Title

Transparency ‘Fixes’ for Local Public Services: Field Experimental Evidence from Bangalore's Water Sector

Authors

Kumar, Tanu
Post, Alison E
Ray, Isha

Publication Date

2017-01-19

Flows, Leaks and Blockages in Informational Interventions: A Field Experimental Study of Bangalore's Water Sector

Tanu Kumar[*1]
tkumar@berkeley.edu

Alison E. Post[1]
aepost@berkeley.edu
Charles and Louise Travers Department of Political Science, University of California, Berkeley, Berkeley, California 94720.

Isha Ray[2]
isharay@berkeley.edu

May 15, 2017

[*] Corresponding author

[1] Charles and Louise Travers Department of Political Science, University of California, Berkeley, Berkeley, California 94720.

[2] Energy and Resources Group, University of California, Berkeley, Berkeley, California 94720

ABSTRACT

Under what circumstances might providing citizens with information compensate for unreliable public services? We present a field-experimental evaluation of a program that provided households in Bangalore with advance notification of intermittently provided piped water. The implementers expected that increasing service predictability would reduce wait times for water, reduce costs related to waiting, and improve citizen-state relationships. As many citizens did not receive accurate information, our study detected no impacts on household wait times for water or state-citizen relations. Nonetheless, our study suggests that notifications about water timing reduced stress, especially among low income populations. These findings indicate that greater attention should be paid to both psychological outcomes and the information production and dissemination chain in information interventions. We introduce a causal framework for analyzing "information pipelines" to enable such efforts.

KEYWORDS:

intermittent water supply, transparency, frontline worker, stress, India

HIGHLIGHTS

- Evaluation of text-message program providing notifications prior to water arrival in the intermittent water supply system of Bangalore, India
- We detected no impact on household welfare or state-citizen relations
- We detected modest reductions in household stress related to water
- Many households did not register receipt of accurate information, motivating analysis of the information production and dissemination process
- We introduce a framework for finding leaks and blocks between information collection and knowledge receipt in informational interventions

ACKNOWLEDGEMENTS

This project was funded through a "DIL Innovate" grant through the Development Impact Laboratory, Blum Center for Developing Economies, U.C. Berkeley (USAID Cooperative Agreement AID-OAA-A-13- 00002). We thank C. Sharada Prasad, Zachary Burt, Francesc Pardo, Mary Rose Fissinger, and Megan Otsuka for their considerable contributions to this project; Maria Chang, Eugenia Giraudy, S. Krishna Kumar, Rajani Sampath, and Christine Thao Tyler for research assistance; and Ben Arnold, Adam Auerbach, Jennifer Bussell, Pradeep Chhibber, Thad Dunning, Christopher Hyun, Francesca Jensenius, Katerina Linos, Aila Matanock, Edward Miguel, Gareth Nellis, Malini Ranganathan, Jasjeet Sehkon, Laura Stoker, Rebecca Weitz-Shapiro, Erik Wibbels, Matthew Winters, seminar participants at U.C. Berkeley, Duke and Claremont McKenna, and especially Joel Middleton and Guadalupe Tunon for advice. We also thank NextDrop, the Public Affairs Foundation (esp. Lalita Pulavarti and Narendra LV), and the Bangalore Water Supply and Sewerage Board (BWSSB) for their collaboration with the project. The project has been approved through IRB protocol 2014-04-6259 and approved by BWSSB. Pre-Analysis Plan 20150514AA registered with egap.org.

INTRODUCTION

Following the 2004 *World Development Report's* influential call, economists and political scientists have analyzed several programs that provide citizens, especially the poor, with more information about public services (World Bank 2004). Citizens armed with information about corruption or service deficiencies, it is argued, can vote underperforming politicians out of office and hold public sector bureaucracies more firmly to account. Providing information thereby sets in motion a virtuous cycle leading to improvements in service quality and access. A growing body of scholarship has evaluated the extent to which, and circumstances under which, these propositions hold (for reviews, see Pande 2011; Lieberman, Posner, and Tsai 2014). Much less attention has been paid to the production and dissemination of the information itself, a process that is fundamental to the success or failure of such interventions. In this paper, we assess both the impact of an informational intervention intended to make intermittently-supplied urban water services more predictable, and the process of producing the information central to the intervention.

Throughout Asia and Africa, intermittency is the hallmark of public service delivery: buses do not run on a standard schedule, water supplies stop and start again, and electricity blackouts occur regularly. Moreover, the poor state of the underlying urban infrastructure—prone to pipe leaks and power outages—means that services are not only intermittent, but are also often unpredictable. Over 300 million people receive piped water intermittently, mostly in South Asia and Sub-Saharan Africa (Kumpel and Nelson 2016, 543). Yet few studies analyze intermittency and unpredictability in service sectors such as domestic water, and none have considered how better information can help reduce the coping costs associated with them.

In this paper, we assess the impact of an informational intervention intended to make intermittently provided water services more predictable; we also assess the process of producing the information central to the intervention. Through a cluster-randomized experiment in Bangalore, we evaluate the household-level impacts of a service developed by NextDrop, a social enterprise, which sent households text-message notifications on water arrival times and supply cancellations. These notifications went to both households who had their own taps and those who collected water from communal standpipes.[3]. Our study is the first experimental evaluation of a program designed to make intermittent urban water supplies more predictable.

Water utilities in the developing world often do not possess sensors that allow them to monitor where water is flowing within their network. NextDrop developed a novel system in which the valvemen, or utility employees who operate the valves that channel water into areas of 50-200 households at a time, notified them when they were opening and closing valves. NextDrop then sent notifications to individual households, letting them know when their water would be turned on.

Assuming that these notifications successfully reached households, there were strong reasons to expect that increasing service predictability would improve household welfare, particularly for low-income households. When water comes only once or twice a week, for a few hours at a time, the individuals responsible (typically women) for managing household supply must be ready to fill storage containers the moment it arrives, even if this means waiting long periods because timing is unpredictable. This waiting can prevent them from spending time outside the home at work or in the community. It can also be stressful, given the potential financial and domestic consequences of missing a day's supply. Substitute sources, after all, are many times more expensive, and often of poorer quality.

Service unpredictability may also weaken bonds between individuals and the state. Citizens who cannot depend upon regular services may be less likely to view government service providers as competent. If citizens see (or think) that those with stronger political connections find it easier to obtain information about public services, they will not see the state as universalistic. They may direct inquiries to more approachable local intermediaries, such as neighborhood leaders or party brokers, rather than to state agencies themselves.

[3] A "standpipe" is a (usually free) public tap, shared by several households.

Changes in these outcomes can only be expected if accurate and timely information reaches individuals responsible for managing household water supply. Given the complexity of Nextdrop's system to collect and disseminate water timing information, we could not assume that this would happen. We therefore develop a causal framework that can be used by researchers and policy analysts to identify blockages in the "information pipeline" for programs that attempt to change household behaviors through information-based interventions. Our framework complements Lieberman et. al. (2014), which begins with households receiving information, and examines why the information may not generate changes in political behavior. Upon finding NextDrop's program to have no robust impacts on most measures of household welfare or of state-society relations, we use our framework to diagnose weaknesses in NextDrop's information pipeline. These include both valveman failure to submit accurate notifications and household failure to leave the family's cellphone with the individual waiting for water. Understanding the process by which the information used in such interventions is collected and disseminated to households is crucial if they are to be implemented effectively and taken to scale.

While we find no impact of NextDrop's program on outcomes such as time spent waiting for water, expenditures on substitutes, or citizen relations with the utility, the program did trigger modest reductions in stress among low-income households. This result highlights the importance of examining non-monetary aspects of household welfare when evaluating the impact of development programs; these aspects could be psychological (as in this study), cultural or social.

In this paper, we first describe the informational intervention we evaluate, and outline our expectations regarding the potential effects of this effort. We draw mainly on literatures in behavioral economics and political science to develop our central hypotheses. As these hypotheses are based on the assumption of delivery of accurate information, we also introduce our information pipeline as a framework for causal chain analysis of informational interventions. We next describe our experimental research design and present our analysis of the program's impact. We draw on our causal framework to diagnose how the information production and dissemination broke down when the NextDrop program was rolled out, and to evaluate the effect of increasing the predictability of services on those who actually did receive accurate notifications. We conclude with the broader implications of our findings.

AN INTERVENTION TO IMPROVE SERVICE PREDICTABILITY: WATER ARRIVAL NOTIFICATIONS

Intermittent water deliveries are sometimes predictable but are more often irregular and unreliable (Galaitsi et al. 2016). This study evaluates the impact of a text-message (or SMS) based program to provide prior notifications regarding water arrival times, and thus reduce the costs of coping of water intermittency. We examine the program's impact on household welfare and on individual political attitudes and behavior. The system was developed for urban India, where cell phone penetration rates are high, and the likelihood that water systems will soon be upgraded to continuous service is low. The service is thus potentially useful in much of urban Asia and Africa, where intermittency is rife, local governments cannot adequately fund water systems, and cell phone penetration is increasing rapidly.

In NextDrop's system, the utility employees ("valvemen") were asked to notify the company by calling a toll-free number whenever they opened and closed valves. NextDrop then sent free text-message notifications with expected water arrival times to individual households, which it had cataloged by "valve area". Notifications could be sent before water flowed through household taps, because it takes some time for water to flow into a valve area and fully pressurize that portion of the network. NextDrop reasoned that a notification 30 to 60 minutes prior to water arrival would be sufficient to allow customers to return from nearby schools and shops, though not from distant workplaces; notifications arriving fewer than 30 minutes in advance would not help them unless they were already at home.

To correctly place households in valve areas, NextDrop collected GPS coordinates for households and created valve area maps, which Indian utilities typically do not possess. It drew on valvemen's tacit knowledge regarding the area boundaries, accompanying them on walks around the edges and taking GPS readings. Each polygon in Figure 1 is an example valve area from Bangalore: the city has thousands of these, for which valvemen turn water off and on by manually adjusting a valve. After receiving a valve opening call, NextDrop could automatically let households know when their water would arrive based on their location. NextDrop also assembled valvemen-reported data indicating where water was flowing in the network into a 'dashboard' that they shared with the utility's engineers.

[Figure 1 here]

Figure 1. Example valve areas in BWSSB Subdivision E3

The company piloted its system in Hubli-Dharwad, India, a city of approximately 1 million where most residents received water services every 4-5 days (Burt and Ray 2014). NextDrop surveys of, and relationships with, their Hubli-Dharwad customers suggested that they valued notifications enough to pay a fee for them.[4] In 2013, NextDrop expanded to the megacity of Bangalore, with the support of the city's water utility, the Bangalore Water Supply and Sewerage Board (BWSSB). The company had already understood the importance of securing valveman cooperation. For their Bangalore rollout, NextDrop negotiated an MOU with the utility that made the notifications part of the valvemen's job description; the notifications were delivered free to the customers. The company hand-delivered reports on valveman notification rates to their supervisors on a weekly basis, so managers had the information they needed to enforce compliance. From NextDrop's perspective, this was a scalable approach. We worked with NextDrop to structure its rollout in Bangalore as a cluster-randomized trial so that the effects of their services could be measured more precisely.

EXPECTED PROGRAM IMPACTS AND THE INFORMATION PIPELINE

[4] NextDrop conducted a survey of 200 households in June 2010 in Hubli-Dharwad; according to its records, respondents indicated a willingness to pay INR 8 per month on average for water arrival information (1USD = INR 45 in 2010). In June 2011, it surveyed 60 enrolled households in its pilot, asking whether they would prefer to remain enrolled or receive an INR 5 cell phone recharge. 94% preferred to continue with the service. By August 2013, the firm had 15,844 paying customers, suggesting significant demand for its services.

In this section, we first outline our prior expectations regarding the potential effects of the NextDrop program in Bangalore. As the program was piloted in Hubli-Dharwad, we recognized that it might not work in the same way in a different location; there are several steps between the production of information and its delivery to the intended recipient, all of which are context-dependent in practice. We therefore develop a causal framework highlighting the key steps upon which such informational interventions depend. We present this information pipeline in the second part of this section.

We draw on a diverse set of literatures to develop our hypotheses because there is little research directly on our topic. The causes and consequences of intermittent and unreliable water supplies have been more the domain of engineering and public health research than of the social sciences (Galaitsi et al. 2016). The economics scholarship on urban water has focused on expenditures needed for expanded water access (e.g. Araral 2008; Hutton and Bartram 2008), access to household connections (e.g., Gamper-Rabindran, Khan, and Timmins 2010; Devoto et al. 2012), and the demand for and impacts of better water quality (e.g., Jalan and Somanathan 2008; Madajewicz et al. 2007). Urban political ecology has discussed the inequities embedded in water networks (e.g., Bakker 2003; Swyngedouw 1997) and the ways in which the poor strategize for reliable access (e.g., Ranganathan 2014; Björkman 2015). The (scant) water policy literature on intermittency has analyzed the effects and coping costs of intermittent water (e.g. Burt and Ray 2014; Kumpel et al 2017) but has not examined interventions that improve predictability within intermittent systems.

We build on behavioral economics, political science, and urban water policy insights to develop our hypotheses on the effects of providing accurate information to improve the predictability of water services on: a) household welfare; and b) citizens' relationship with the state.[5] As our discussion of potential heterogeneous treatment effects will suggest, however, household welfare effects and political

[5] Our specific hypotheses and research design were recorded in pre-analysis plan 20150514AA, registered with EGAP prior to the receipt of our baseline data. See also the Online Appendix, Section 2.

effects could be related. Households that experience a change in welfare indicators, such as reduced time waiting for water, may also exhibit changes in political attitudes and behavior.

Household Welfare Effects

The water policy literature has shown that intermittent water supply imposes significant costs on households, especially with respect to water quality and human health (Kumpel and Nelson 2013; Ercumen et al. 2015). Studies that have focused explicitly on the *unpredictable* nature of water delivery in intermittent systems suggest that unpredictable supplies impose substantial costs, especially upon household members tasked with managing the water (Pattanayak et al. 2005; Subbaraman et al. 2015; Zérah 2000). In low-income households that cannot afford maids or automatically-filling storage tanks, household members—particularly women—may stay near the home for substantial periods of time to ensure they can collect and then store water when services finally commence. The "waiter" thus devotes time to waiting that might otherwise be spent on work, or on community, social and family events. We therefore hypothesized that the receipt of accurate prior notifications regarding water delivery times and service disruptions would reduce the amount of time spent waiting for water, allow for more frequent participation in community and social activities, and result in fewer foregone earnings.

The literature on urban water supply also documents the extent to which municipal water costs less than substitutes such as bottled water and water from vendors (e.g. Estache, Gomez-Lobo, and Leipziger 2001; Kjellén and McGranahan 2006). We therefore hypothesized that notifications regarding the timing of water delivery would reduce the reliance on substitutes, because they would decrease the probability of missing a supply.

Finally, supply unpredictability may impose psychological costs. Given that domestic water is a vital resource, the household member responsible for managing water supply may be under stress when services are unpredictable or water storage cannot be planned. This argument builds directly on empirical studies of psychosocial stress related to water insecurity (e.g. Wutich and Ragsdale 2008; Stevenson et al. 2012), as well as the behavioral economics literature, which has shown that scarcity imposes cognitive stress (Mullainathan and Shafir 2013). This led us to hypothesize that receiving accurate prior notifications would reduce respondents' worry about missing a water supply, and the extent to which respondents found themselves thinking about water while doing chores or other work.

While these effects may be observable across the entire urban population in cities with water intermittency, we expected them to be particularly pronounced for *low-income households*, because the cost of substitutes for piped water as a fraction of household income is greater, and because, as noted earlier, poverty itself may exacerbate stress. We also expected effects to be greater for households that do not have automatically-filling overhead or underground tanks ("sumps"); overhead tanks in particular cannot be supported on structures of poor construction quality.

Political Effects

There is also reason to expect that--even in the absence of substantive service improvements--better information alone leads to a more favorable view of the local state and its agencies. Here, we build on a broader literature investigating how citizens "see" and relate to the state (Corbridge et al. 2005; Ferguson and Gupta 2002; Evans 2008). Our research questions connect the political science and local public goods literatures to a rich body of research on the role of information and communication technologies in development (or ICTD). The ICTD literature has argued that better and cheaper information on government schemes, commodity prices, water quality, etc., directly influences citizens' views of the state (though not necessarily in a positive direction) (e.g., Madon and Sahay 2002; Tolbert and Mossberger 2006).

Our first intuition was that increasing the predictability of services would improve citizens' image of the state, given that NextDrop's messages in Bangalore were sponsored by the state-run utility. Even if services remain less frequent than citizens may desire, receiving accurate, prior information regarding service timing (and cancellations) should not only make services easier to access, but also convey greater state capacity and control to consumers. In addition, the innovative application of text messaging to disseminate service schedule information to citizens could make the state seem more "modern," and thus improve citizen perceptions of governmental competence (see Harriss 2006; Kuriyan and Ray 2009; Ghertner 2011) .

We also hypothesized that these notifications would shift perceptions regarding who is responsible for addressing citizens' concerns. The literature on citizen-state interactions in the developing world suggests that ordinary citizens often turn to political intermediaries or direct action when they have service problems.[6] Groups of households and community leaders may collectively protest at government offices.[7] These patterns may slowly shift with the introduction of a universally administered notification system (like Nextdrop's) that connects citizens more directly to the service provider. Citizens may become more likely to view government agencies themselves, rather than local intermediaries, as responsible for addressing their problems; they may also feel that the state is more universalistic, as notifications would be sent not just to favored groups, but to all citizens. Changes in perceptions of government agencies might also stem from the automatic arrival of information that relieve citizens from spending time and effort, or relying on personal connections or bribes, to gain information.[8]

We expect that effects, while relevant across the entire urban population, may be particularly strong among households that consider themselves marginal in religious, social (e.g. caste), or linguistic terms. These segments are less likely to have politically influential intermediaries prior to the intervention.[9] We also expect effects to be stronger for populations that experience greater coping costs, such as households with less money to buy non-tap water, and those without automatically-filling storage tanks.

Information Production and Dissemination

[6] On Bangalore, see Ranganathan (2014); on similar dynamics elsewhere in India, see Berenschot (2010) and Krishna (2011). On intermediaries within clientelistic party systems in the developing world, see Stokes *et al.* (2013).

[7] See Ranganathan (2014) regarding water protests in Bangalore; also Auerbach (2016) regarding urban services in India.

[8] Our emphasis here diverges from the political science and economics literatures on informational interventions, which examine whether information increases political participation or changes votes, thus fueling bottom-up pressure to make service providers more accountable (see Pande 2011; Lieberman, Posner, and Tsai 2014).

[9] We expect this to be the case even though recent studies suggest that India's marginal populations are now better able to exert pressure upon the state than previously (e.g. Corbridge et al. 2005; A. Banerjee and Somanathan 2007).

These hypotheses can be expected to hold only if the household "waiter" receives regular, accurate information from NextDrop. Yet there is a substantial literature suggesting that this assumption may be naïve. The broadest critiques argue that information cannot be reified into a "thing" with well-defined properties and effects, existing independent of collection, codification, dissemination and interpretation (e.g., Nunberg 1996; Srinivasan, Finn, and Ames 2017). Information is produced and received, and these processes are both political and situated in the given context. Delivering accurate information delivered on time is thus a complex undertaking that can easily fail. Even when successfully executed, it can only have impact if the recipient has both the ability and willingness to act on it (Lieberman et al 2014).

To guide our analysis, as well as that of other researchers analyzing informational interventions, we develop an information pipeline, or causal framework intended to capture potential sources of leakage in informational interventions. While Lieberman *et al*. (2014) provide a framework for assessing the conditions under which accurate, effectively-disseminated information will generate a causal effect, our framework spells out key nodes in the information production and dissemination process itself, or the stages prior to Lieberman *et al*.'s framework. Understanding the production and dissemination parts of the causal chain in the real world is crucial (see Pande 2011); major shifts in the informational environments in the developing world will eventually require transformed public sector bureaucracies rather than a host of small, researcher-administered experiments.

Figure 2 displays six ways in which informational interventions can break down. First, the entity responsible for collecting the information to be disseminated may not obtain it, or obtain only a subset thereof. This could occur, for instance, if frontline workers fail to supply the information requested of them (Hyun, Post, and Ray Forthcoming). Second, the entity responsible for providing information may not be able to analyze or compile the information in the desired format. Third, the entity charged with disseminating the information may not in fact send it; this can occur for a variety of reasons, including poor incentives and technical glitches. Fourth, information may be sent, but respondents may not receive it. Messages may be lost in transmission or go to the wrong person because phone numbers change or because the person for whom the message is intended does not keep the household cell phone. Fifth, the person for whom the message is intended may technically receive messages, but not *register* receipt. This can occur, for example, if information is sent in the wrong language, or if he or she is inundated with messages, or if the information is so useless that the respondent simply stops paying attention. Finally, the information sent may actually be inaccurate. Inaccuracies could reflect deliberate efforts to conceal information, carelessness, or a lack of measurement ability. Researchers evaluating informational interventions must consider each of these possibilities, though all nodes may not be relevant for all interventions. Understanding and collecting data on these steps in the production and and receipt of information can help researchers understand why a particular intervention may not have had the expected impact, and to identify ways in which the programs they are evaluating—even if successful—can be improved.

[Figure 2 here]

Figure 2. Information pipelines for transparency interventions

RESEARCH DESIGN

We evaluated the effectiveness of the NextDrop system through a cluster-randomized experiment in Bangalore, a city of over 8 million that is often called India's Silicon Valley. Prior scholarship on domestic water, particularly on water intermittency, has not been field-experimental in nature. The handful of empirical studies on the coping costs and inefficiencies associated with unreliable water supply have either been observational or stated-preference based experiments (Akram and Olmstead 2010; Baisa et al. 2010; Dauda, Yacob, and Radam 2014; Pattanayak et al. 2005; Subbaraman et al. 2015; Zérah 2000).

To design our evaluation, we worked closely with NextDrop and the Bangalore water utility, BWSSB. We structured the study to evaluate the efficacy of NextDrop's system in a real world setting—their existing efforts to scale up in new cities—which meant deferring to some of their implementation decisions, so long as these did not interfere with our ability to randomize assignment and ensure noninterference between treatment and control. This section outlines the main features of our study, which we conducted in two waves in 2015.

Study Site Characteristics

We conducted our impact evaluation in a socio-economically diverse section of Bangalore chosen to maximize similarities with other Indian urban centers. NextDrop had recently received approval from BWSSB to introduce services across the city.[10] Because NextDrop wanted to respond to BWSSB's request to roll out quickly, we were asked to restrict our evaluation to only one of the utility's 32 subdivisions. We limited our search to subdivisions not scheduled for immediate expansion, which meant areas mostly outside the city center. Because water services are typically more intermittent and unreliable further from the center, this meant we were considering areas where the intervention was more likely to have an impact. It also meant that our study location was more typical of water service in urban India more broadly, as BWSSB performs well in relation to other South Asian utilities (Connors 2005; McKenzie and Ray 2009).

After review of the limited (and somewhat inaccurate) government data on low-income settlements and population densities in Bangalore, and extensive site visits throughout the city in 2014, we chose to conduct our evaluation in BWSSB subdivision E3 (Figure 3). Our exploratory fieldwork increased our confidence that the impacts of NextDrop's services were likely to be higher in low-income areas with residential structures of one to two stories, because the roofs for such buildings typically cannot support overhead storage tanks. However, we (and NextDrop) were interested in understanding how impacts varied across different income groups. We therefore sought a subdivision that contained a diverse population (where roughly 1/3 of our sample could be drawn from the bottom third of the city's income distribution), and that possessed a reasonable number of low-rise structures.

[10] May 2014 memorandum of understanding.

[Figure 3 here]

Figure 3. BWSSB Subdivision E3

The subdivision contained low-rise residential neighborhoods, as well as several low- and middle-income neighborhoods, in sufficient numbers. Data from our baseline survey suggests that approximately 33% of the area's residents – 14% of whom included recent migrants from states such as Tamil Nadu and Andhra Pradesh – could be classified as Bangalore's bottom third of the income distribution. Over 85% of residents received water services once or twice a week, which is common in urban India, and also frequent enough that we could detect the effect of the notifications on our outcomes of interest, if they were indeed useful. Moreover, 28% of residents possessed neither an automatically filling overhead tank nor sump, requiring one to be present at the time of water arrival to store water for use between supplies. The area thus appeared suitable for an analysis of how the impact of NextDrop's intervention may vary according to the variety of criteria we had outlined, and in a setting that is representative for urban India.

Randomization and Sampling Strategies

Within BWSSB subdivision E3, we employed a cluster-randomized experimental design. We opted for cluster- rather than household-level randomization, because of concerns regarding information sharing between treatment and control households. We further separated clusters of households in our study from one another by at least two streets so as to create physical buffers preventing information sharing between our treatment and control groups (Figure A.1, Online Appendix). Moreover, spillovers were unlikely because information on water arrival times is relevant only to individuals within the same valve areas (typically 50 to 200 households); individuals would have little incentive to share information with those living in other clusters.[11] Our study, powered using pilot data to detect a reduction of 30-45

[11] In an ideal world, we would have randomized assignment to valve areas rather than clusters we ourselves designated. After discussions with our survey team, we realized that, because valve area boundaries are not visible above ground and do

minutes in the time spent waiting for water per week, included 120 clusters comprised of 300 households in total. [12]

Because blocking on a variable associated with the outcome of interest can improve the precision of causal estimates in cluster-randomized experiments (Imbens 2011), we employed a geographic approach to stratification. Blocking on socio-economic geography also enabled analyses of subsets corresponding to areas where we expected to observe stronger effects: those with poorer residents and with poorer quality water infrastructure. Based on extensive site surveys, we designated 30 geographic blocks with a particular socio-economic character, either low income (10 blocks) or mixed income (20 blocks). Each block included four clusters that we expected to be similar not only in socio-economic terms, but also in terms of the state of the underlying water infrastructure. Within each block, we randomly assigned two clusters receive treatment and two to the control condition.[13]

Data and Measurement

We measured the impact of the intervention through two surveys administered to the treatment and control groups. A baseline survey was conducted prior to the intervention in April and May of 2015, and an endline survey was conducted in October and November 2015, after treated households had received services for four months. We ran the trial for four months to give households enough time to adapt their daily routines to the service. Enumerators asked to speak with the individual who managed and stored water for the household. If the "waiter" was unavailable, his or her name was noted and an

---

not follow the street layout, survey enumerators would have had difficulty following even boundaries drawn on maps. Substituting cluster-level for valve area randomization led to only minimal spillovers (see below).

[12] For more details, see our pre-analysis plan (Online Appendix). These calculations presumed that we would lose approximately 20% of our sample through attrition and that 20% of households would refuse to sign up for services.

[13] Given the lack of accurate state data on the existence and location of the city's numerous and scattered small slums, identifying an area with a suitable demographic mix required significant on-the-ground legwork by our team. We included four clusters per block rather than two following Imbens (2011). Blocks and clusters were designated following exploratory site visits throughout E3. Within each block, we outlined four clusters separated by two streets or lanes from one another. Within each cluster, we followed a systematic sampling plan with a skip of three between households on every street. After piloting the survey in low-income areas, we decided that a skip of three would be sufficient to avoid group interview sessions in which neighbors "help" respondents answer survey questions. The skip of three also allowed us to get a sample of 25 households per cluster, even though our clusters were small in area given the need to ensure socio-economic comparability between clusters within a block.

appointment made to return for an interview. Because women typically manage water, 80% of our respondents were women.

Enumerators concluded the baseline survey by offering all households the opportunity to enroll in NextDrop services, when they became available in their area, by submitting their cell phone numbers and offering consent. Offering services to both treatment and control allowed us to employ a placebo design to help identify compliers—i.e., those who would accept treatment—in both groups. Respondents were given the option of signing up for text or voicemail notifications in English, Kannada, Telugu, and Tamil, and informed that the service was sponsored by BWSSB, the state water utility.[14] NextDrop enrolled the households in our treatment group following the completion of the baseline survey, and waited until the end of our study to enroll the control group.

Enumerators also collected GPS coordinates from each household. This allowed NextDrop to correctly place treatment group households in valve areas (so they received accurate information), and helped our team to verify that enumerators had not strayed outside cluster boundaries. Coordinates also assisted with returning to the same households in the second wave. To ensure that GPS readings had the 5m precision required by NextDrop, we configured our survey software to prevent enumerators from beginning each survey until sufficiently precise coordinates had been obtained.

Comparing key characteristics between our treatment and control groups, we see that our cluster-randomized design achieved balance between treatment and control with respect to household characteristics, water supply conditions, and political factors (Online Appendix, Table A.2). As is to be expected with two-wave designs, our sample did experience attrition: we lost 16% of our initial sample between waves 1 and 2, often because households (usually renters) had moved to other neighborhoods. Attrition did not affect covariate balance (Table A.2).

[14] Forms describing the NextDrop service were also translated into these languages. Interviews were typically conducted in Kannada (the primary language spoken in Bangalore), but were conducted in Telugu or Tamil when relevant (usually for recent migrants).

Conditions at Baseline and Implications for Aggregate Impacts

A first condition for the intervention to generate an effect would be that treatment group members indeed faced costs due to unpredictable water services. If households were not spending time waiting for water, relied infrequently on substitutes for tap water, viewed the utility as highly competent, and already contacted the utility directly regarding service problems, the intervention could not generate shifts in behavior and attitudes. We conducted pilot surveys in subdivision E3, prior to confirming our choice of this area for the impact evaluation, and these pilots suggested room for movement on key outcomes of interest.[15]

Analysis of population means from our baseline survey shows that there was indeed room for movement on outcomes such as waiting times for water, use of substitutes for piped water, and tendency to contact the utility directly regarding service problems. Moreover, our baseline data suggest that many households faced difficult water supply conditions. A full 69% of our households reported that their water did not come at a specific time. Additionally, 43% reported that they simply learned that water had arrived when it began to come out of their taps, rather than knowing when to expect the water from the supply schedule, valvemen, or local leader. There was less room for movement on outcomes such as missing work due to waiting for water, or attitudes towards the utility, which were already quite favorable (Table 1).

Table 1. Conditions at Baseline

[Table 1 here]

RESULTS

Though baseline conditions provided ample room for NextDrop's notifications to have household-level impacts, intention-to-treat (ITT) estimates for average treatment effect show no statistically significant change for any outcome variables across the entire study population, except those

[15] For outcomes such as wait time and expenditures on substitutes we also consulted survey data from Hubli-Dharwad, collected during one of the authors' previous research efforts, which suggested substantial room for movement.

related to worry and stress (Table 2).[16] They also show no effects when we restrict our analysis to our ten low-income blocks, which contain lower-income populations and less variability in water infrastructure. In addition, they show no robust effects for our target group -- low-income households without automatically filling tanks. Tests for many other heterogeneous effects outlined in our pre-analysis plan also do not yield statistically significant or substantively important effects.[17] In other words, the NextDrop program failed to generate discernible impacts on household welfare or state-society relations in the treatment group. This could be because the estimated average treatment effect on (reported) time spent waiting for water is very small—about 2.5 minutes for the overall population. Further analysis suggests that we should have been able to detect reasonably sized effects, had they been present: the minimum detectable effect for our wait time outcome is 9 minutes.[18] It is unlikely, however, that reductions of this size would impact household wellbeing.[19]

We do detect a small but measurable decrease in worrying about water and thinking about water during the day among the overall population and within our low-income blocks.[20] It is possible that there was an effect for the target population, but we simply do not know, as the experiment was not powered for these outcomes. While these results may seem surprising given that the intervention did not affect wait times for water -- the variable through which we expected psychological effects to be mediated -- it may be that individuals worried less because they *felt* more informed with NextDrop notifications. Indeed, about 85% of our treatment group reported in our endline survey that they found

---

16 See the Online Appendix for results without covariate adjustment (Table A.3).

17 Results available upon request. See the Online Appendix for our pre-analysis plan.

18 We estimate power as Power = 1- ϕ (1.96 - Effect size/SE) + ϕ (-1.96 - Effect size/SE), where ϕ is the cumulative distribution function for a standard normal random variable, and SE is the standard error for the average effect size where standard errors are clustered at the cluster level. The reported power is the probability that the null hypothesis of a zero average treatment effect is rejected at the 5% level. The minimum detectable effect (MDE) is an estimate of the smallest effect size that would yield a test with 80% power (see e.g. Miguel et al. (2016)). We used the following formula: MDE= $(1.96+0.84)\sigma$, where σ is the standard error of the coefficient on the treatment indicator in a regression including the treatment indicator, relevant covariates, and clustered standard errors (the regression model we use in our analysis throughout the paper).

19 These results do not appear to be driven by spillovers between the treatment and control groups. In our endline survey, we specifically asked respondents who reported that they had received notifications whether they had received them from the utility, BWSSB, or other sources. Only 11 respondents reported receiving notifications from anyone other than BWSSB.

20 Note these decreases are slightly less significant statistically after adjustments for multiple hypothesis testing.

the NextDrop notifications “useful;” roughly 74% of individuals who claimed that notifications were rarely or never accurate nevertheless found them useful.[21]

Table 2. Intent-to-Treat Estimates (with covariate adjustment)
[Table 2 here]

Leakage in the Information Pipeline

To explain these results, we turn to our causal framework for analyzing information flows in transparency interventions. Using this framework, we identify several breakdowns in NextDrop’s system: valvemen often failed to submit notifications to NextDrop, many household “waiters” either did not receive or did not register receipt of NextDrop notifications, and many notifications were inaccurate. These observations suggest that failure in the production and dissemination of information played a large role in the program’s failure. They also suggest that conducting tests of our hypotheses about the effect of receiving *accurate* prior notifications would require focusing on the subset of households actually receiving quality information.[22]

Following our framework, we identify the main blockages in the NextDrop pipeline, and how each contributed to reductions in the effective size of our treatment group (Figure 4). We first examine the extent to which valvemen submitted information by calling NextDrop’s automated voice mail system to log water valve opening times—the key factor in NextDrop’s ability to collect the information it aimed to disseminate. Comparing our geo-coded survey responses with logs of the valvemen reports to NextDrop, we find that valvemen sent reports to NextDrop approximately 70% of the time.[23] Reporting

---

[21] This somewhat significant effect is visible in our CACE analysis as well, with larger effects for households actually receiving accurate notifications.

[22] We registered our revised empirical strategy as an amendment to our pre-analysis plan prior to the receipt of data from our endline survey (see the Online appendix).

[23] We reached the same percentage through two different calculations. First, we analyzed the number of valvemen reports a week to NextDrop as a percentage of expected reports for each valve area, based on the official utility supply schedule, for 4 weeks prior to the endline survey. In addition, we compared household survey responses naming the last water supply day with valveman reports for each valve area for the week preceding the endline survey. The geo-coded nature of our data facilitated this analysis. Moreover, our parallel, ethnographic study of the valvemen in the NextDrop intervention (in a different subdivision) also found that they did not submit information regularly, or sometimes submitted inaccurate notifications—e.g., sending a round of notifications during tea breaks rather than when actually turning on water valves (AUTHOR, Forthcoming).

rates were equivalent for the treatment and control groups.[24] Valveman non-reporting, it appears, accounted for significant leakage in the information pipeline, reducing the effective size of our treatment group from 1193 to 854.

For information to have an impact, it must not only be sent, but must also be received. Only 38% of treatment group members (453 out of 1193) reported receiving notifications at least once every two weeks, a much lower percentage than the 70% rate at which valvemen were regularly submitting notifications (Figure 4). This gap between treatment assignment and actual receipt of messages becomes larger for the populations for which we expect the intervention to have the greatest effect, i.e., lower income households without automatically-filling tanks. For this target group, only 25% of treatment households reported receiving notifications.

[Figure 4 here]

Figure 4. NextDrop's leaky information pipeline and treatment group attrition

What explains the low rate at which treatment group households reported receiving messages? We must first consider whether information was lost in the transmission process. There is one clear way in which this occurred in our study: many household "waiters" for water did not possess the household cellphone registered with NextDrop. Among our 854 treatment group households that were regularly sent notifications, this was the case for 207, reducing our effective treatment group size to 647. Gender differentials in mobile access, then, may have prevented many individuals responsible for managing household water from actually receiving notifications, and diluted any impact even accurate notifications might have had (Figure 4).[25] The marginal drop-off associated with differential access, however, was not as large as that associated with valveman non-reporting.

[24] Reports were sent at least 70% of the time to 72% of the treatment group households, and to 75% of the control group households. Control group households did not receive NextDrop notifications. Household refusals totaled only 3% of our sample. Because of our placebo design for enrollment, we can identify the set of non-compliers for both the treatment and control groups

[25] Women did not possess the household cell phone in similar proportions among poor and middle class households.

Given that cellular phone services are quite reliable in urban India, we infer that much of the remaining discrepancy between the number of households sent messages and registry of message receipt can be attributed to respondents simply not noticing NextDrop's notifications. They may have been inundated with text messages, or not bothered with notifications if the first notification they received did not appear useful.[26] Non-registry of receipt appears to account for a reduction of 647 to 453 in our effective sample size (Figure 4).

Finally, to be useful, information must be accurate. Only 289 of our 1193 treatment group respondents reported that the information received in NextDrop notifications was either always or usually accurate. To understand the extent of information inaccuracies, we compared household survey responses about the last day they had received water and the average time of water arrival to the time-stamped and geo-coded valveman reports for the relevant valve areas. Our analysis shows that, where valvemen had submitted reports in the week prior to our endline survey, 36% of households reported receiving water on a different day than that reported by the valveman. Furthermore, a comparison of household reports regarding average water arrival times with valveman report data suggests that 62% of households received reports *after* the water arrived (Figure A.2, Online Appendix).[27] These inaccuracies further reduced our effective treatment group size from 453 to 289. They may also explain why so many households who were sent notifications did not register receiving them; those charged with waiting for water may have not understood their purpose or begun ignoring them. Thus serious problems of data accuracy compounded already significant problems with non-reporting.

Effects of Receiving Accurate Information

[26] We did not ask survey respondents how many text messages they received on an average day. Study participants could choose to receive notifications via text, or voicemail, in their chosen languages, so it seems unlikely that treatment group members did not understand the notifications they received.

[27] This percentage was calculated for the 33% percentage of households who were able to name a specific time when their water typically arrived residing in valve areas where valvemen had issued reports in the month prior to their survey interview.

Given the extensive leakage in NextDrop's information pipeline, evaluating our original hypotheses, rather than simply assessing the impact of NextDrop services more broadly, requires calculating effects for those actually receiving accurate information. As described above, only 453 of our 1193 treatment group households reported receiving NextDrop notifications regularly. Of these, 289 reported receiving *accurate* notifications. Given these high rates of noncompliance, or "incomplete administration of treatment" (Angrist 2006), a precise assessment of the impact of receiving notifications entails analyzing the causal effects for compliers. We define compliers as those who actually received the treatment of interest, namely accurate information.

Table 3 presents CACE estimates for the subset of households reporting that they received accurate notifications. The tables include results for the entire study population, low-income blocks, and our target households. Similar to the ITT results, we observe no significant differences between the treatment and control groups for most outcomes.[28] Meanwhile, CACE estimates calculated based on message accuracy for only low-income households without any sort of water storage tank suggest that target group households reduce reliance on substitutes and are more likely to contact the utility directly.[29] Accurate notifications do, however, appear to have reduced levels of stress and worry among low-income households to a greater extent than suggested by the ITT results. A larger effect size among those receiving accurate information supports our original hypotheses about the effects of predictability on stress and worry.

While we can clearly report that the NextDrop system had modest effects at best within our study area, our conclusions regarding the effects of accurate notifications are less certain. The various forms of noncompliance we describe eroded the power of our study. As discussed in our pre-analysis plan, our

---

[28] For results without covariate adjustment, see the Online Appendix Table A.6. CACE estimates for the somewhat larger subset of households simply reporting that they received notifications—accurate or inaccurate—are similar (Tables A.3 and A.4). As a robustness check, we calculated the average treatment effect for those receiving accurate messages based on a different measure of message accuracy: whether or not geo-coded survey responses regarding the last day water had arrived, and the time water usually arrived, corresponded with valveman reports from the preceding week. These analyses suggested modest impacts at best (Online appendix, Tables A.6 and A.7.)

[29] Results available upon request. This analysis was not specified in our pre-analysis plan.

experiment sample size was calculated to achieve 80% power based on the assumptions of an 80% compliance rate along with a 20% attrition rate between the baseline and endline surveys. Yet if we define a complier as one who received the intended treatment, namely accurate information about water arrival, our compliance rate is 23.5% (Online appendix, Table A.9). We estimated our ability to detect a CACE given this percentage of compliers and the distribution of these compliers across clusters using simulations, following Arnold *et al*. (2011).[30] Our simulations suggest that we are substantially underpowered to detect a reduction in wait time for low-income households that lack automatically-filling overhead water tanks (Online Appendix, Table A.10).[31] We would expect changes in most of our outcome variables to be triggered by reductions in the time spent waiting for water, and the ATE for wait time is large and negative across different analyses, so it may be that a larger sample size would have allowed us to detect significant impacts on several dimensions.

---

30 The method described is amended slightly such that 1) only those defined as compliers are simulated to receive a treatment effect and 2) the simulations calculate the ability of the experiment to detect a CACE using two staged least squares as opposed to detecting an ITT effect using standard regression analysis, and is based on the same assumptions regarding effect size, etc. as our original power calculations.

31 For example, when calculating a CACE where those who receive accurate messages are classified as compliers, simulations were able to detect an effect significant at the 5% level only between 9.5% and 32.8% percent of the time.

Table 3. CACE for Households Receiving Accurate Notifications (with covariate adjustment)

[Table 3 here]

CONCLUSION AND DISCUSSION

Populations throughout the developing world suffer from unpredictable service delivery. Erratic public transportation makes commuting across congested cities difficult, while unpredictable power outages make it difficult to complete everyday household tasks. Our study focuses on the predictability of perhaps *the* most important service for human development outcomes: piped water for domestic use. We highlight the costs imposed by intermittent and erratic water services (see Table 1 for examples); we also emphasize the need for scholars of development, and political economy more broadly, to expand their usual focus on expenditures and access to include measures of service quality such as frequency and predictability. These are familiar analytical variables in engineering, public health, and energy and water policy, but much less so for the mainstream social sciences.

In our study, we examine the impact of an informational intervention designed to reduce the coping costs associated with water intermittency and provide a causal framework for analyzing the ways in which such interventions can break down. As Pande (2011) notes, while a large literature examines the political effects of providing information, we know little about how broad shifts in the information environments in developing countries may be achieved. Our framework highlights six potential blockages between information generation and actionable knowledge: failure to collect the intended information, failure to perform required analyses, failure to disseminate the information; citizen non-receipt of information due to technical or other factors; citizen non-registration of information received because of logistical or linguistic factors; and the provision of inaccurate information. NextDrop's information pipeline in particular remained ineffective mainly because the person waiting for water often did not possess the family cell phone, and because valvemen often submitted inaccurate information when they submitted it at all. Future studies of informational interventions should be structured to collect data on each of these potential bottlenecks. This will not only allow researchers to explain null results, but also identify points of leverage for program improvements.

In spite of breakdowns in NextDrop's information pipeline, our experimental evaluation of the impact of NextDrop's water notification system during its rollout in Bangalore suggests that the main impact of the program was a modest reduction in stress levels associated with managing household water among low income households. These results indicates that stress as an indicator of household welfare should be examined more frequently in development interventions; this in an important, and often gendered, outcome in its own right.

Aside from changes in stress levels, our experimental results suggest that NextDrop's program failed to trigger changes in household welfare or state-society relations in Eastern Bangalore. However, our power to detect the impact of accurate notifications, and indeed the long run viability of NextDrop's programmatic model, was eroded by failures in the production and dissemination of the information around which this program was conceived. To the extent that our findings result partly from the one cell phone in the household being kept mainly by the men in our study site, our study reminds us that, too often, development interventions still treat households as a unitary construct, undifferentiated by internal gender dynamics (see Alderman et al. 1995). To the extent that our findings result partly from non-cooperation by valvemen, our study serves as a reminder that frontline workers in public sector bureaucracies will play key roles when small-scale interventions are brought to scale. It makes more sense to study interventions in real world settings, using government bureaucracies, and in light of existing household structures, than in artificial settings with implementers or respondents answering to the research team. As external evaluators, we did not incentivize valvemen to cooperate with Next Drop because we wished to study the actual intervention that NextDrop and BWSSB planned to implement.[32] We also did not consider incentivizing men to leave their phones at home when they went to work.

Frontline workers are crucial to the successful implementation not only of informational interventions, but development programs more broadly. Yet many influential experimental evaluations in the development literature mention the roles of the human last mile in their methods sections, but do not return to their potentially critical roles when discussing (and explaining) their findings. This is especially the case for studies that report a "successful" result. For example, Cohen and Dupas (2010) find that free bednets significantly raise the numbers of Kenyan households that use insecticide-treated bednets; Banerjee et al. (2010) find that small incentives such as a bag of lentils increase the probability that a mother will take her baby to a clinic to be immunized; Blattman *et. al* (2014) find a significant positive impact on negotiated conflict resolution in a joint UN- and NGO-led experiment in Liberia. None of these papers discuss the skills and characteristics of the NGO workers or local intermediaries in explaining their positive results. The social sciences have a rich tradition of analyzing the motivations, performance and discretionary power of street-level bureaucrats, or last mile workers, on access to public services such as education, healthcare, policing and social security (e.g. Lipsky 1980; Portillo and Rudes 2014), upon which such analyses could draw.

Our findings also highlight the importance of systematically investigating the extent to which both null *and* positive field experimental results depend upon the process of information production and other key elements of program implementation. Development research has recognized the importance of understanding when and why particular interventions fail, but success is important to explain as well (Karlan and Appel 2016). If an experiment suggests that a program is effective, funders may decide to replicate it elsewhere (to test for external validity) or to roll it out more generally (in a "similar" context to that of the experiment). Yet as we and others (e.g., Ananthpur, Malik, and Rao 2014; Bold et al. 2013)

[32] After struggling with valveman noncompliance in Bangalore and Mysore during 2015, NextDrop decided to switch models. Just as we ended our experiment, they decided to measure water arrival times directly through low cost sensors installed in household storage tanks. This proved unworkable, and they discontinued their service completely in May 2016.

demonstrate, successful replication in another context is not guaranteed. Among other factors, replication would entail successful implementation by front-line workers whose cooperation we cannot take for granted; not every town, village or community has experienced NGOs or trusted facilitators or charismatic elders. Furthermore, rollout could be met by different levels of cooperation and adherence by the intendend beneficiaries themselves (see e.g. Rammelt and Leung (2017)). These problems are difficult to anticipate before implementation; they require careful analysis of each step of the information pipeline for experiments whose results support the research hypotheses as well as in those whose do not.

REFERENCES

Akram, Agha Ali, and Sheila M. Olmstead. 2010. "The Value of Household Water Service Quality in Lahore, Pakistan." *Environmental and Resource Economics* 49 (2): 173–98. doi:10.1007/s10640-010-9429-7.

Alderman, H., P. A. Chiappori, L. Haddad, J. Hoddinott, and R. Kanbur. 1995. "Unitary versus Collective Models of the Household: Is It Time to Shift the Burden of Proof?" *World Bank Research Observer* 10: 1–19.

Ananthpur, Kripa, Kabir Malik, and Vijayendra Rao. 2014. "The Anatomy of Failure: An Ethnography of a Randomized Controlled Trial to Deepen Democracy in Rural India." Policy Research Working Paper. Washington D.C.: World Bank.

Angrist, Joshua D. 2006. "Instrumental Variables Methods in Experimental Criminological Research: What, Why and How." *Journal of Experimental Criminology* 2 (1): 23–44.

Araral, Jr, Eduardo. 2008. "Public Provision for Urban Water: Getting Prices and Governance Right." *Governance* 21 (4): 527–49.

Arnold, Benjamin F., D. Hogan, John M. Colford, and A. Hubbard,. 2011. "Simulation Methods to Estimate Design Power: An Overview for Applied Research."." *BMC Medical Research Methodology* 11 (94).

Auerbach, Adam Michael. 2016. "Clients and Communities." *World Politics* 68 (01): 111–48.

Baisa, Brian, Lucas W. Davis, Stephen W. Salant, and William Wilcox. 2010. "The Welfare Costs of Unreliable Water Service." *Journal of Development Economics* 92 (1): 1–12.

Bakker, Karen. 2003. "Archipelagos and Networks: Urbanization and Water Privatization in the South." *The Geographical Journal* 169 (4): 328–41.

Banerjee, Abhijit, and Rohini Somanathan. 2007. "The Political Economy of Public Goods: Some Evidence from India." *Journal of Development Economics* 82 (2): 287–314.

Banerjee, Abhijit Vinayak, Esther Duflo, Rachel Glennerster, and Dhruva Kothari. 2010. “Improving Immunisation Coverage in Rural India: Clustered Randomised Controlled Evaluation of Immunisation Campaigns with and without Incentives.” *BMJ* 340: c2220.

Berenschot, Ward. 2010. “Everyday Mediation: The Politics of Public Service Delivery in Gujarat, India.” *Development and Change* 41 (5): 883–905.

Björkman, Lisa. 2015. *Pipe Politics, Contested Waters: Embedded Infrastructures of Millennial Mumbai*. Durham, NC: Duke University Press.

Blattman, Christopher, Alexandra C. Hartman, and Robert A. Blair. 2014. “How to Promote Order and Property Rights under Weak Rule of Law? An Experiment in Changing Dispute Resolution Behavior through Community Education.” *American Political Science Review* 108 (01): 100–120.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng’ang’a, and Justin Sandefur. 2013. “Scaling Up What Works: Experimental Evidence on External Validity in Kenyan Education.” Working Paper 321. Washington D.C.: Center for Global Development.

Burt, Zachary, and Isha Ray. 2014. “Storage and Non-Payment: Persistent Informalities within the Formal Water Supply of Hubli-Dharwad, India.” *Water Alternatives* 7 (1): 106–20.

Cohen, Jessica, and Pascaline Dupas. 2010. “Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment.” *The Quarterly Journal of Economics* 125 (1): 1–45.

Connors, Genevieve. 2005. “When Utilities Muddle through: Pro-Poor Governance in Bangalore’s Public Water Sector.” *Environment and Urbanization* 17 (1): 201–218.

Corbridge, Stuart, Glyn Williams, Manoj Srivastava, and René Véron. 2005. *Seeing the State: Governance and Governmentality in India*. Cambridge: Cambridge University Press. http://ebooks.cambridge.org/ref/id/CBO9780511492211.

Dauda, Suleiman Alhaji, Mohd Rusli Yacob, and Alias Radam. 2014. “Household’s Willingness to Pay for Heterogeneous Attributes of Drinking Water Quality and Services Improvement: An Application of Choice Experiment.” *Applied Water Science* 5 (3): 253–59. doi:10.1007/s13201-014-0186-6.

Devoto, Florencia, Esther Duflo, Pascaline Dupas, William Pariente, and Vincent Pons. 2012. “Happiness on Tap: Piped Water Adoption in Urban Morocco.” *American Economic Journal: Economic Policy* 4 (4): 68–99.

Ercumen, Ayse, Benjamin F. Arnold, Emily Kumpel, Zachary Burt, Isha Ray, Kara Nelson, and John M. Colford. 2015. “Upgrading a Piped Water Supply from Intermittent to Continuous Delivery and Association with Waterborne Illness: A Matched Cohort Study in Urban India.” *PLoS Medicine* 12 (10). doi:10.1371/journal.pmed.1001892.

Estache, Antonio, Andres Gomez-Lobo, and Danny Leipziger. 2001. “Utilities Privatization and the Poor: Lessons and Evidence from Latin America.” *World Development* 29 (7): 1179–1198.

Evans, Peter. 2008. “In Search of the 21st Century Developmental State.” *The Centre for Global Policity Economy, University of Sussex Working Paper* 4.

Ferguson, James, and Akhil Gupta. 2002. “Spatializing States: Toward an Ethnography of Neoliberal Governmentality.” *American Ethnologist* 29 (4): 981–1002.

Galaitsi, S.E., Robert Russell, Amahl Bishara, John L. Durant, Jennifer Bogle, and Annette Huber-Lee. 2016. “Intermittent Domestic Water Supply: A Critical Review and Analysis of Causal-Consequential Pathways.” *Water* 8 (274): 1–25.

Gamper-Rabindran, Shanti, Shakeeb Khan, and Christopher Timmins. 2010. “The Impact of Piped Water Provision on Infant Mortality in Brazil: A Quantile Panel Data Approach.” *Journal of Development Economics* 92 (2): 188–200.

Ghertner, D. Asher. 2011. “Gentrifying the State, Gentrifying Participation: Elite Governance Programs in Delhi: Gentrifying the Local State in India.” *International Journal of Urban and Regional Research* 35 (3): 504–32.

Harriss, John. 2006. “Middle-Class Activism and the Politics of the Informal Working Class: A Perspective on Class Relations and Civil Society in Indian Cities.” *Critical Asian Studies* 38 (4): 445–65.

Hutton, Guy, and Jamie Bartram. 2008. “Global Costs of Attaining the Millennium Development Goal for Water Supply and Sanitation.” *Bulletin of the World Health Organization* 86 (1): 13–19.

Hyun, Christopher, Alison E. Post, and Isha Ray. Forthcoming. “Frontline Worker Compliance with Transparency Reforms: Barriers Posed by Family and Financial Responsibilities.” *Governance*.

Imbens, Guido W. 2011. “Experimental Design for Unit and Cluster Randomized Trials.” Working Paper, Harvard University.

Jalan, Jyotsna, and E. Somanathan. 2008. “The Importance of Being Informed: Experimental Evidence on Demand for Environmental Quality.” *Journal of Development Economics* 87 (1): 14–28.

Karlan, Dean, and Jacob Appel. 2016. *Failing in the Field: What We Can Learn When Research Goes Wrong*. Princeton, N.J.: Princeton University Press.

Kjellén, Marianne, and Gordon McGranahan. 2006. *Informal Water Vendors and the Urban Poor*. International Institute for Environment and Development London. http://www.mumbaidp24seven.in/reference/informal_water_vendors.pdf.

Krishna, Anirudh. 2011. “Gaining Access to Public Services and the Democratic State in India: Institutions in the Middle.” *Studies in Comparative International Development* 46 (1): 98–117.

Kumpel, Emily, and Kara L. Nelson. 2013. “Comparing Microbial Water Quality in an Intermittent and Continuous Piped Water Supply.” *Water Research* 47 (14): 5176–88.

———. 2016. “Intermittent Water Supply: Prevalence, Practice, and Microbial Water Quality.” *Environmental Science & Technology* 50 (2): 542–553.

Kuriyan, Renee, and Isha Ray. 2009. “Outsourcing the State? Public–Private Partnerships and Information Technologies in India.” *World Development* 37 (10): 1663–73.

Lieberman, Evan S., Daniel N. Posner, and Lily L. Tsai. 2014. “Does Information Lead to More Active Citizenship? Evidence from an Education Intervention in Rural Kenya.” *World Development* 60 (August): 69–83.

Lipsky, Michael. 1980. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service*. New York, NY: Russell Sage Foundation.

Madajewicz, Malgosia, Alexander Pfaff, Alexander van Geen, Joseph Graziano, Iftikhar Hussein, Hasina Momotaj, Roksana Sylvi, and Habibul Ahsan. 2007. “Can Information Alone Change Behavior? Response to Arsenic Contamination of Groundwater in Bangladesh.” *Journal of Development Economics* 84 (2): 731–54.

Madon, S., and S. Sahay. 2002. “An Information-Based Model of NGO Mediation for the Empowerment of Slum Dwellers in Bangalore.” *The Information Society* 18 (1): 13–19.

McKenzie, David, and Isha Ray. 2009. “Urban Water Supply in India: Status, Reform Options and Possible Lessons.” *Water Policy* 11 (4): 442–460.

Miguel, Edward A., Kevin Croke, Joan Hamory Hicks, Eric Hsu, and Michael Robert Kremer. 2016. “Does Mass Deworming Affect Child Nutrition ? Meta-Analysis, Cost-Effectiveness, and Statistical Power.” WPS7921. The World Bank. http://documents.worldbank.org/curated/en/926411482169164466/Does-mass-deworming-affect-child-nutrition-meta-analysis-cost-effectiveness-and-statistical-power?cid=DEC_PolicyResearchEN_D_INT.

Mullainathan, Sendhil, and Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means so Much*. New York, NY: Times Books, Henry Holt and Company.

Nunberg, Geoffrey. 1996. “Farewell to the Information Age.” In *The Future of the Book*, edited by Geoffrey Nunberg. Berkeley, CA: University of California Press.

Pande, Rohini. 2011. “Can Informed Voters Enforce Better Governance? Experiments in Low-Income Democracies.” *Annual Review of Economics* 3 (1): 215–37.

Pattanayak, Subhrendu K., Jui-Chen Yang, Dale Whittington, and K. C. Bal Kumar. 2005. “Coping with Unreliable Public Water Supplies: Averting Expenditures by Households in Kathmandu, Nepal: COPING WITH UNRELIABLE PUBLIC WATER SUPPLIES.” *Water Resources Research* 41 (2): n/a-n/a.

Portillo, Shannon, and Danielle S. Rudes. 2014. “Construction of Justice at the Street Level.” *Annual Review of Law and Social Science* 10 (1): 321–34.

Rammelt, Crelis F., and Maggi W. H. Leung. 2017. “Tracing the Causal Loops Through Local Perceptions of Rural Road Impacts in Ethiopia.” *World Development* 95 (July): 1–14. doi:10.1016/j.worlddev.2017.02.024.

Ranganathan, Malini. 2014. “‘Mafias’ in the Waterscape: Urban Informality and Everyday Public Authority in Bangalore.” *Water Alternatives* 7 (1). http://search.proquest.com/openview/8bd2275e590fe373749d8d00c2d0ee3e/1?pq-origsite=gscholar.

Srinivasan, Janaki, Megan Finn, and Morgan Ames. 2017. “Information Determinism: The Consequences of Faith in Information.” *The Information Society* 33 (1).

Stevenson, Edward G.J., Leslie E. Greene, Kenneth C. Maes, Argaw Ambelu, Yihenew Alemu Tesfaye, Richard Rheingans, and Craig Hadley. 2012. “Water Insecurity in 3 Dimensions: An Anthropological Perspective on Water and Women’s Psychosocial Distress in Ethiopia.” *Social Science & Medicine* 75 (2): 392–400.

Stokes, Susan C., Thad Dunning, Marcelo Nazareno, and Valeria Brusco. 2013. *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. New York, NY: Cambridge University Press.

Subbaraman, Ramnath, Laura Nolan, Kiran Sawant, Shrutika Shitole, Tejal Shitole, Mahesh Nanarkar, Anita Patil-Deshmukh, and David E. Bloom. 2015. “Multidimensional Measurement of

Household Water Poverty in a Mumbai Slum: Looking Beyond Water Quality." *PLOS ONE* 10 (7): e0133241.

Swyngedouw, E. 1997. "Power, Nature, Adn the City. The Consequent of Water and the Political Ecology of Urbanization in Guayaquil, Ecuador: 1880-1990." *Environment and Planning A* 29 (2): 311–32.

Tolbert, Caroline J., and Karen Mossberger. 2006. "The Effects of E-Government on Trust and Confidence in Government." *Public Administration Review* 66 (3): 354–69.

World Bank. 2004. "World Development Report 2004: Making Services Work for Poor People." Washington D.C.: World Bank.

Wutich, Amber, and Kathleen Ragsdale. 2008. "Water Insecurity and Emotional Distress: Coping with Supply, Access, and Seasonal Variability of Water in a Bolivian Squatter Settlement." *Social Science & Medicine* 67 (12): 2116–25.

Zérah, Marie-Hélène. 2000. "Household Strategies for Coping with Unreliable Water Supplies: The Case of Delhi." *Habitat International* 24 (3): 295–307.

TABLES

Table 1. Conditions at Baseline

Outcome	Overall population	Low income blocks	Target group
Household conditions			
Time spent waiting for water (*hrs. per supply day*)	0.92	1.26	1.30
Missing community events (*fraction of respondents*)	0.20	0.25	0.26
Missing work (*hrs. missed last 6 months*)	2.36	3.95	2.20
Need for substitutes (*fraction of respondents unable to store enough on supply days*)	0.25	0.25	0.23
Psychological conditions			
Worrying about water *(ranges from 1= often, to 4=not at all)*	2.43	2.38	2.25
Thinking about water during the day *(ranges from 1= often, to 4=not at all)*	2.46	2.29	2.23
Political attitudes			
Perception that providers are competent *(ranges from 1= agree to 3=disagree)*	1.38	1.40	1.29
Perceptions that providers are innovative and modern (*same as above)*	1.38	1.41	1.32
Perception that providers care about "people like us" (*same as above)*	1.48	1.50	1.44
Contacting			
Contacting providers directly about problems with service *(fraction of respondents contacting utility rather than others)*	0.07	0.05	0.05
Holding state water providers directly responsible for service *(fraction of respondents naming utility rather than others)*	0.15	0.08	0.11

Table 2. Intent-to-Treat Estimates (with covariate adjustment)[1]

Outcome[1]	Overall Population[2]			Low Income Group[3]			Target population[4]		
	Control mean[5]	ATE	p[6]	Control mean	ATE	p	Control mean	ATE	p
Household welfare effects									
Time spent waiting for water	0.51	-0.04	0.54 [0.94]	0.79	-0.05	0.71 [0.94]	0.71	0.01	0.91 [0.94]
Missing community events	0.13	0.00	0.89 [0.94]	0.16	0.03	0.46 [0.94]	0.14	0.00	0.91 [0.94]
Hours of work missed	2.28	-0.71	0.33 [0.94]	3.00	-0.05	0.94 [0.94]	3.30	0.15	0.90 [0.94]
Need for substitutes[6]	0.16	-0.02	0.18 [0.94]	0.23	-0.03	0.25 [0.94]	0.22	-0.07	0.03 [0.34]
Psychological effects[7]									
Worrying about water	2.63	0.11	0.04 [0.08]	2.51	0.12	0.07 [0.10]	2.53	0.05	0.31 [0.31]
Thinking about water during the day	2.77	0.10	0.04 [0.08]	2.52	0.22	0.00 [.02]	2.60	0.07	0.23 [0.27]
Political effects									
Perception that providers are competent	1.35	-0.05	0.16 [0.99]	1.34	-0.08	0.25 [0.99]	1.33	-0.04	0.54 [0.99]
Perceptions that providers are innovative and modern	1.50	-0.02	0.61 [0.99]	1.47	0.00	0.97 [0.99]	1.48	-0.02	0.75 [0.99]
Perception that providers care about "people like us"	1.66	0.00	0.99 [0.99]	1.63	0.01	0.93 [0.99]	1.66	-0.05	0.41 [0.99]
Contacting									
Contacting providers directly about problems with service	0.09	0.00	0.84 [0.91]	0.05	-0.03	0.08 [0.23]	0.02	0.03	0.11 [0.23]
Holding state water utility directly responsible for service	0.22	0.00	0.91 [0.91]	0.16	-0.06	0.02 [0.12]	0.13	0.02	0.55 [0.83]
N		2440			848			642	
N Treated		1227			426			336	

Note: [1] Unless noted otherwise hypothesis testing has been conducted using two tailed tests. [2] The covariates included are a block indicator, baseline outcome variable, an indicator for whether or not a household is low income, whether household receives Kaveri water supply, whether a household receives water supply every 2-4 days, whether a household receives supply without regularity (everyday supply is the omitted category), whether or not the household has an overhead tank/sump, and with the exception of row 1, the time reported waiting for water in wave 1. [3] Covariates included are the same as those included for the overall population. [4] Covariates included are the same as those included for the overall population, with the exception of whether the household has a tank and whether the household is low income. [5] Mean for control group in wave 2 of survey. [6] Calculated using Fisher exact tests. P-values adjusted using Benjamini-Hochberg adjustments for multiple testing are in brackets. [7] Hypothesis testing based on one-tailed tests.

Table 3. CACE for Households Receiving Accurate Notifications (with covariate adjustment)

Outcome[1]	Overall Population[2]			Low Income Group[3]			Target population[4]		
	Control mean[5]	CACE	SE[6] [p][7]	Control mean	CACE	SE [p]	Control mean	CACE	SE [p]
Household welfare effects									
Time spent waiting for water	0.49	-0.03	0.21 [0.90]	0.75	-0.09	0.70 [0.90]	0.67	0.61	0.95 [0.90]
Missing community events	0.13	-0.01	0.06 [0.90]	0.16	0.14	0.19 [0.90]	0.14	-0.03	0.19 [0.90]
Hours of work missed	2.01	-1.68	2.28 [0.90]	2.75	1.73	5.41 [0.90]	2.68	2.54	9.41 [0.90]
Need for substitutes[8]	0.15	-0.06	0.07 [0.80]	0.23	-0.18	0.22 [0.80]	0.22	-0.60	0.24 [0.07]
Psychological effects[8]									
Worrying about water	2.64	0.40	0.21 [0.07]	2.51	0.65	0.40 [0.08]	2.54	-0.41	0.64 [0.31]
Thinking about water during the day	2.78	0.37	0.20 [0.07]	2.54	1.16	0.43 [0.02]	2.61	-0.23	0.58 [0.34]
Political effects									
Perception that providers are competent	1.35	-0.18	0.12 [0.41]	1.34	-0.45	0.30 [0.41]	1.32	-0.25	0.33 [0.90]
Perceptions that providers are innovative and modern	1.50	-0.06	0.12 [0.90]	1.47	-0.02	0.34 [0.95]	1.48	-0.22	0.36 [0.90]
Perception that providers care about "people like us"	1.66	0.02	0.15 [0.95]	1.63	0.07	0.39 [0.95]	1.66	-0.58	0.35 [0.41]
Contacting									
Contacting providers directly about problems with service	0.09	0.01	0.04 [0.97]	0.05	-0.16	0.08 [0.10]	0.02	0.28	0.14 [0.10]
Holding state water utility directly responsible for service	0.22	0.00	0.08 [0.97]	0.16	-0.27	0.12 [0.10]	0.13	0.20	0.16 [0.29]
N[9]		2364			811			612	
N Treated		1193			403			319	
N compliers		289			71			45	

Note: [1] Unless noted otherwise hypothesis testing has been conducted using two tailed tests. [2] The covariates included are a block indicator, baseline outcome variable, an indicator for whether or not a household is low income, whether household receives Kaveri water supply, whether a household receives water supply every 2-4 days, whether a household receives supply without regularity (everyday supply is the omitted category), whether or not the household has an overhead tank/sump, and with the exception of row 1, the time reported waiting for water in wave 1. [3] Covariates included are the same as those included for the overall population. [4] Covariates included are the same as those included for the overall population, with the exception of whether the household has a tank and whether the household is low income. [5] Mean for control group in wave 2 of survey. [6] Standard errors clustered at the cluster level. [7] P-values adjusted using Benjamini-Hochberg multiple testing corrections shown in brackets. [8] Hypothesis testing based on one-tailed tests. [9] Only those units in both treatment and control groups that agreed to sign up for NextDrop's services have been included.

FIGURES

Figure 1. Example valve areas in BWSSB Subdivision E3

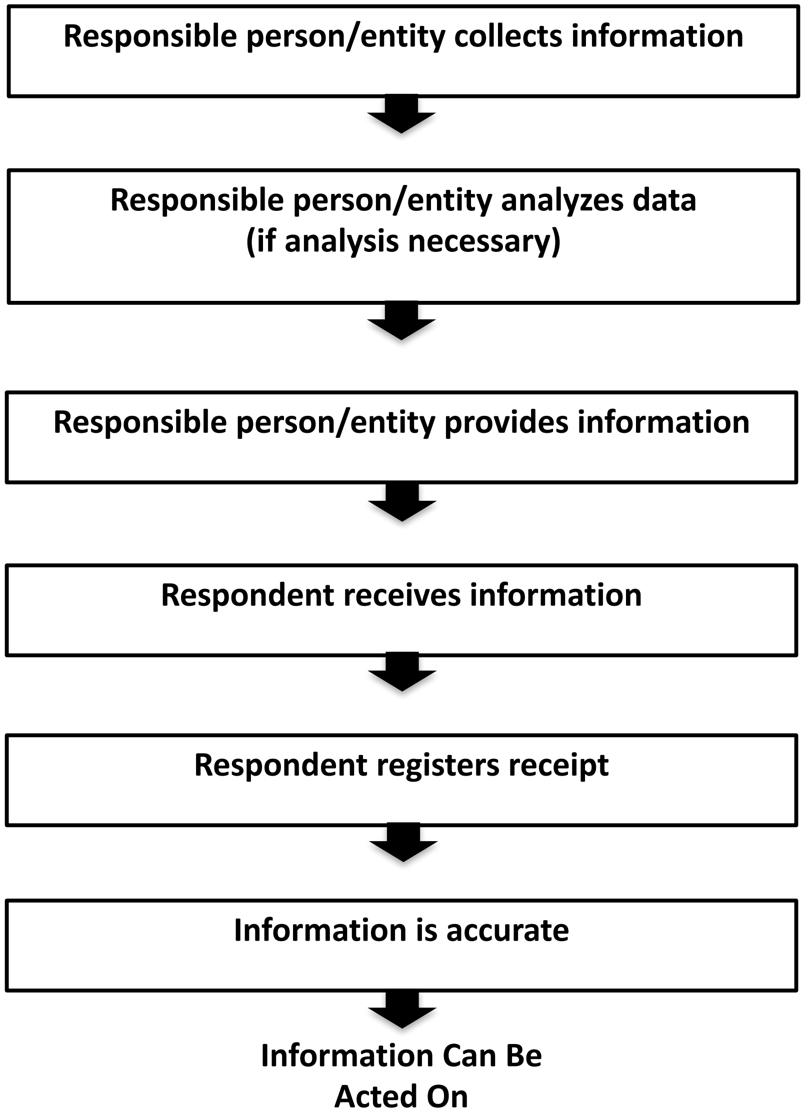


Figure 2. Information pipelines for transparency interventions

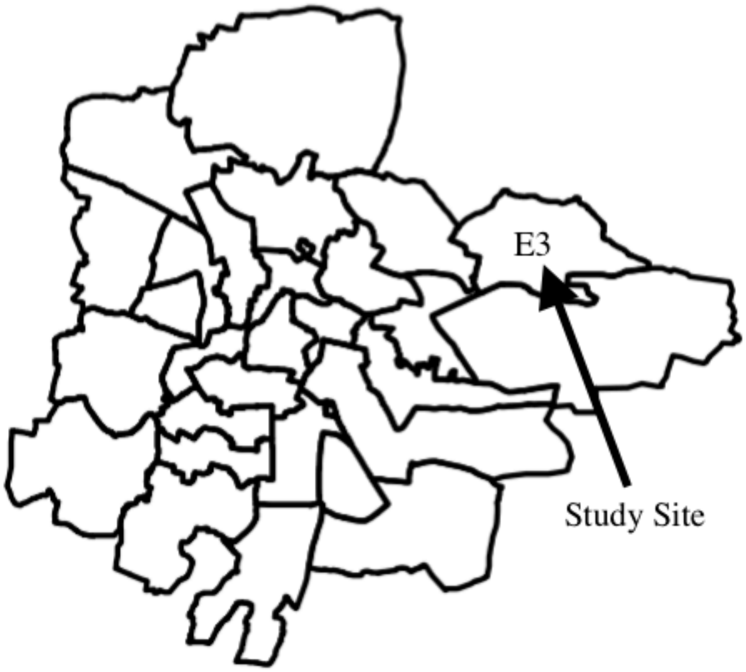


Figure 3. BWSSB Subdivision E3

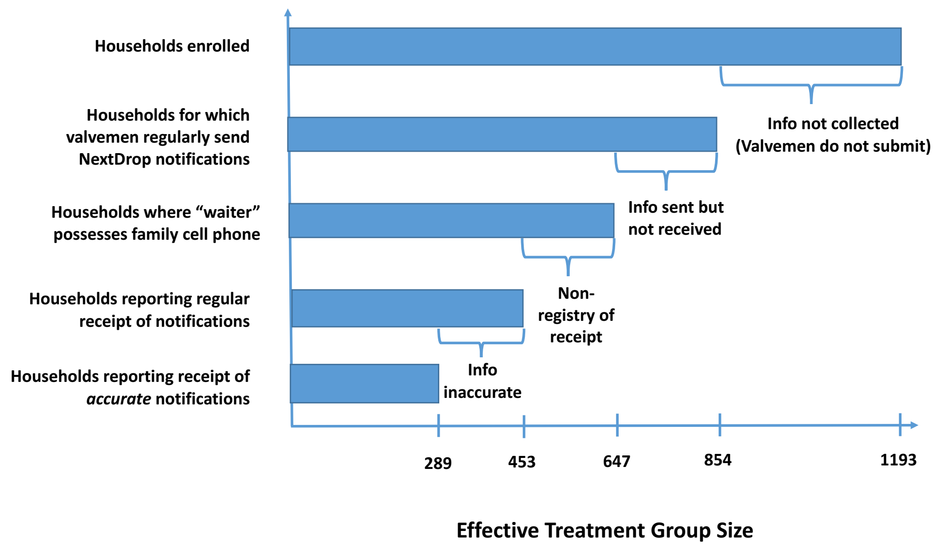
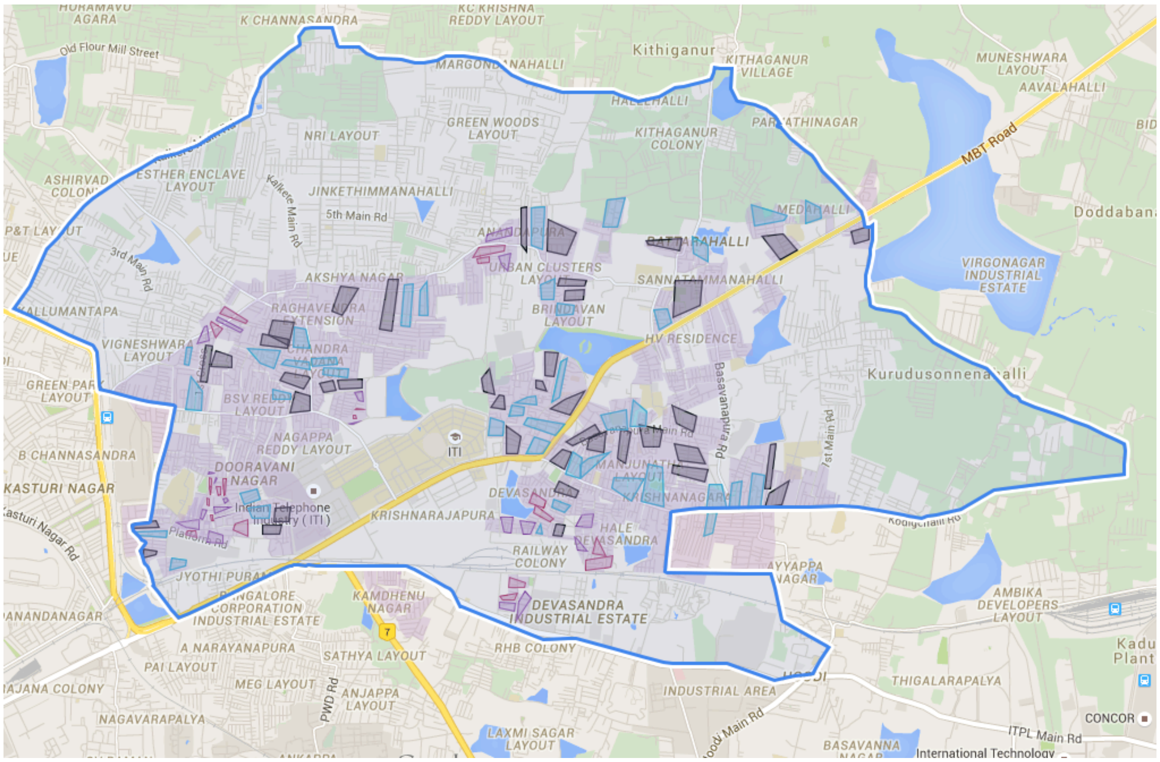


Figure 4. NextDrop’s leaky information pipeline and treatment group attrition

APPENDIX

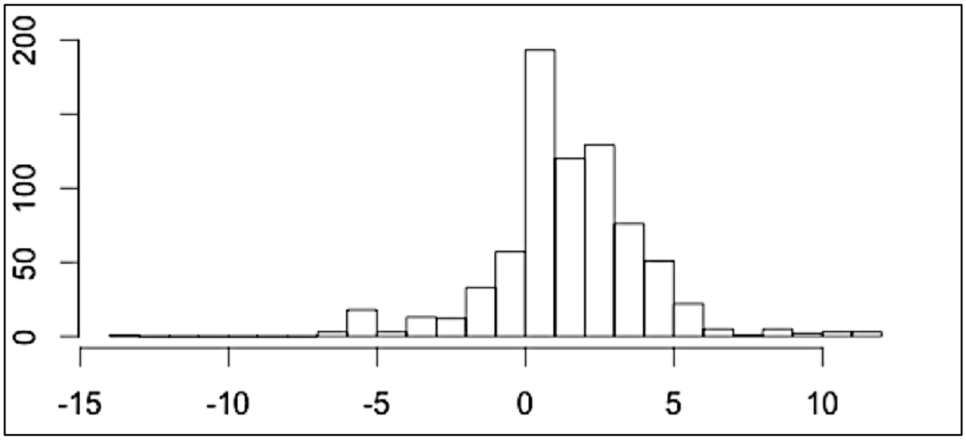
Section 1: Supplementary Figures and Tables

Figure A.1 Low and Mixed Income Clusters within Study Area



Note: The BWSSB E3 subdivision boundary is shown in blue, while areas receiving piped water supply are denoted in lavender. Pink and purple polygons denote low-income clusters (treatment and control); black and blue polygons denote mixed income clusters (treatment and control). There are four clusters per block.

Figure A.2. Histogram of discrepancies between household reports of water arrival times and valveman report data



Note: Calculated for the 33% percentage of households who were able to name a specific time when their water typically arrived, and residing in valve areas where valvemen had issued reports in the month prior to their survey interview. Unit on the X-axis is the difference in hours between a household report and valveman report regarding water arrival time. A negative number indicates that valveman reported time is prior to household reported time, while a positive number indicates that household reported time is prior to that of the valveman. A "2", for example, indicates that a valveman-sent report arrived two hours after a household reported water arrival, while a "0" indicates that both sources report the same water arrival time. Y-axis denotes frequency of reports.

Table A.1. Variable Definitions and Characteristics

Variable	Definition	Baseline mean	Max	Min	sd.
Dependent variables					
Time spent waiting for water[1]	Hours spent waiting per supply	0.92	12.00	0.00	1.70
Missing community events	Indicator for whether the household has missed community events in the past six months	0.20	1.00	0.00	0.40
Hours of work missed	Number of hours missed in the past six months	2.36	480.00	0.00	13.53
Need for substitutes	Indicator whether the household has needed substitutes for water in the past six months	0.25	1.00	0.00	0.44
Worrying about water	1-4 scale, with 1 indicating that the respondent worries about water arrival often, and 4 indicating that the respondent never worries about water arrival	2.43	4.00	1.00	0.98
Thinking about water during the day	1-4 scale, same as above	2.46	4.00	1.00	0.96
Perception that providers are competent	1-3 scale, where 1=I agree, 2=Don't know, 3=Disagree	1.37	3.00	1.00	0.60
Perception that providers are innovative and modern	1-3 scale, same as above	1.38	3.00	1.00	0.59
Perception that providers care about "people like us"	1-3 scale, same as above	1.48	3.00	1.00	0.65
Contacting providers directly about problems with service	1 if response to the question "When your water does not arrive at the right time or on the right day, whom, if anyone, do you contact?" is BWSSB, 0 otherwise	0.07	1.00	0.00	0.25
Holding state water utility directly responsible for service	1 if response to the question "Who do you feel is responsible for making sure that the piped water supply comes on the right day and time?" is BWSSB, 0 otherwise	0.15	1.00	0.00	0.36
Covariates					
Low-income	Indicator for whether or not a household earns up to Rs. 10,000 (150-200 USD)	0.33	1.00	0.00	0.47
Kaveri water supply	Indicator for whether or not the household receives Kaveri water supply	0.90	1.00	0.00	0.30
Water supplied everyday	Indicator for whether the household receives water supply everyday	0.02	1.00	0.00	0.15
Water received every 2-4 days	Indicator for whether the household receives water supply every 2-4 days	0.64	1.00	0.00	0.48
No regularity to water supply	Indicator for whether a household receives water supply with no regularity	0.67	1.00	0.00	0.47
House possesses overhead tank and/or sump	Indicator for whether or not the house possesses an overhead tank and/or sump	0.72	1.00	0.00	0.45
Tank fills automatically	Indicator for whether or not tanks fill automatically (applies only to houses with tanks)	0.40	1.00	0.00	0.49

Note: [1] Responses of waiting for 24 hours or more have been omitted from the sample. Results are robust to their inclusion and available upon request.

Table A.2. Baseline Covariate Balance[1]

Variable	Control Mean	Treatment Mean	Randomization inference p-value[2]	KS-test p-value
Household characteristics				
Altitude	908.94	907.00	0.44	0.45
Household size	4.48	4.61	0.13	0.32
State of origin is Karnataka	0.86	0.86	0.83	1.00
Muslim or SC/ST	0.27	0.29	0.23	0.97
In bottom two income categories	0.32	0.35	0.05	0.45
Have load bearing roof, motorized vehicle, and fridge	0.58	0.55	0.11	0.73
Water supply characteristics				
Have piped Kaveri water	0.88	0.92	0.00	0.27
Have overhead tank and/or sump	0.72	0.72	0.94	1.00
Water supplied everyday	0.03	0.02	0.57	1.00
Water supplied every 2-4 days	0.67	0.62	0.00	0.18
No regularity to water supply	0.69	0.64	0.00	0.13
Hours spent waiting for water	0.52	0.47	0.37	1.00
Political Characteristics				
Number of local leaders	0.74	0.77	0.68	0.77

Note: [1] N=2364. Balance for full sample, accounting for attrition between waves 1 and 2. Results are similar for balance for all households from wave 1. [2] Calculates Fisher exact tests. Based on 10,000 randomizations and assumes clustered nature of data.

Table A.3. Intent-to-Treat Estimates (without covariate adjustment)

Outcome[1]	Overall Population			Low Income Group			Target population		
	Control mean[2]	ATE	p[3]	Control mean	ATE	p	Control mean	ATE	p
Household welfare effects									
Time spent waiting for water	0.51	-0.04	0.51 [0.99]	0.79	-0.03	0.82 [0.99]	0.71	0.00	0.99 [0.99]
Missing community events	0.13	0.00	0.96 [0.99]	0.16	0.02	0.51 [0.99]	0.14	0.01	0.81 [0.99]
Hours of work missed	2.28	-0.62	0.36 [0.99]	3.00	-0.03	0.95 [0.99]	3.30	0.02	0.95 [0.99]
Need for substitutes[4]	0.16	-0.02	0.15 [0.91]	0.23	-0.03	0.28 [0.99]	0.22	-0.08	0.02 [0.19]
Psychological effects[4]									
Worrying about water	2.63	0.07	0.15 [0.31]	2.51	0.07	0.26 [0.39]	2.53	0.00	0.48 [0.48]
Thinking about water during the day	2.77	0.07	0.12 [0.31]	2.52	0.17	0.05 [0.28]	2.60	0.04	0.35 [0.42]
Political effects									
Perception that providers are competent	1.35	-0.04	0.23 [0.95]	1.34	-0.06	0.36 [0.95]	1.33	-0.01	0.90 [0.95]
Perceptions that providers are innovative and modern	1.50	-0.01	0.68 [0.95]	1.47	0.02	0.79 [0.95]	1.48	0.00	0.95 [0.95]
Perception that providers care about "people like us"	1.66	0.01	0.86 [0.95]	1.63	0.02	0.85 [0.95]	1.66	-0.04	0.54 [0.95]
Contacting									
Contacting providers directly about problems with service	0.09	0.00	0.98 [0.98]	0.05	-0.03	0.02 [0.08]	0.02	0.03	0.09 [0.18]
Holding state water utility directly responsible for service	0.22	0.00	0.96 [0.98]	0.16	-0.06	0.03 [0.08]	0.13	0.02	0.46 [0.69]
N		2440			848			642	
N Treated		1227			426			336	

Note: [1] Unless noted otherwise hypothesis testing has been conducted using two tailed tests. [2] Mean for control group in wave 2 of survey. [3] Calculated using Fisher exact tests. P-values adjusted using Benjamini-Hochberg adjustments for multiple testing are in brackets. [4] Hypothesis testing based on one-tailed tests.

Table A. 4. CACE for All Households Receiving Notifications (no covariates)

Outcome[1]	Overall Population			Low Income Group			Target population		
	Control mean[2]	CACE	SE[3] [p][4]	Control mean	CACE	SE [p]	Control mean	CACE	SE [p]
Household welfare effects									
Time spent waiting for water	0.49	-0.50	0.35 [0.58]	0.75	-1.23	0.95 [0.58]	0.67	-0.53	0.91 [0.59]
Missing community events	0.13	-0.11	0.10 [0.58]	0.16	-0.22	0.23 [0.58]	0.14	-0.12	0.16 [0.59]
Hours of work missed	2.01	-2.81	2.73 [0.58]	2.75	-4.83	8.48 [0.59]	2.68	-2.63	4.82 [0.59]
Need for substitutes[5]	0.15	0.04	0.07 [0.58]	0.23	0.00	0.21 [0.59]	0.22	-0.14	0.18 [0.58]
Psychological effects[5]									
Worrying about water	2.64	0.32	0.23 [0.16]	2.51	0.25	0.39 [0.26]	2.54	-0.45	0.44 [0.18]
Thinking about water during the day	2.78	0.27	0.22 [0.16]	2.54	0.52	0.39 [0.16]	2.61	-0.54	0.42 [0.16]
Political effects									
Perception that providers are competent	1.35	-0.15	0.15 [0.88]	1.34	-0.49	0.39 [0.88]	1.32	-0.32	0.29 [0.88]
Perceptions that providers are innovative and modern	1.50	-0.05	0.14 [0.90]	1.47	-0.24	0.39 [0.90]	1.48	-0.23	0.33 [0.90]
Perception that providers care about "people like us"	1.66	0.02	0.15 [0.90]	1.63	-0.07	0.44 [0.90]	1.66	-0.05	0.38 [0.90]
Contacting									
Contacting providers directly about problems with service	0.09	-0.05	0.05 [0.46]	0.05	-0.17	0.11 [0.46]	0.02	0.08	0.12 [0.53]
Holding state water utility directly responsible for service	0.22	0.09	0.10 [0.46]	0.16	-0.15	0.15 [0.46]	0.13	0.15	0.17 [0.46]
N[6]		2364			811			612	
N Treated		1193			403			319	
N compliers		453			119			84	

Note: [1] Unless noted otherwise hypothesis testing has been conducted using two tailed tests. [2] Mean for control group in wave 2 of survey. [3] Standard errors clustered at the cluster level. [4] P-values adjusted using Benjamini-Hochberg multiple testing corrections shown in brackets. [5] Hypothesis testing based on one-tailed tests. [6] Only those units in both treatment and control groups that agreed to sign up for NextDrop's services have been included.

Table A.5. CACE for Households Receiving Notifications (with covariate adjustment)[1]

Outcome[1]	Overall Population[2]			Low Income Group[3]			Target population[4]		
	Control mean[5]	CACE	SE[6] [p][7]	Control mean	CACE	SE [p]	Control mean	CACE	SE [p]
Household welfare effects									
Time spent waiting for water	0.49	-0.02	0.14 [0.90]	0.75	-0.05	0.42 [0.90]	0.67	0.32	0.49 [0.90]
Missing community events	0.13	-0.01	0.04 [0.90]	0.16	0.08	0.11 [0.90]	0.14	-0.01	0.10 [0.90]
Hours of work missed	2.01	-1.07	1.44 [0.90]	2.75	1.02	3.17 [0.90]	2.68	1.30	4.64 [0.90]
Need for substitutes[8]	0.15	-0.04	0.05 [0.81]	0.23	-0.11	0.13 [0.81]	0.22	-0.31	0.12 [0.05]
Psychological effects[8]									
Worrying about water	2.64	0.25	0.13 [0.07]	2.51	0.38	0.23 [0.07]	2.54	-0.21	0.33 [0.31]
Thinking about water during the day	2.78	0.24	0.13 [0.07]	2.54	0.68	0.25 [0.02]	2.61	-0.12	0.30 [0.34]
Political effects									
Perception that providers are competent	1.35	-0.12	0.07 [0.39]	1.34	-0.27	0.18 [0.39]	1.32	-0.13	0.18 [0.90]
Perceptions that providers are innovative and modern	1.50	-0.04	0.08 [0.90]	1.47	-0.01	0.20 [0.95]	1.48	-0.12	0.19 [0.90]
Perception that providers care about "people like us"	1.66	0.01	0.10 [0.95]	1.63	0.04	0.23 [0.95]	1.66	-0.30	0.19 [0.39]
Contacting									
Contacting providers directly about problems with service	0.09	0.00	0.03 [0.97]	0.05	-0.09	0.05 [0.09]	0.02	0.15	0.07 [0.09]
Holding state water utility directly responsible for service	0.22	0.00	0.05 [0.97]	0.16	-0.16	0.07 [0.09]	0.13	0.11	0.08 [0.28]
N[9]		2364			811			612	
N Treated		1193			403			319	
N compliers		453			119			84	

Note: [1] Unless noted otherwise hypothesis testing has been conducted using two tailed tests. [2] The covariates included are a block indicator, baseline outcome variable, an indicator for whether or not a household is low income, whether household receives Kaveri water supply, whether a household receives water supply every 2-4 days, whether a household receives supply without regularity (everyday supply is the omitted category), whether or not the household has an overhead tank/sump, and with the exception of row 1, the time reported waiting for water in wave 1. [3] Covariates included are the same as those included for the overall population. [4] Covariates included are the same as those included for the overall population, with the exception of whether the household has a tank and whether the household is low income. [5] Mean for control group in wave 2 of survey. [6] Standard errors clustered at the cluster level. [7] P-values adjusted using Benjamini-Hochberg multiple testing corrections shown in brackets. [8] Hypothesis testing based on one-tailed tests. [9] Only those units in both treatment and control groups that agreed to sign up for NextDrop's services have been included.

Table A.6. CACE for Households Receiving Accurate Notifications (without covariate adjustment)[1]

Outcome[1]	Overall Population			Low Income Group			Target population		
	Control mean[2]	CACE	SE[3] [p][4]	Control mean	CACE	SE [p]	Control mean	CACE	SE [p]
Household welfare effects									
Time spent waiting for water	0.49	-0.78	0.55 [0.57]	0.75	-2.06	1.58 [0.57]	0.67	-1.04	1.77 [0.59]
Missing community events	0.13	-0.17	0.15 [0.57]	0.16	-0.36	0.37 [0.57]	0.14	-0.23	0.32 [0.59]
Hours of work missed	2.01	-4.41	4.28 [0.57]	2.75	-8.13	14.27 [0.59]	2.68	-5.12	9.39 [0.59]
Need for substitutes[5]	0.15	0.06	0.11 [0.57]	0.23	0.00	0.36 [0.59]	0.22	-0.28	0.34 [0.57]
Psychological effects[5]									
Worrying about water	2.64	0.50	0.35 [0.16]	2.51	0.42	0.67 [0.26]	2.54	-0.87	0.84 [0.18]
Thinking about water during the day	2.78	0.43	0.34 [0.16]	2.54	0.87	0.67 [0.16]	2.61	-1.03	0.78 [0.16]
Political effects									
Perception that providers are competent	1.35	-0.24	0.23 [0.88]	1.34	-0.82	0.65 [0.88]	1.32	-0.62	0.56 [0.88]
Perceptions that providers are innovative and modern	1.50	-0.08	0.23 [0.90]	1.47	-0.40	0.66 [0.90]	1.48	-0.44	0.64 [0.90]
Perception that providers care about "people like us"	1.66	0.03	0.24 [0.90]	1.63	-0.12	0.74 [0.90]	1.66	-0.09	0.72 [0.90]
Contacting									
Contacting providers directly about problems with service	0.09	-0.08	0.08 [0.46]	0.05	-0.28	0.19 [0.46]	0.02	0.15	0.23 [0.52]
Holding state water utility directly responsible for service	0.22	0.14	0.16 [0.46]	0.16	-0.25	0.24 [0.46]	0.13	0.28	0.31 [0.46]
N[6]		2364			811			612	
N Treated		1193			403			312	
N compliers		289			71			45	

Note: [1] Unless noted otherwise hypothesis testing has been conducted using two tailed tests. [2] Mean for control group in wave 2 of survey. [3] Standard errors clustered at the cluster level. [4] P-values adjusted using Benjamini-Hochberg multiple testing corrections shown in brackets. [5] Hypothesis testing based on one-tailed tests. [6] Only those units in both treatment and control groups that agreed to sign up for NextDrop's services have been included**.**

Table A.7. Average Treatment Effects for Valve Areas Receiving Accurate Notifications (Defined Based on Water Arrival Time)

Outcome[1]	Overall Population			Low Income Group			Target population		
	Control mean[2]	ATE	p[3]	Control mean	ATE	p	Control mean	ATE	p
Household welfare effects									
Time spent waiting for water	0.64	-0.11	0.34 [0.56]	0.87	0.04	0.83 [0.91]	0.87	0.09	0.70 [0.84]
Missing community events	0.12	-0.03	0.25 [0.56]	0.14	0.00	0.93 [0.93]	0.16	-0.07	0.26 [0.56]
Hours of work missed	1.93	0.30	0.68 [0.84]	3.10	1.36	0.26 [0.56]	4.19	2.02	0.35 [0.56]
Need for substitutes[4]	0.16	-0.05	0.06 [0.36]	0.22	-0.02	0.38 [0.56]	0.25	-0.13	0.01 [0.16]
Psychological effects[4]									
Worrying about water	2.63	0.14	0.15 [0.45]	2.45	0.00	0.47 [0.47]	2.45	-0.05	0.40 [0.47]
Thinking about water during the day	2.69	0.16	0.09 [0.45]	2.50	0.07	0.31 [0.47]	2.47	-0.03	0.43 [0.47]
Political effects									
Perception that providers are competent	1.32	-0.07	0.34 [0.99]	1.32	0.00	0.99 [0.99]	1.35	-0.04	0.73 [0.99]
Perceptions that providers are innovative and modern	1.47	-0.05	0.56 [0.99]	1.48	0.02	0.88 [0.99]	1.56	-0.02	0.89 [0.99]
Perception that providers care about "people like us"	1.66	-0.05	0.70 [0.99]	1.60	0.10	0.66 [0.99]	1.73	-0.06	0.65 [0.99]
Contacting									
Contacting providers directly about problems with service	0.07	0.04	0.11 [0.68]	0.03	0.00	0.95 [0.95]	0.04	0.03	0.37 [0.69]
Holding state water utility directly responsible for service	0.16	0.04	0.46 [0.69]	0.16	-0.04	0.56 [0.69]	0.10	0.03	0.58 [0.69]
N		646			352			210	
N Treated		321			151			96	

Note: Calculated for the 40% of our sample that said their water came at a specific time, and was able to provide a time. (Effects thus look somewhat different than for the overall population.) Accuracy based on modal response (within a valve area) to "At what hour does your water usually start?" Valve areas are considered "accurate" if the last valveman report was between 2 hours before and 1 hour after this time. [1] Unless noted otherwise hypothesis testing has been conducted using two tailed tests. [2] Mean for control group in wave 2 of survey. [3] Calculated using Fisher exact tests. P-values adjusted using Benjamini-Hochberg adjustments for multiple testing are in brackets. [4] Hypothesis testing based on one-tailed tests.

Table A.8. Average Treatment Effects for Valve Areas Receiving Accurate Notifications (Defined Based on Day of Water Supply)

Outcome[1]	Overall Population			Low Income Group			Target population		
	Control mean[2]	ATE	p[3]	Control mean	ATE	p	Control mean	ATE	p
Household welfare effects									
Time spent waiting for water	0.53	0.01	0.97 [0.97]	0.75	0.09	0.46 [0.82]	0.64	0.27	0.07 [0.37]
Missing community events	0.13	-0.01	0.79 [0.97]	0.15	0.03	0.54 [0.82]	0.16	-0.03	0.48 [0.82]
Hours of work missed	3.13	-1.62	0.14 [0.37]	3.07	0.14	0.87 [0.97]	3.68	0.06	0.95 [0.97]
Need for substitutes[4]	0.18	-0.03	0.15 [0.37]	0.27	-0.07	0.15 [0.37]	0.26	-0.11	0.02 [0.23]
Psychological effects[4]									
Worrying about water	2.59	0.08	0.20 [0.32]	2.45	0.06	0.32 [0.32]	2.48	-0.11	0.17 [0.32]
Thinking about water during the day	2.72	0.05	0.29 [0.32]	2.46	0.17	0.07 [0.32]	2.54	-0.10	0.22 [0.32]
Political effects									
Perception that providers are competent	1.35	-0.04	0.46 [0.88]	1.37	-0.08	0.41 [0.88]	1.36	-0.06	0.49 [0.88]
Perceptions that providers are innovative and modern	1.49	0.02	0.76 [0.88]	1.49	0.03	0.78 [0.88]	1.50	0.00	0.94 [0.94]
Perception that providers care about "people like us"	1.64	0.07	0.24 [0.88]	1.65	0.06	0.69 [0.88]	1.65	-0.04	0.66 [0.88]
Contacting									
Contacting providers directly about problems with service	0.07	0.01	0.58 [0.66]	0.03	-0.01	0.63 [0.66]	0.02	0.03	0.16 [0.66]
Holding state water utility directly responsible for service	0.20	0.02	0.66 [0.66]	0.14	-0.03	0.34 [0.66]	0.12	0.05	0.28 [0.66]
N		1411			536			390	
N Treated		688			272			180	

Note: Accuracy defined based on modal response (within a valve area) to "On which day does your water usually arrive?" Valve areas are considered "accurate" if the last valveman report was on this same day. [1] Unless noted otherwise hypothesis testing has been conducted using two tailed tests. [2] Mean for control group in wave 2 of survey. [3] Calculated using Fisher exact tests. P-values adjusted using Benjamini-Hochberg adjustments for multiple testing are in brackets. [4] Hypothesis testing based on one-tailed tests.

Table A.9. Rates of Compliance According to Different Definitions

	Predicted	HH enrolling in ND	HH enrolling in ND and receiving notifications	HH enrolling in ND and receiving accurate notifications[1]
N in treatment group (after attrition)	1227	1227	1227	1227
N compliers	981	1193	453	289
% compliers	80%	97%	37%	24%

Note: [1] Accuracy defined based on survey responses.

Table A.10, Power of experiment to detect CACE effects at 5% significance among the target group

Definition of compliance	Covariate adjustment	Power range[1]
Receipt of notifications	No	0.210 - 0.551
Receipt of notifications	Yes	0.193 - 0.531
Receipt of accurate notifications	No	0.125 - 0.232
Receipt of accurate notifications	Yes	0.103 - 0.259

Note: [1] For each definition of compliance, the simulations yield a range of power estimates that accommodate upper and lower bounds on true treatment effects, standard deviations, and intra-cluster correlation coefficients. As in the pre-analysis plan, we assume that the true treatment effect lies between a 30-minute (standard deviation of one hour) and two (standard deviation of two hours) hour reduction in wait time. The intra-cluster correlation coefficient is assumed to lie between .01 and .15. Results based on 1,000 simulations. Results are not corrected for multiple hypothesis testing—power simulations conducted with multiple testing adjustments are likely to yield even lower ranges for power.

Section 2: Pre-Analysis Plan

The Household Welfare and Political Impacts of Increasing Service Predictability: An Experimental Intervention in Bangalore's Water Sector

May 13, 2015

Abstract: Throughout the developing world, intermittency and unpredictability are the hallmarks of public service delivery: buses do not run on a standard schedule, water supplies are variable in terms of arrival times, and electricity blackouts occur unexpectedly. Surprisingly, the causes and consequences of unpredictable urban services have received much less attention than patterns of access to, or government expenditures on, these services. While addressing the underlying causes of service unpredictability tends to be very costly (e.g., replacing leaky water pipes or increasing capacity to improve water pressure levels), low-cost informational interventions can potentially help households to cope with service unpredictability. Alleviating coping costs may also change the way in which citizens relate to their local governments.

Through a cluster-randomized experiment in Bangalore, we will evaluate a text-message based notification scheme providing households with advance warning of the timing of water services and supply cancellations, and a dedicated contact number for reporting problems. We assess whether the notification system reduces: a) the time spent waiting for water; b) expenditures on substitutes for piped water services; and c) stress levels on account of uncertain and irregular deliveries and uncertainty. We hypothesize that lower- income households will see greater welfare impacts because they have few affordable and accessible alternative water sources. We also examine if, and how, the receipt of real-time information changes how citizens "see the state," whom they hold responsible for service quality and problems, and whom they approach about service concerns. We hypothesize that the advance notification system will prompt citizens to see the state as more modern and responsive when compared

with the status quo, in which the state can only be accessed through intermediaries and low level bureaucrats. This research thus develops a framework within which the material, stress-related, and political impacts of greater predictability in urban services can be assessed.

Introduction

Throughout the developing world, intermittency and unpredictability are the hallmarks of public service delivery: buses do not run on a standard schedule, water supplies are variable in terms of arrival times, and electricity blackouts occur unexpectedly. For instance, worldwide, 400 million people rely on intermittent water, often receiving services only a few days a week for a few hours (van den Berg and Danilenko 2011). The poor state of the underlying urban infrastructure—prone to unexpected pipe leaks and power outages that put pumps out of service—often means that services are not only intermittently delivered, but are also unpredictable. Yet the causes and consequences of unpredictable urban services have received much less attention than patterns of access to, or government expenditures on, these services.

Service unpredictability can be particularly onerous for low-income populations. Relying on buses that do not arrive on time or on a standard schedule can make it difficult to consistently arrive at work on time; lower income populations who cannot afford substitutes such as personal cars or private taxis are more likely to develop reputations for unreliability with employers under such circumstances.[33] Coping with electricity blackouts is also more difficult for poorer households, who often cannot afford private generators.[34] Similarly, low-income households receiving intermittent and unpredictable water services suffer in a number of ways: they must spend time waiting at home for water arrival so as to be able to fill household storage containers, as substitutes such as vended water tend to be much more

[33] E.g. Smith (2007).

[34] United Nations Development Programme. 2010. Human Development Report. New York: UNDP.

expensive than municipal water. Higher income households, in contrast, can afford pumps that automatically fill household tanks when water services commence, as well as the load-bearing roofs that such tanks require.

There are many reasons to believe that service unpredictability, in addition to imposing material and financial hardship, also weakens bonds between individuals the state. Citizens who cannot depend upon regular services, one might argue, will be less likely to view government service providers as competent and respectful of citizen concerns. If citizens must attempt to contact low-level bureaucrats or local political bosses on multiple occasions in order to obtain information regarding when services might resume, such feelings will only be magnified. Often, lay citizens do not know whom to contact at all. Under such circumstances, we would also expect citizens receiving unpredictable services to be far less willing to pay user fees--especially ones in line with delivery costs--than those receiving services with regularity.[35]

This project examines the effect of increasing service predictability on household welfare and on citizens' relationship with the local state. It does so through an impact analysis of an information-based intervention being piloted in India's water sector. Through a cluster-randomized experiment in Bangalore, we evaluate the impact of a service developed by a social enterprise called NextDrop, which provides households with text message notifications regarding water arrival times and supply cancellations. Water utilities in urban India typically do not possess sensors that allow them to monitor exactly where water is within their network, so NextDrop has developed a novel system in which the utility employees, who operate the valves allowing water to flow into areas of 50-200 households, notify them when they are opening and closing valves. NextDrop then sends notifications to individual households, which they have cataloged by "valve area" through the collection of household GPS coordinates, letting them know when their water will arrive.

[35] See, for instance, Savedoff and Spiller (1999) regarding the water and sanitation sector.

Our field experiment—which is taking place in the context of NextDrop's rollout in Bangalore—examines two sets of potential impacts. The study is designed to assess household level welfare impacts such as: a) the time spent waiting for water; b) expenditures on substitutes for piped water services; and c) stress levels on account of uncertain deliveries. The study will also examine how (and if) NextDrop delivery notifications impact how citizens "see the state," [36] whom they hold responsible for service quality and problems, and whom they approach about service concerns.

This paper proceeds as follows. First, we provide a more thorough description of the intervention to be evaluated. We then locate our study within the theoretical perspectives that inform our hypotheses. We proceed to outline the main hypotheses we will examine in the course of the study with respect to our outcomes of interest, including our expectations regarding heterogeneous treatment effects. Finally, we describe our research design and review the power calculations used to determine sample size, and the allocation of our sample between different populations of interest.

The Intervention Our Study Will Evaluate

We evaluate the impact of a particular intervention that may increase the predictability of urban services: text-message based notifications regarding the timing of water delivery in intermittent water systems. NextDrop, a start-up venture launched by recent UC Berkeley graduates (among others), has developed and initiated an SMS-based notification system to help consumers and small businesses reduce the coping costs of water intermittency. NextDrop's cell-phone based system is innovative in its use of employee- and crowd-sourced data collected through low-cost and ubiquitous cell phones. The notification system potentially represents a more reliable form of information to households regarding when water will be available than the published schedules in newspapers or on the walls of local service

[36] Corbridge *et al*. (2005).

stations, which our preliminary data collection in Bangalore suggests depart significantly from actual practice.

NextDrop obtains information regarding the timing of water delivery in particular "valve areas" of 50-200 households by collecting water flow information from valvemen—the individuals responsible for opening and closing the valves controlling water into particular districts—and disseminating notifications to NextDrop customers 30 minutes to an hour before the water turns on. The notifications are provided free of charge to households.[37]

NextDrop has already piloted their system in a second-tier Indian city, Hubli-Dharwad (population 1 million), in the southern Indian state of Karnataka. The enterprise has since refined its incentive schemes for valvemen (whose cooperation is essential for the accurate data collection),[38] refined the ability to notify customers via a geocoding system and created a 'dashboard' for the utility for real-time information about water flow and allocation within the system. It is currently rolling out services in the cities of Mysore and Bangalore.

Theoretical Discussion and Hypotheses

This project draws on two literatures that speak to the potential effects of improving the predictability of urban services—and NextDrop's notifications in particular—on: a) household welfare; and b) citizens' relationship with the state. Building on these literatures, we propose to examine eight related hypotheses regarding household welfare effects and six on political effects.

---

[37] NextDrop's revenue model involves charging utilities for information-based services, including real-time information of water flows and sending water arrival notifications to consumers.

[38] The accuracy of valvemen reports is monitored through a system of callbacks with consumers. We are evaluating the efficacy of alternative incentives schemes for the water utility's valvemen in a related project.

A. *Household Welfare Effects*

The water policy literatures suggest that intermittent water supply imposes significant costs on households. While some researchers have studied the effects of water service intermittency on water quality and human health (e.g. Kumpel and Nelson 2013; Ercumen *et al*. forthcoming), few have quantified or modeled the coping costs and inefficiencies of unreliable water deliveries.[39] This is the first study, to our knowledge, that will empirically estimate the impacts of greater predictability of intermittent water services. Our hypotheses thus build on a broader literature, including behavioral economics and development studies.

The first set of costs we examine relate to the time spent waiting for water services to commence. In low income households that cannot afford maids or automatically-filling storage tanks, household members—and particularly women—may need to restrict their activities to the household for substantial periods of time in order to ensure they are at home, and thus able to store water, whenever water services commence (e.g. Zérah 2000). The "waiter" thus devotes time to waiting for water that might otherwise be spent on work, community activities, religious functions, etc. Receiving notifications regarding the time of water delivery on a given supply day, or a notification of supply cancellation, it stands to reason, would reduce the amount of time spent waiting and allow more time for these other activities. This leads us to a first set of hypotheses:

H1: Household members in charge of managing a household's water supply will spend less time waiting for water on a weekly basis if they receive accurate prior notifications regarding delivery times and service disruptions.

[39] For exceptions, see Zérah (2000) or Baisa *et al*. (2010).

H2: *Household members in charge of managing water supply will be better able to participate in community, social, or religious activities if they receive accurate prior notifications regarding delivery times and service disruptions.*

H3: Household members in charge of managing water supply will be less likely to forego earnings if they receive accurate prior notifications regarding delivery times and service disruptions.

The literature on urban water supply also documents the extent to water obtained through a city's piped water network costs less than substitutes such as bottled water and bulk water from water vendors (e.g. Estache *et al*., 2001). Obtaining substitute sources, especially at short notice, requires time and energy spent fetching and carrying water from wells, or arranging for special deliveries with vendors, etc. (Kjellén and McGranahan 2006). In light of these studies, we would expect that notifications regarding the timing of water delivery would reduce the necessity to rely upon substitutes, because they would decrease the probability of missing a supply period. This leads us to a second set of hypotheses:

H4: Receiving accurate prior notifications regarding water delivery times and service disruptions will reduce household expenditures on substitutes, such as bottled, vended or tanker water.

H5: Receiving accurate prior notifications regarding water delivery times and service disruptions will reduce the effort households spend securing alternative sources of water.

Finally, waiting for water and supply unpredictability may impose psychological costs. Given that water is such a vital resource for households, and that substitutes for piped water can be much more expensive or less desirable, the household member responsible for obtaining and managing a household's water supply may incur a great deal of stress when services are unpredictable or water storage cannot be planned. This argument builds directly on empirical studies of water stress (e.g. Wutich and Ragsdale 2008), as well as the behavioral economics literature, which has shown that many dimensions of poverty impose cognitive and other stresses upon the poor (e.g. Mullainathan and Shafir 2013). Such stresses manifest themselves in a number of ways: consciously worrying about water provision, preoccupation (i.e. thinking about water delivery while doing other things), or stress incurred through the hassle of repeatedly trying to obtain information regarding supply timing from street level bureaucrats or local intermediaries. This leads us to a third set of hypotheses:

H6: Receiving accurate prior notifications regarding water delivery times and service disruptions will reduce respondents' sense that missing or delayed water supply is a constant worry.

H7: Receiving accurate prior notifications regarding water delivery times and service disruptions will reduce the extent to which respondents finding themselves thinking about water supply while doing other things, such as household chores or paid work.

H8: Receiving accurate prior notifications regarding water delivery times and service disruptions will reduce the effort households must expend to secure important information regarding the timing of water deliveries.

While we hypothesize that these effects may be observable across the entire urban population in developing country cities with water intermittency, we expect these effects to be particularly pronounced under certain circumstances:

- For *low-income households*, because the cost of substitutes for piped water as a fraction of household income is greater, and because poverty itself exacerbates stress (see Mullainathan and Shafir 2013). We will measure income not only in terms of monthly cash income, but also in terms of the types of assets that a household possesses. Income may correlate strongly with religion or scheduled caste/tribe status; if so, we are likely to observe that effects vary along these lines as well);

- For *households living in structures without automatically-filling overhead tanks*, which often cannot be supported in 1-2 level structures of poor construction quality; [40]
- For *households where someone spends a significant amount of time waiting or worrying about water* prior to the intervention. (We expect this to be closely related to the above conditions.)
- For *households receiving water services that typically arrive on a scheduled supply day within an interval of less than 4 hours*. Our intuition is that notifications are most useful when they occur within an interval of four hours or less, because individuals would be more likely to stay at home and wait for water under such circumstances. If arrival times are less predictable than this, household members are unlikely to stay at home waiting for the water to arrive.
- In *households where the person responsible for managing and storing water is of working age*, and especially where this person is a male of working age, receiving accuracy notifications from NextDrop may reduce the economic cost of water intermittency. On the other hand, stress levels for women "waiters" may be reduced more dramatically than those for male "waiters" because domestic water is traditionally a "female" responsibility. Therefore if the water does not arrive on time, or does not arrive at all, it is more likely to be the female head-of-household who will need to find alternative sources.

We thus expect to observe *heterogeneous treatment effects* when we subset our analysis according to these intuitions.[41]

[40] There may be overlap between low-income households and households living without automatically filling storage tanks.

B. Political Effects

In this project, we also ask whether--even in the absence of substantive service improvements, such as better water quality, or more frequent deliveries--better information alone lead to a more favorable view of the local state and its agencies. In doing so, we build on a broader literature investigating what determines how citizens "see" and relate to the state (Corbridge *et al*. 2005; Ferguson and Gupta 2002; Evans 2008). This question connects both the political science and public goods literatures to a rich body of research on the role of information, and information technologies, in development. Here it has been argued that better information, through information technologies, on government schemes, commodity prices, or water quality, directly influences citizens' views of the state (though not necessarily in a positive direction) (Madon and Sahay 2002; Tolbert and Mossberger 2006).

Our first intuition is that increasing the predictability of services will lead citizens to revise their judgments of state competence. Even if services are still delivered intermittently, and with less frequency than citizens may desire, receiving accurate, prior information regarding service timing (and cancellations) should not only make services easier to access, but also convey greater state capacity and control. In addition, the innovative application of text messaging to disseminate service schedule information to citizens could make the state seem more "modern," and thus further improve citizen perceptions of governmental competence (Harriss 2006; Kuriyan and Ray 2009; Ghertner 2011). This leads us to the following hypotheses:

H9: Receiving accurate prior notifications regarding water delivery times and service disruptions will improve citizen perceptions of the competence of state water providers.

---

[41] In some cases we may not have sufficient sample size for heterogeneous treatment effects to be visible or statistically significant.

H10: Receiving accurate prior notifications regarding water delivery times and service disruptions will lead citizens to perceive state water providers as more innovative and modern than previously.

We also expect notifications that increase service predictability to shift citizen perceptions regarding who is responsible for addressing their concerns regarding services, and the state's level of universalism. The literature on citizen-state interactions in the developing world suggests that citizens (and especially low and middle income populations) tend to turn to lower level bureaucrats or political intermediaries regarding service problems.[42] When such mechanisms do not work, a group of households and community leaders may collectively protest at government offices.[43] These patterns may slowly shift with the introduction of a universally administered notification system that connects citizens more directly to the urban service provider. Citizens are more likely to view government agencies themselves, rather than their local intermediaries, as responsible for addressing their problems, and redirect their complaints (or lodge additional complaints) with government bureaucracies.

This shift will be tied, at least in part, to an increasing sense that notifications are sent not just to favored groups, but also to marginalized populations. It will also stem from the fact that information is arriving automatically, without effort on the part of the citizen; it no longer requires cumbersome and time-consuming efforts to proactively solicit information—which may even involve reliance upon personal connections, favors, or bribes. To the extent that this perceptual shift occurs, we would expect that more people will lodge individual (rather than group-only) complaints, thinking that their complaints will be heard. These intuitions lead us to the following hypotheses:

---

[42] On the Bangalore case, see Ranganathan (2014); on the importance of intermediaries within clientelistic party systems in the developing world, see Stokes *et al*. (2013) for a review.

[43] See Ranganathan (2014) regarding water protests in Bangalore; see also Auerbach (2014) regarding urban services in Northern Indian cities.

H11: Receiving accurate prior notifications regarding water delivery times and service disruptions will lead citizens to be more likely to perceive state water providers as more universalistic service providers that care about "people like us."

H12: Receiving accurate prior notifications regarding water delivery times and service disruptions will lead citizens to view state water providers as directly responsible for correcting service problems than previously, when local political leaders and/or intermediaries might have been held responsible.

H13: Receiving accurate prior notifications regarding water delivery times and service disruptions will lead citizens to be more likely to contact the central water bureaucracy directly (through text or phone) than previously.

Building on the preceding arguments, it stands to reason that when water utility services become easier to utilize because they become more predictable, and as perceptions of state competence and universalism increase, citizens may be more disposed to pay for services. This line of argument is consistent with, yet extends, the consensus in the urban water literature, which suggests that poor service quality detracts from consumer willingness to pay for services (Whittington, et al 1990; Hensher, et al 2005). These intuitions also build on the literature on the politics of the welfare state, which suggests that citizens of countries with more universalistic welfare states are more supportive of devoting general tax expenditures to social services *because* they perceive themselves as benefiting or potentially benefiting from the service regime (e.g. Esping-Anderson 1990). Hence, we may observe the following:

H14: Receiving accurate, prior notifications regarding water delivery times and service disruptions will lead citizens to pay their water bills at a greater rate than citizens who *do not receive such notifications*.

While we hypothesize that these effects may be observable across the entire urban population in developing country cities with water intermittency, we expect these effects to be particularly pronounced under certain circumstances:

- For *marginal households (low income, religious minority, low caste, etc.)*, because they are less likely to have had influential political intermediaries prior to the intervention.[44]
- For *low-income households living in structures without automatically filling overhead tanks* (see previous section)
- For *households where someone spends a significant amount of time waiting or worrying about water* prior to the intervention. (We expect this to be closely related to the above conditions.)
- For *households receiving water services that typically arrive on a scheduled supply day within an interval of less than 4 hours*. Our intuition is that notifications are most useful when they occur within an interval of four hours or less, because individuals would be more likely to stay at or near home and wait for water under such circumstances. If arrival times are less predictable than this, household members are unlikely to stay close to home waiting for the water to arrive.
- For *effects to vary by neighborhood* type (level of cohesion, presence of neighborhood association, etc.)[45]

As in the previous section, then, we expect to observe *heterogeneous treatment effects* when we subset our analysis according to these intuitions.

Research Design

---

[44] We expect this to be the case even though recent studies suggest that India's marginal populations are now better able to exert pressure upon the state than previously (e.g. Corbridge *et al*. 2005; Banerjee & Somanathan 2007).

[45] We are currently assessing whether or not it will be possible for us to gather data on relevant neighborhood characteristics.

The effects of this informational intervention will be assessed in the context of a cluster-randomized experiment to be conducted in Bangalore, India, between May 2015 and December 2015. This section provides more detail on the nature of the intervention to be evaluated, the rationale chosen for our study site, our randomization and sampling strategy, the specific indicators to assess the variety of impacts we have outlined above, and power calculations for determination of sample size.

A. *Study Site and Timeline*

We will evaluate the hypothesized household-level impacts of NextDrop's services in the context of its ongoing rollout in the Indian city of Bangalore, a megacity of over 8 million, often called India's Silicon Valley. NextDrop signed a memorandum of understanding with the state water utility providing water and sanitation services in the city, Bangalore Water Supply and Sewerage Board (BWSSB), in May 2014, allowing it to collect information from BWSSB valvemen and provide notifications throughout the city.

Our research design takes advantage of the fact that NextDrop's rollout is taking place in stages. We focus our evaluation on a portion of the city where NextDrop does not yet operate: BWSSB's E3 subdivision, a socio-economically diverse district (roughly 20 km squared) in the eastern part of the city. Because NextDrop is required to serve at least 2/3 of the city by the end of 2014 under the terms of a DFID grant—in addition to the firm's desire to respond to BWSSB's request to roll out quickly—we were asked to restrict our evaluation to only one of the utility's 32 subdivisions. Exploratory fieldwork during summer 2014 increased our confidence in our original intuitions that the impact of NextDrop's services were likely to be higher in low income areas with residential structures of 1-2 stories, and where water delivery is variable in timing (yet not *completely* unpredictable). We thus sought a utility subdivision that contained a diverse population (where roughly 1/3 of our sample could be drawn from the bottom third of the city's income distribution), and that still possessed a reasonable number of low rise structures—a type of building environment that is growing less common in megacities like

Bangalore, but is typical of urban India more generally. We limited ourselves to consideration of subdivisions that BWSSB had not specifically requested NextDrop to expand into immediately, because we did not want our research to interfere with NextDrop's relationship with Bangalore's water utility.

After review of the limited (and only somewhat accurate) available government data on low-income settlements and population densities in Bangalore, and extensive site visits throughout the city in the summer and fall of 2014, we chose utility subdivision E3. This subdivision possessed the ideal combination of some low-rise residential neighborhoods, as well as several low and middle-income neighborhoods of sufficient size.[46] Our estimates suggest that approximately 20% of the area's residents – who include recent migrants from Tamil Nadu and Andhra Pradesh -- could be classified as Bangalore's bottom third of the income distribution. We estimate that roughly 25% of the population lives in low-rise structures (and therefore do not have overhead water tanks). Residents receive services 1-2 times a week, which is typical for urban India and also frequent enough that it is likely we will observe the notifications having an effect on our outcomes of interest if they are indeed useful for households. The area thus promises to allow us to analyze how the impact of NextDrop's intervention may vary according to the variety of criteria we outline, and in a setting that is reasonably representative for urban India.[47]

The main means of evaluating the effects of the intervention involves two surveys to the treatment group as well as to the control group, a baseline one prior to the intervention and a follow-up one for both groups after the treatment group has received services for 6 months. We will conduct the baseline survey and enroll households into the survey in May of 2015. We will then conduct the follow-up survey in November of 2015.We chose to run the trial for 6 months because this will allow

---

[46] As subsequent sections will explain, we employ a skip of 3 between sampled households, which increases the size of the low-income areas required to obtain a sufficiently large sample from low-income groups.

[47] Data from our baseline survey will allow us to compare in more precise terms the E3 population with the Indian urban population more generally.

households sufficient time to adapt their daily routines to the service. Allowing the trial to run longer would mean that we risk losing a larger portion of our study participants due to household moves, etc.

B. *Varying Treatment "Dosage"*

We make three disclaimers about the nature of the "treatment" in BWSSB subdivision E3. As mentioned above, NextDrop depends on the water utility's valvemen to send it information regarding water valve opening and closing times. NextDrop is working closely with both the valvemen and BWSSB to ensure that information is actually sent to NextDrop in the first place, and that such information is completely accurate. Yet the accuracy of this information cannot be assumed, and must be investigated explicitly—something that we will do in this study through questions in our follow-up survey. We expect that notifications will have a greater effect when they are accurate.

Second, NextDrop only provides notifications for water supplied by Bangalore's central water utility, BWSSB. Bangalore, however, has been growing outwards for several decades and has been incorporating existing townships and municipalities into its borders. As a result, households in some areas—including some within E3—receive piped water from both the main BWSSB network and a local source.[48] In such cases, NextDrop only provides notifications for water supplied by BWSSB. This means that while some households in Bangalore will receive notifications on every water supply day, others may only receive notifications every second supply day. Therefore, we must also investigate the frequency of NextDrop notifications relative to the frequency of supply.

---

[48] BWSSB supplies households with piped water from two sources in E3: "Cauvery" or "Kaveri" water (surface water from the Cauvery river) and "CMC water" (water piped from local borewells). CMC water used to be administered by local municipalities. Following the Bangalore government's annexation of outlying towns, BWSSB has been assuming control of CMC systems. It is in the process of assuming responsibility for the CMC systems in E3. In E3, NextDrop trains valvemen to send notifications for Cauvery water, and for CMC water where no Cauvery water is supplied. Where households receive both Cauvery and CMC water, BWSSB valvemen only send notifications for Cauvery water.

Third, NextDrop notifications may not always arrive a significant amount of time before water services commence. NextDrop estimates that notifications arrive 30 – 60 minutes before water service commence in Bangalore, but given that our study is taking place in a new area, detailed data on the elapsed time between notifications and water services is unavailable. There may also be some variation within E3 associated with differing distances between households and the main valve for a particular valve area. We will obtain data on the amount of warning households typically receive through their NextDrop notifications through our follow-up survey. We expect that households receiving notifications more in advance will value them more.

We plan to obtain data on the accuracy, frequency, and timing of NextDrop's services in order to examine the extent to which households receive varying "doses" of our treatment through specific questions in our follow-up survey. We conceive of the largest "dose" as one consisting of accurate and frequent notifications issues more than 30 minutes before water arrives. If significant variation exists, we will examine the extent to which stronger effects as associated with larger "doses" of our treatment. As this "dosage" is likely to vary mainly with some variable related to geography, we may use propensity score analysis to estimate counterfactuals through a single scalar of propensity scores (Imbens 2000; Joffe and Rosenbaum 1999; Rosenbaum and Rubin 1983).

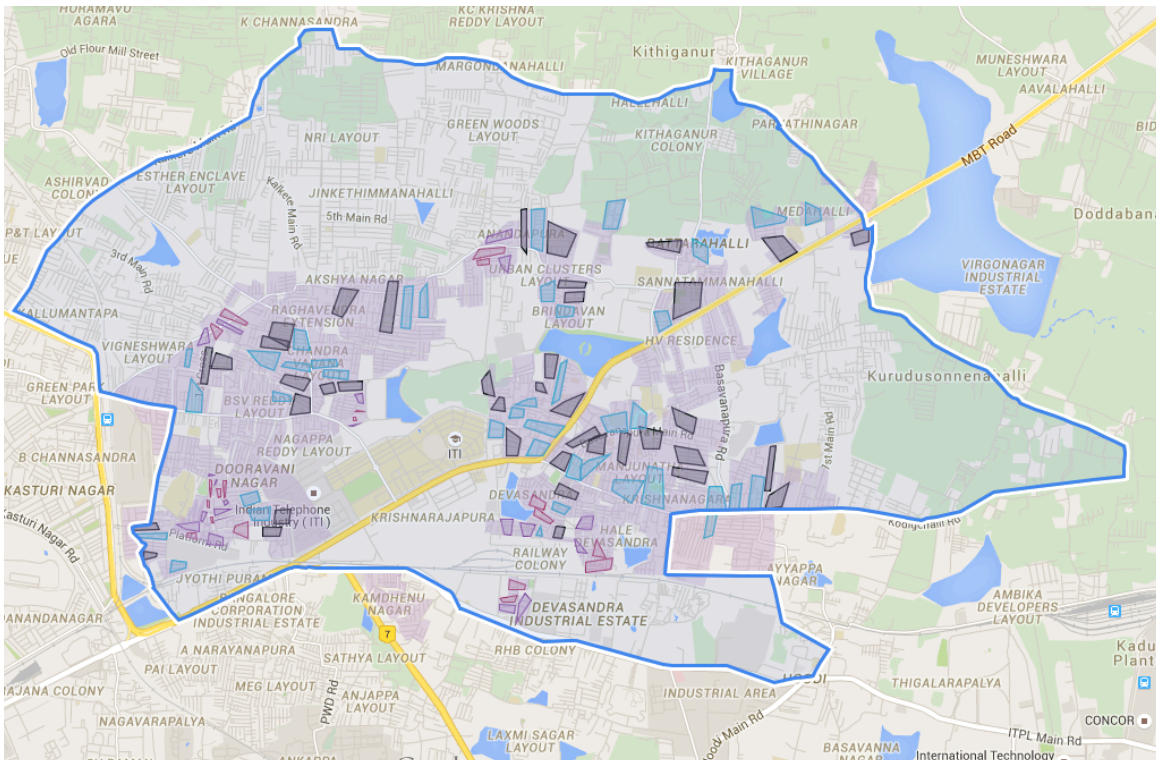
C. *Randomization and Sampling Strategy*

Within BWSSB subdivision E3, we are evaluating the impact of NextDrop's services through a cluster-randomized experiment. We chose a cluster-randomized design, rather than one with household-level randomization, because of concerns regarding information sharing between treatment and control households.[49] The 120 clusters of households in our study are separated from one another by at least two

[49] Our concerns regarding interference under household-level randomization were heightened when we conducted a small phone survey of existing NextDrop consumers in utility subdivision NE3, which specifically asked if information was shared

streets so as to create a physical buffer preventing information sharing between our treatment and control groups (see Figure 1). Within each cluster, we will utilize systematic sampling—i.e., employing a skip—between households in order to choose study participants. After piloting the survey in low-income areas, we decided that a skip of at least 3 between households would be sufficient to avoid "group interview" sessions in which neighbors group together to help respondents answer survey questions.

Figure 1. Low and Mixed Income Clusters within Study Area



Note: The BWSSB E3 subdivision boundary is shown in blue, while areas receiving piped water supply are denoted in lavender. Pink and purple polygons denote low-income clusters (treatment and control), while black and blue polygons denote mixed income clusters (treatment and control). There are four clusters per block.

---

with immediate neighbors. Small survey pilots in E3, conducted by the authors, also found that information of various types is shared between immediate neighbors, particularly in low-income neighborhoods.

Because blocking on a variable associated with the outcome of interest can improve the precision of causal estimates in the context of cluster-randomized experiments (Imbens 2011), we employed a geographic approach to stratification, or "blocking." Based on site surveys, we designated 30 geographic areas with a particular socio-economic character, either low income (10 blocks) or mixed income (20 blocks). We expect blocks in either one of these two categories to be similar not only in socio-economic terms, but also in terms of the state of the underlying water infrastructure. By blocking on socio-economic geography, we thus lay the groundwork for analyses of subsets of the data corresponding to areas where we expect to observe stronger effects: areas with poorer residents and with poorer quality water infrastructure.

Blocks were designated by our survey research organization following preliminary fieldwork by our team based on extensive site visits throughout E3.[50] Within each of these blocks, we outlined four clusters separated by two streets or lanes from one another.[51] Cluster boundaries were drawn so that clusters within a block were very similar to one another in terms of socio-economic mix. Within each block, we randomized two clusters to treatment and two to control. As mentioned above, within each cluster, survey enumerators will visit every third household. For both treatment and control groups, interviews will be conducted with the person responsible for managing and storing water for the household. Enumerators will interview a constant number of respondents per cluster.[52] GPS coordinates

[50] As mentioned previously, fine-grained census data on income does not exist for urban India. Our team therefore relied on visual cues such as dwelling type, number and type of vehicles parked along streets, and conversations with local residents to designate low and mixed income blocks.

[51] We include four clusters per block rather than two (the pairwise approach to blocking), following Imbens (2011).

[52] Note that attrition between wave 1 and wave 2 may mean that final counts will differ a little bit between clusters. Analyses will be weighted to account for this. The aim, however is equalize the N per cluster. We are collecting phone numbers for households in both the treatment and control group, which should help us avoid uneven attrition between treatment and control if treatment group members are enabled to work more outside the home because of NextDrop's notifications. (Individuals can be contacted via phone and enumerators can meet them after work.)

will be collected for each household to ensure compliance with the cluster design, to allow for the enrollment of treatment households in NextDrop's system (and especially the correct placement of treatment group members in valve areas), and facilitate follow-up surveys after six months.

Noncompliance is always a challenge for experimental researchers. In the case of this evaluation, the concern is that some households in our treatment group may not desire to enroll in NextDrop's notification services when our survey enumerators give them the opportunity to do so. In theoretical terms, what is of most interest to us is the impact of receipt of NextDrop notifications upon those that genuinely want to receive them, or the compliers. We will therefore employ a placebo design that involves asking respondents in both treatment and control group households whether they would like to receive notifications *when services are available in their area*, and will collect contact information for "compliers" in both the treatment and control group. NextDrop will enroll the treatment group compliers as soon as they receive their mobile phone numbers and GPS coordinates; they will then enroll the control group compliers in their system after six months, after we have conducted the follow-up survey that will be used for our impact analysis and collected other related data (details below, section E)). This research design will thus allow us to not only calculate treatment effects utilizing the ITT, but also calculate the Complier Average Causal Effect (CACE).

D. *Power Calculations*

Our power estimates focus on *one* of our outcome variables: reductions in the amount of time spent waiting for water on a weekly basis.[53] We focus on wait time because any changes in this are likely to generate changes in other outcomes of interest, such as stress levels and political attitudes. It is also an outcome measure for which we could obtain preliminary data prior to our study. Analyses

[53] One typically powers a study for one primary outcome of interest, even if multiple outcomes will be analyzed. As a result, we may not obtain statistically significant differences between treatment and control groups for some of our outcomes of interest.

regarding sample size, including the allocation of households between low-income and other clusters, were calculated on the basis of our preliminary estimates of the effects for the sub-population that we hypothesize will be most affected: low-income households, and particularly those living in low-rise structures. Based on our site visits to E3, and census data, we identified five major low-income areas, which could be divided into *10 blocks, each containing 4 clusters* of similar socio-economic character. (Making clusters smaller would not have been feasible given the small size of these areas, the need to accommodate at least four clusters within each low-income area for blocking, and the need to leave two streets between clusters to prevent information sharing.)

In addition to the number of clusters, we also need estimates for the size and spread of the effects we hope to capture through our study. Unfortunately, prior surveys on wait times for water in Bangalore do not exist, and the effect of an intervention like that offered by NextDrop has not been evaluated in another location. We therefore relied on a few different data sources to estimate the expected size and standard deviation of our expected effects for this population. First, we conducted a very small pilot survey within two low-income areas in E3, within which we asked how much time, on average, household members currently spend at home on supply days because they expect the water to arrive. We then calculated expected reductions for these households based on the assumption that wait time could be reduced to one hour on a given supply day if households received NextDrop notifications.[54] These inquiries yielded estimates of 30-minute (s.d. = 1 hour) and 45 minute (s.d. = 1.76 hour) weekly reductions for the two areas. We also analyzed NextDrop's valvemen report data regarding valve opening times for a subdivision in which they have operated for over a year, NE3. We drew a systematic sample from among the valve areas, and estimated a mean wait time for each valve area based on deviations from the beginning of the most common start time for services on supply days. (In other words, in the absence of household level data, we sampled from valve areas, assuming these serve

[54] We assumed that households would stop waiting after 3.5 hours, and would not bother staying at home continuously if arrival times were so unpredictable that one could not anticipate if services would start in the morning or afternoon.

as an acceptable proxy for household level data for preliminary estimates.) Relying on similar assumptions as in the previous exercise, we calculated a mean reduction of 2 hours a week (s.d. = 2). We treat this set of estimates from different sources as lower and upper bound estimates for our treatment effect in power calculations.

For cluster-randomized experiments, power calculations must also include estimates of the intra-cluster correlation (ICC). In this case, as before, we could not build on insights from pre-existing surveys. We therefore consulted the literature on health studies related to household interventions in contexts with varying socio-economic background conditions and from various countries. We observed that ICC values typically fall in the 0.01 – 0.15 range, and thus use these as lower and upper bounds for our calculations.[55]

To detect a CACE of the size discussed above through an experiment with forty clusters, with 80% power, and at the 0.05 significance level (two-tailed test), calculations suggest we would need between 150 and 500 respondents falling in our "high impact" population, or 4-13 households per cluster. Given our estimates that only 2/3 – 4/5 of the population in these low income clusters actually fall in our hypothesized "high impact" group, this suggests the need to sample 5-16 households per cluster. If we assume an 80% compliance rate with our intervention (agreement to enroll in NextDrop services), and a 20% attrition rate between the baseline and follow-up services, this suggests we need to sample 7-22 households per cluster to detect an effect within this group in these areas.[56]

To detect an effect through an ITT analysis would require more households per cluster, as our original expectations regarding the size of the average effect for treatment areas would be lower due to the inclusion of noncompliers, which we estimate to be roughly 20% of our sample. When we down-

[55] E.g., Groves *et. al.* (2013, Chapter 4); Smeeth *et al*. (2002).

[56] We expect to experience some attrition because households may move to another part of the city or back to their home villages. To capture the ITT, we would need a roughly 20% larger sample size in these areas, as the mean expected effect would be reduced to account for the inclusion of noncompliers.

weight the size of our expected effect (and modify the standard deviation) accordingly and conduct power analyses, calculations suggest that we would need to observe 230-700 households, or between 6 and 17.5 households per cluster, in order to observe an effect. Note that these calculations assume that *everyone* surveyed in these clusters will experience a reduction in wait time. We have estimated that only 2/3 – 4/5 of the population in our identified low-income areas falls in our target population (low income and low rise). So to obtain a yield of 6-17.5 low income/low rise households within these clusters, we should ultimately survey (and retain as compliers) 8-25 households per cluster. Moreover, if we experience 20% attrition between wave 1 and wave 2, we would need to survey 10 – 31 households per cluster to detect an effect. Overall, this analysis suggests that an original *sample size of 25 HH per cluster in our 40 low income clusters*, or 1000 HH total, will be sufficient for us to detect an effect (either CACE or ITT) of the size we expect if there is indeed one at work.

Because the remit for our study includes evaluating the effect of NextDrop's intervention *across income groups*, we also examined how many households would be required to detect an effect roughly half the size of that we observed through our pre-piloting among low income groups. Overall, this analysis suggests that sample size of 2000 HH, split between 80 clusters with 25 HH each, would allow us to assess whether or not NextDrop is having an impact on wait times within the mixed income clusters.

E. *Specific Indicators*

The impacts of NextDrop's services will be assessed by observing differences between our treatment and control groups on behavioral and survey-based indicators. While most impacts will be measured through responses to questions in the baseline and post-intervention surveys, we plan to draw on at least one behavior measure. This behavioral measure is outlined in Table 1.

Table 1: Behavioral Indicator for Treatment Effects

Hypothesis	*Indicators*
H14: Receiving prior notifications regarding water delivery times and service disruptions will lead citizens to pay their water bills at a greater rate than citizens who do not receive such notifications.	-We will examine differences in changes in monthly payment rates and arrears between our treatment and control groups through examining BWSSB billing records

References (submitted with Pre-Analysis Plan)

Auerbach, Adam. 2014. "Clients and Communities: The Political Economy of Party Network Organization and Development in India's Urban Slums." Manuscript.

Baisa, B, L.W. Davis, S. W. Salant, and W. Wilcox. 2010. The welfare costs of unreliable water service. *Journal of Development Economics*, 92: 1 – 12.

Banerjee, A & R. Somanathan. 2007. The political economy of public goods: Some evidence from India. *Journal of Development Economics*, 82: 287 – 314.

Corbridge, Stuart, Glyn Williams, Manoj Srivastava, Réne Véron. 2005. *Seeing the State: Governance and Governmentality in India*. New York: Cambridge University Press.

Ercumen, A., B F. Arnold, E Kumpel, Z Burt, Isha Ray, KL Nelson, and J M Colford Jr. (in press). Upgrading a Piped Water Supply from Intermittent to Continuous Delivery: Impact on Waterborne Illness in Urban India

Esping-Andersen, Gosta. 1990. *The Three Worlds of Welfare Capitalism*. Cambridge, UK: Polity.

Estache, Antonio, Gómez-Lobo, Andrés, and Daniel Leipziger. 2001. "Utilities Privatization and the Poor: Lessons and Evidence from Latin America." *World Development*. 29(7): 1179-1198.

Evans, Peter. 2008. In search of the 21st century developmental state. Working Paper 4. Center for Global Political Economy. University of Sussex.

Ferguson, J., and A. Gupta. 2002. Spatializing states: towards an ethnography of neoliberal governmentality. *American Ethnologist*, 29: 981 – 1002.

Ghertner, D. Asher. 2011. "Gentrifying the State, Gentrifying Participation: Elite Governance Programs in Delhi." *International Journal of Urban and Regional Research*. 35(3): 504-532.

Groves, Robert M., Floyd J. Fowler, Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, Roger Turangeau. 2013. *Survey Methodology*. (Second edition). Hoboken, NJ: John Wiley and Sons.

Harriss J., 2006. Middle class activism and the politics of the informal working class. *Critical Asian Studies*, 38: 445 – 465.

Hensher, D, et al. 2005. Households' Willingness to Pay for Water Service Attributes. *Environmental and Resource Economics*, 32: 509 – 531.

Imbens, Guido. 2011. ""Experimental Design for Unit and Cluster Randomized Trials." Manuscript prepared for the International Initiative for Impact Evaluation, 3i.

Imbens, Guido. 2000. The role of the propensity score in estimating dose-response functions. *Biometrika*, 87: 706-710.

Joffe, M. & Rosenbaum, P. 1999. Invited commentary: propensity scores. *American Journal of Epidemiology*, 150: 1-7.

Kjellén, M, and G. McGranahan. 2006. Informal water vendors and the urban poor. *Human Settlements Discussion Paper Series*. IIED. http://www.iied.org/HS/publications.html

Kumpel, E. and Nelson, K.L. 2013. Comparing microbial water quality in an intermittent and continuous piped water supply. *Water Research* 47(14): 5176-5188.

Kuriyan, Renee and Isha Ray. 2009. "Outsourcing the State? Public-Private Partnerships and Information Technologies in India." *World Development*. 37(10): 1663-1673.

Madon, S and S. Sahay. 2002. An information based model of NGO mediation for the empowerment of slum-dwellers in Bangalore. *The Information Society*, 18: 13 – 19.

Mullainathan, S and E. Shafir. 2013. *Scarcity: Why Having Too Little Means So Much*. New York: Henry Holt & Company.

Ranganathan, Malini. 2014. 'Mafias' in the waterscape: Urban informality and everyday public authority in Bangalore. *Water Alternatives*, **7**:89 – 105.

Rosenbaum, P. and Rubin, D. B. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41-55.

Savedoff, William and Pablo Spiller. 1999. *Spilled Water: Institutional Commitment in the Provision of Water Services*. Washington, D.C.: Inter-American Development Bank.

Smeeth, Liam and Edmond Siu-Woon Ng. 2002. "Intraclass correlation coefficients for cluster randomized trials in primary care: data from the MRC Trial of the Assessment and Management of Older People in the Community." *Controlled Clinical Trials*. 23: 409-421.

Smith, Sandra. 2007. *Lone Pursuit: Distrust and Defensive Individualism among the Black Poor.* New York: Russell Sage Foundation.

Stokes, Susan C., Thad Dunning, Marcelo Nazareno, and Valeria Brusco. 2013. *Brokers, Voters, and Clientelism: The Puzzle of Distributive Politics*. New York: Cambridge University Press.

Tolbert, C and K Mossberger. 2006. "The Effects of E-Government on Trust and Confidence in Government." *Public Administration Review,* 66: 354-369

United Nations Development Programme. 2010. Human Development Report. New York: UNDP.

van den Berg, C and A Danilenko. 2011. *The IBNET Water Supply and Sanitation Performance Blue Book*. The World Bank: Washington DC.

Whittington, D., et al. 1990. Estimating the Willingness to Pay for Water Services in Developing Countries. *Economic Development and Cultural Change*, 38: 293 – 311.

Wutich, A and K. Ragsdale. 2008. Water insecurity and emotional distress: Coping with supply, access, and seasonal variability of water in a Bolivian squatter settlement. *Social Science and Medicine*. 67: 2116 – 2125.

Zérah, Marie-Helène. 2000. Household strategies for coping with unreliable water supplies: the case of Delhi. *Habitat International*, **24**: 295 – 307.

Appendix including copy of survey, flyers, etc. to be included with the Pre-Analysis Plan upon registration with EGAP during Winter 2014

Amendment to Pre-Analysis Plan 20150514AA

December 11, 2015

In May of 2015, we registered a cluster-randomized experiment titled "The Household Welfare and Political Impacts of Increasing Service Predictability: An Experimental Intervention in Bangalore's Water Sector."[57] Briefly, the experiment aims to assess the impact of SMS notifications about water supply arrival and cancellation on household behavior in Bangalore, a context in which water supply is intermittent. The baseline survey was conducted in the spring of 2015, after which treatment was administered by NextDrop, a Bangalore based social enterprise. The endline survey is currently on-going; not only have we not finished collecting data on outcomes, but we have not yet begun to clean or analyze this data. Nevertheless, the baseline data, interviews with NextDrop, and data from NextDrop have given us information we did not have when registering our original pre-analysis plan. This amendment discusses our plan to drop a block of treatment and control units from the study along with our planned analysis of partial non-compliance (also known as "treatment dilution"[58]) and heterogeneous treatment effects.

1. Units to drop from the study

We will need to drop one block of treatment and control clusters due to the fact that NextDrop decided (after we had designated clusters and randomized units to treatment and control) that it would be too difficult administratively for them to solicit notifications for a small set of the valve areas in E3. In particular, valve areas in utility subdivision E3 fall in two categories: a) CMC valve areas which draw on borewell water, and which pre-date BWSSB's administration of services in the area; and b) Cauvery

[57] ID 20150514AA.

[58] See Angrist, Joshua D. 2006. "Instrumental Variables Methods in Experimental Criminology Research: What, Why, and How." *Journal of Experimental Criminology.* 2:23-44.

valve areas, which draw on water from the Cauvery river. Subsequent to the design of our study, NextDrop found it more difficult to work with the CMC area valvemen, so decided not to solicit notifications from them. This problem affects one of our treatment clusters; as a result, we will be dropping all four clusters (two treated, two control) that form the block within which this particular treatment cluster lies.

2. Dilution of Treatment/Partial compliance

As described in the original design registration, NextDrop obtains information regarding the timing of water delivery in particular "valve areas" by collecting water flow information from valvemen—the individuals responsible for opening and closing the valves controlling water into particular districts—and disseminating notifications to NextDrop customers. NextDrop was able to solicit this information successfully in its pilot location, the city of Hubli-Dharwad. Our study, however, capitalizes upon NextDrop's rollout in a new location, Bangalore, where the enterprise set aside a section of the city for the purposes of our impact evaluation. When it began soliciting notifications in our study area, NextDrop found valvemen to be less willing to provide notifications consistently than in Hubli-Dharwad. During the course of administering treatment in our study area, we have found that valvemen may provide NextDrop with information infrequently in some cases, directly affecting the frequency of information received by customers. Furthermore, certain households lie in areas serviced *both* by BWSSB and CMC valve areas, as described above (see Figure 1). Because we expect NextDrop to have received fewer notifications from valvemen operating CMC valves, it is very likely that households in these areas receive notifications a smaller fraction of the time than households located just in Cauvery valve areas.

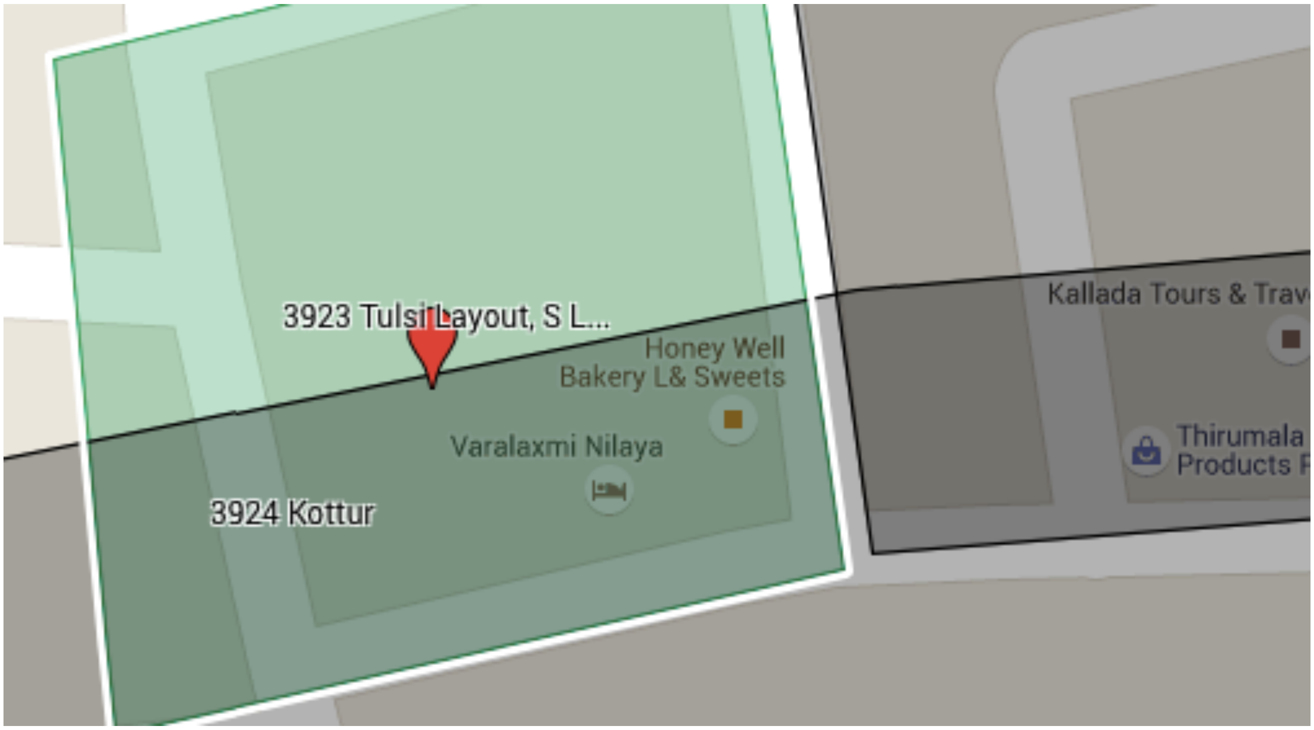


Fig 1. This household lies in an area covered by BWSSB and CMC valve areas

This problem appears to be a case of “dilution of treatment” (Angrist 2006), or what is referred to in more recent literature as partial compliance (Jin and Rubin 2008). The original design registration acknowledged one source of non-compliance, namely households’ desire to enroll into NextDrop’services, and outlined a plan to calculate both intention-to-treat (ITT) and complier average causal effects (CACE). We will continue to follow this plan while including this new source of non-compliance. Importantly, we have information on treatment dilution for both treatment and control groups.

If feasible, we will also analyze the effects of different levels of compliance or dilution of treatment using the method of principal stratification as described by Frangakis and Rubin (2002) and Jin and Rubin (2008). If we do so, we will create strata based on levels on non-compliance, and analyze causal effects within these strata. In addition to household reported frequency of notifications, we will use NextDrop's data on valveman reports to confirm customer reports of notification frequency and place control households within strata.

3. Heterogeneous treatment effects

An analysis of our baseline survey data reveals variation in underlying water supply conditions for our sample. For example, the number of water supply days per week can range from 1 to 7, with some respondents claiming that there is no regularity to the supply. It is likely that treatment effects will be larger for households with poorer supply conditions. In addition to the variables for which we identified potential heterogeneous treatment effects in the original design registration, we will analyze treatment effects conditional on the following variables: frequency of water supply and frequency of supply cancellation. Data on these variables was collected before treatment in the baseline survey.

We also expect to see variation in respondent employment status and distance to the respondent's workplace. It is possible that treatment effects will be greater for the employed or for those who work far from home. As a result, we will calculate treatment effects conditional on these variables. Data on these variables was also collected before treatment in the baseline survey.

Additionally, while possessing a household mobile phone was a prerequisite for enrolling in our study, the individual within a household who keeps this phone may not be responsible for waiting for water. It is unlikely that NextDrop's messages will have an effect on household behavior unless the individual waiting for water also keeps the mobile phone. In both waves of our survey, we ensure that the individual being interviewed is also the household member responsible for collecting and storing water. In the second wave of the interview, we also ask if this individual keeps the mobile phone that was registered for the receipt of NextDrop notifications. We will analyze treatment effects conditional on whether or not the survey respondent keeps the mobile phone.

Finally, as described above, the behavior of valvemen is directly responsible for the information disseminated to customers. In addition to variation in the frequency of message delivery, we may also see variation in the accuracy or quality of information received by households. We expect to see greater effects for households to whom accurate information is delivered, and have asked households about the accuracy of NextDrop's notifications in the second wave of our survey. We will calculate treatment effects conditional on household responses to this question.

References

Angrist, Joshua D. 2006. "Instrumental Variables Methods in Experimental Criminology Research: What, Why, and How." *Journal of Experimental Criminology*. 2:23-44.

Frangakis, C.E. and Rubin, D.B. (2002). "Principal stratification in causal inference." *Biometrics* 58(1): 21-29

Jin, H. and Rubin, D.B. (2008). "Principal stratification for causal inference with extended partial compliance." *Journal of the American Statistical Association* 103 (481): 101-111.

VanderWeele T.J., and Hernan, M. A. (2013). "Causal inference under multiple versions of treatment." *Journal of Causal Inference* 1 (1): 1-20.