Fawcett B (2008) The Last Taboo: Opening the Door on the Global Sanitation Crisis. London: Earthscan.

8. Bateman OM, Smith S, Roark P (1993) A comparison of the health effects of water supply and sanitaiton in urban and rural areas of five African countries. WASH Field Report 398, USAID.

9. Bateman OM, Smith S (1991) A comparison of the health effects of water supply and sanitaiton in urban and rural Guatemala. WASH Field Report 352, USAID.

10. Bongartz P, Chambers R (2009) Beyond subsidies: Triggering a revolution in rural sanitation. IDS In Focus 10.

11. Lamba S, Spears D (2013) Caste, 'cleanliness,' and cash: Effects of caste-based political reservations in Rajasthan on a sanitation prize. Journal of Development Studies 49: 1592-1606.

12. Kitagawa EM (1955) Components of a difference between two rates. Journal of the American Statistical Association 50: 1168-94.

13. Schempf A, Becker S (2006) On the application of decomposition methods. American Journal of Public Health 96: 1899.

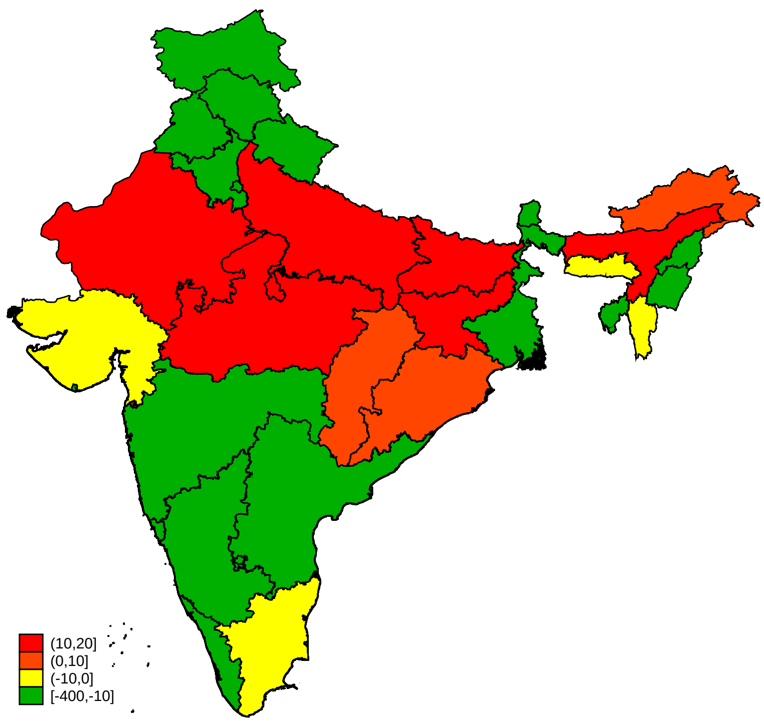

Figure 1. Percent change in open defecation density by state. Each state is shaded according to the percent increase or decrease of its average open defecation density from the 2001 to the 2011 census.

Figure Legends

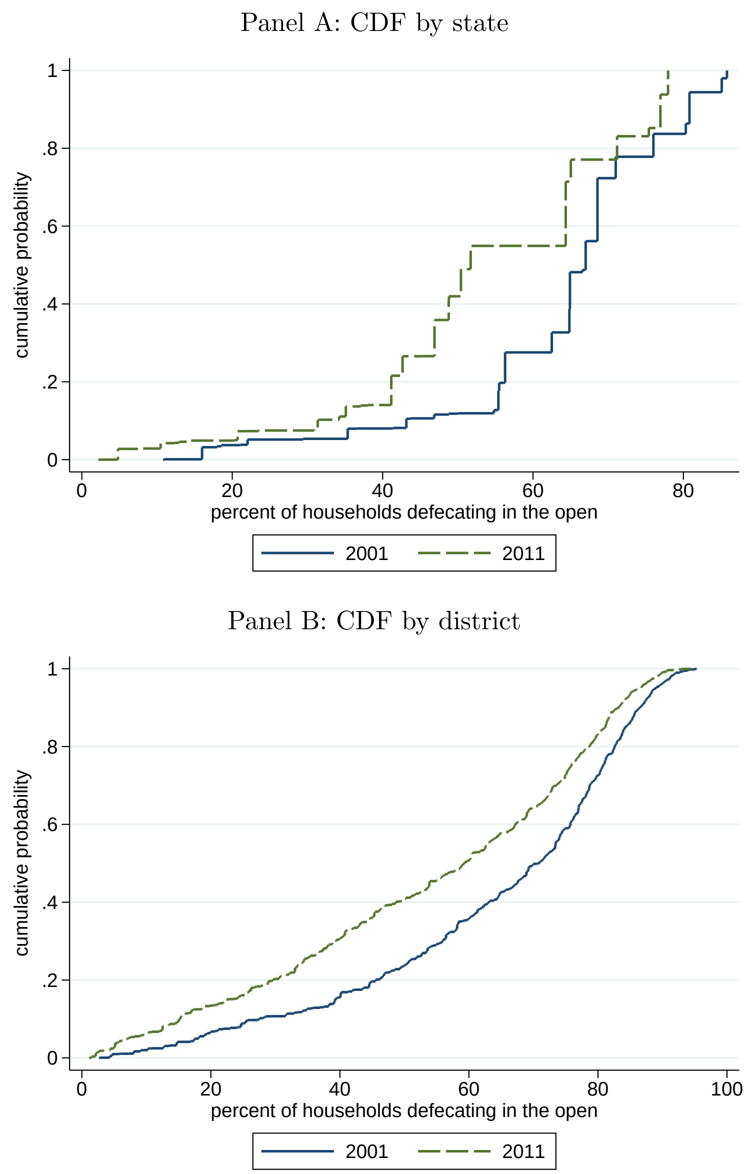

Figure 2. The percent of households defecating in the open declined from 2001 to 2011. Observations are weighted by population.

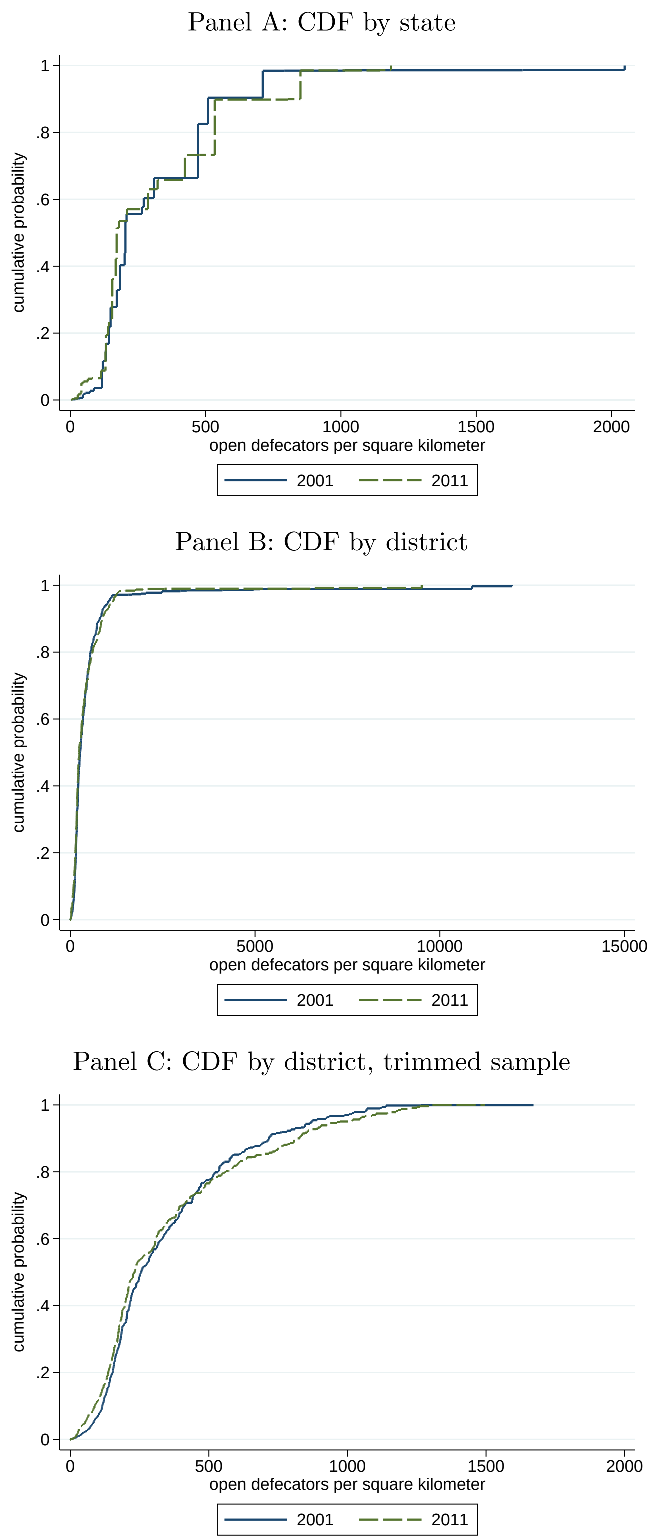

Figure 3. The distributions of average open defecator density are similar for 2001 and 2011. Observations are weighted by population. “Trimmed” sample omits Delhi, Mumbai, and Chennai.

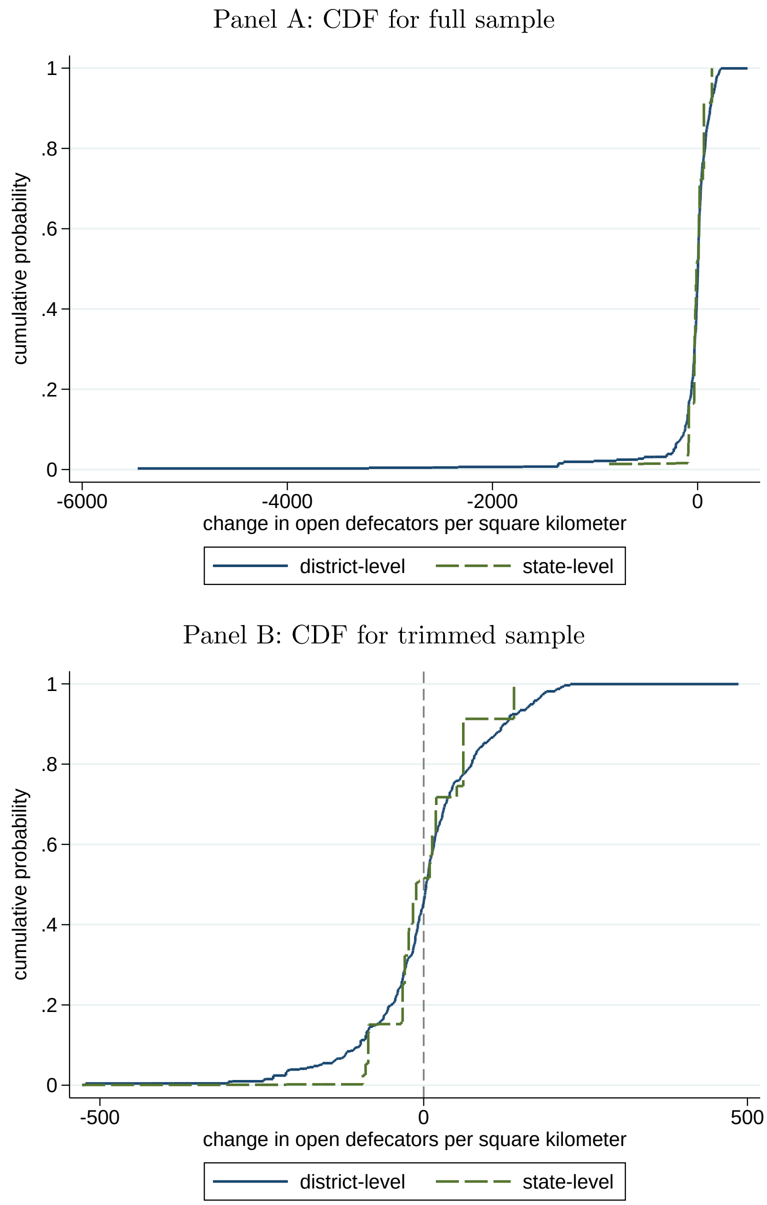

Figure 4. The intercensal change in open defecation density was positive for about as many people as it was negative. Observations are weighted by population. "Trimmed" sample omits Delhi in the state-level distribution and Delhi, Mumbai, and Chennai in the district-level distribution. Only districts which existed in both census years are included in the district-level distribution.

Table 1. Average exposure to open defecation, 2001 and 2011

year:	2001	2011	year:	2001	2011
Panel A: State-level disaggregation			Panel C: District-level disaggregation		
sample:	all India		sample:	all India	
open defecation rate	63.9%	53.9%	open defecation rate	63.8%	54.1%
open defecators/km^2	323.1	319.1	open defecators/km^2	513.2	455.6
rate increased		1.0%	rate increased		5.6%
exposure increased		47.8%	exposure increased		52.9%
median exposure change		-11.6	median exposure change		2.7
n (people):	1.027b	1.210b	n (people):	1.027b	1.210b
Panel B: State-level disaggregation			Panel D: District-level disaggregation		
sample:	Delhi omitted		sample:	Delhi, Mumbai, & Chennai omitted	
open defecation rate	64.4%	54.5%	open defecation rate	64.7%	55.0%
open defecators/km^2	299.6	306.9	open defecators/km^2	346.0	350.9
rate increased		1.1%	rate increased		5.8%
exposure increased		48.5%	exposure increased		54.5%
median exposure change		-11.6	median exposure change		3.6
n (people):	1.013b	1.193b	n (people):	0.998b	1.177b

Source: Author's computations from 2001 and 2011 Census of India. The open defecation rate is the percent of census households who defecate in the open. In "rate increased," "exposure increased," and "median exposure change" computations only states (in Panels A and B) and districts (in Panels B and C) which existed in both census years. "Rate increased" is the percent of the population living in a place (state or district) where the percent of households defecating in the open increased; "exposure increased" is the percent of the population living in a place where open defecation density increased; "median exposure change" is the median person's change in average open defecation density at the state or district level, respectively Observations are persons in all computations, so averages are weighted by the number of people in each state or district in each year.

Tables

Table 2. Change in average exposure to open defecation, by state 2001-2011

state	2001 OD density	2011 OD density	change	2011 population (%)
Bihar	711.1	850.6	139.5	8.6
Uttar Pradesh	472.5	533.7	61.2	16.5
Jharkhand	271.5	322.7	51.2	2.7
Assam	120.2	139.7	19.5	2.6
Madhya Pradesh	149.0	167.7	18.7	6.0
Dadra & Nagar Haveli	302.8	316.8	14.0	0.0
Rajasthan	117.2	130.3	13.1	5.7
Chhattisgarh	132.1	142.5	10.4	2.1
Odisha	200.8	210.2	9.3	3.5
Arunachal Pradesh	5.7	6.3	0.6	0.1
Mizoram	4.6	4.2	-0.4	0.1
Meghalaya	50.3	49.1	-1.2	0.2
Manipur	19.2	14.9	-4.3	0.2
Nagaland	35.3	28.0	-7.3	0.2
Tripura	56.4	48.9	-7.5	0.3
Gujarat	142.9	131.4	-11.6	5.0
Karnataka	171.9	155.4	-16.5	5.0
Sikkim	27.8	11.0	-16.8	0.1
Jammu & Kashmir	46.4	27.6	-18.8	1.0
Uttarakhand	87.1	64.5	-22.6	0.8
Tamil Nadu	310.0	286.8	-23.1	6.0
Andhra Pradesh	184.3	155.0	-29.2	7.0
Maharashtra	203.8	171.3	-32.6	9.3
Himachal Pradesh	72.6	38.1	-34.5	0.6
Goa	150.2	79.9	-70.2	0.1
Haryana	264.7	179.9	-84.8	2.1
West Bengal	508.8	423.2	-85.6	7.5
Kerala	130.9	41.3	-89.7	2.8
Punjab	208.0	114.1	-93.9	2.3
Lakshwadeep	204.6	47.4	-157.2	0.0
Puducherry	1015.6	803.5	-212.1	0.1
Daman & Diu	791.0	477.2	-313.8	0.0
Chandigarh	1671.3	1144.0	-527.3	0.1
Delhi	2048.4	1185.4	-863.0	1.4

Source: Author's computations from 2001 and 2011 Census of India. States are listed in decreasing order of "change," which is the absolute intercensal increase in open defecators per square kilometer. 2011 population is the percent of India's 2011 population living in that state.

Table 3. Decomposition of state-level change in average exposure to open defecation density

Panel A: Interchanging 2001 and 2011 exposure and population distribution			
	2001 exposure	2011 exposure	difference:
2001 distribution	323.1	314.0	-9.1
2011 distribution	326.8	319.1	-7.7
difference:	3.7	5.1	-4.0

Panel B: Kitagawa decomposition			
source of change:	exposure change	distributional change	total change
contribution:	-8.4	4.4	-4.0

Source: Author's computations from 2001 and 2011 Census of India. Panel A computes the average exposure to open defecation density using the 2001 and 2011 state open defecation densities and fractions of the Indian population. Panel B reports a Kitagawa decomposition of the India-wide average change in average exposure to open defecation density into the part due to change in within-state average open defecation density ("exposure change") and change in the distribution of the population among Indian states ("distributional change").