Demand Estimation with Strategic Complementarities: Sanitation in Bangladesh

Raymond Guiteras
NCSU

James Levinsohn
Yale University

Mushfiq Mobarak
Yale University

June 2018

Abstract

For many products, the utility a household derives from its consumption depends on the share of other households that also adopt. We estimate a structural model of demand that allows for these inter-dependencies. We apply our model to the adoption of household latrines – a technology that has large consequences for public health. We estimate the model using data from a large-scale experiment in which over 18,000 households in 380 communities in rural Bangladesh were randomly assigned incentives to purchase latrines. Multiple levels of randomization were designed to identify both the own-price effect on demand as well as a strategic complementarity effect. We conduct counter-factual simulations to explore the role of subsidies in the presence of strategic complementarities. We examine several policy options, including targeting on the basis of household poverty or neighborhood population density, and the policymaker's tradeoffs along the price, saturation and scope margins.

Email: rpguiter@ncsu.edu, james.levinsohn@yale.edu, ahmed.mobarak@yale.edu. For helpful comments we thank Francesco Amodio, Steve Berry, Gilles Duranton, Ganesh Iyer, Laura Lasio, Duncan Thomas, Adrian Tschoegl, seminar participants at Harvard/MIT (Development Economics), McGill, North Carolina State University, University of Cape Town, University of Vienna, Wharton, Yale SOM, Yale Political Science, and conference participants at the New Frontiers in Development Economics Conference at UConn, GGKP Annual Conference (World Bank/OECD/U.N) and the China India Customer Insights Conference at the Cheung Kong GSB. We thank Meir Brooks, Julia Brown, Zhi Chen, Seungmin Lee, Benjamin Lidofsky, Derek Wolfson and Tetyana Zelenska for outstanding research assistance, Innovations for Poverty Action-Bangladesh (Mahreen Khan, Ishita Ahmed, Ariadna Vargas, Rifaiyat Mahbub, Mehrab Ali, Mehrab Bakhtiar, Alamgir Kabir, Ashraful Haque) for field support, the Bill and Melinda Gates Foundation and the International Growth Centre for financial support, and Wateraid-Bangladesh and VERC for their partnership in implementing the interventions.

1 Introduction

The utility of a purchase often depends on the number or share of others making the same purchase. This phenomenon has many labels – 'peer effect,' a 'network externality,' or a 'strategic complementarity in demand' – and there are myriad examples. Strategic complementarity in demand exists when farmers learn from their neighbors and decide whether to adopt a new technology (Conley and Udry 2010). It is present when a consumer decides whether to adopt mobile phone service, and the value of the service depends on how many others are on the same network (Bjorkegren 2015). Economists have identified demand complementarities in decisions about whether to invest in home energy efficiency (Allcott 2011), how much labor effort to expend (Mas and Moretti 2009), whether to migrate (Munshi 2003; Meghir et al. 2015; Akram et al. 2017), purchase insurance (Mobarak and Rosenzweig 2013; Kinnan 2017), or enter the labor force in the presence of gender norms on who works (Iversen and Rosenbluth 2010; Bertrand 2011). This list is illustrative, but hardly exhaustive.

When strategic complementarities are positive, as in all the examples listed above, a policy to promote adoption may be welfare enhancing. In many instances, the key barrier to adoption is price (J-PAL 2011), which makes subsidies an obvious policy lever. However, the precise design of subsidy policy in the presence of strategic complementarities is not straightforward. There are two main challenges – one conceptual and one econometric.

The conceptual challenge relates to understanding the effects subsidies will have in the presence of interdependent decision-making. While subsidies have the usual own-price effect, they also have an indirect spillover effect on others' adoption decisions. That is, while a subsidy will increase my likelihood of adoption, my decision feeds into my neighbors' decisions which in turn impact my own decision. We model this formally below by introducing strategic complementarity into the household adoption decision. This interdependence in decision-making introduces complexities into the prospective analysis of subsidy policy. For example, if our goal is to maximize adoption, should a fixed subsidy budget be widely distributed, or should a smaller number of households be targeted with larger subsidies? When household heterogeneity is modeled, what type of targeting is most efficacious? Indeed, even computing the own-price elasticity of demand, a key primitive to any analysis, is not straightforward when prices affect a household's decision both directly and through the decisions of its peers.

The second challenge to fashioning subsidy policy is econometric and is often referred to as "The Reflection Problem" (Manski 1993).

> "[The Reflection Problem] arises when a researcher observing the distribution of behaviour in a population tries to infer whether the average behaviour in some group influences the behaviour of the individuals that comprise that group."

Manski concluded his influential paper by noting that "Given that identification based on observed behaviour alone is so tenuous, experimental and subjective data will have to play an important role in future efforts to learn about social effects." We address the econometric challenge by taking Manski's advice to heart and collecting experimental data. We estimate key demand parameters of a model of inter-dependent decision-making using a randomized controlled trial (RCT) that was designed to identify the own-price response as well as the effect that interdependency has on demand (Manski's "social effect").

While these challenges of evaluating subsidy policy in the presence of strategic complementarities are present in many contexts, our specific empirical application is the adoption of latrines in a developing country. This is an important policy issue in its own right, for sanitation practices have significant consequences for human health and welfare. About one billion people practice open defecation (OD) (WHO et al. 2014). The attendant health burden falls principally on the poor. Diarrheal disease kills nearly a million people a year (Prüss-Ustün et al. 2014), and is the cause of nearly 20% of deaths of children under five in low income countries (Mara et al. 2010). Latrine use has been shown to improve public health and generate positive externalities (Spears 2012; Pickering et al. 2015; Hathi et al.

2017; Geruso and Spears 2018; Gautam 2017), but adoption rates remain low. Price is an important determinant of latrine adoption decisions[1], making subsidies an important policy lever (Gautam 2018).

Strategic complementarities are likely important in the sanitation adoption decision for at least three reasons. First is an epidemiological or technical complementarity. A household's utility from adoption is higher if neighboring households also adopt, and conversely the benefits of adoption are muted or nullified if neighbors practice OD.[2] Second, social norms may be important. In a community in which OD is the norm, the "social cost" of OD may be low, whereas it may be very high in a community in which everyone is expected to use a toilet. Third, investing in latrines may allow neighbors to learn about the technology and change their perceptions about the net benefits of adoption.

We conducted a large-scale field experiment on sanitation behavior with over 18,000 households in 380 communities in rural Bangladesh. We randomly varied (1) the price specific households faced to identify the direct effect of price on adoption, and (2) subsidy "saturation" – the fraction of each community offered subsidies – to identify the indirect effect of others' adoption decisions. We then estimate a discrete choice model of demand in the presence of strategic complementarities. Both the large number of communities, and the large number of households per community in our experiment are useful for precisely estimating the own price and the demand spillover effects. The estimated structural model then allows us to evaluate several prospective policies.

Our estimates indicate that subsidies encourage latrine adoption, and that adoption decisions are strategic complements. Holding own price constant, a household becomes more likely

[1] Data from both Bangladesh (Guiteras et al. 2015) and Indonesia (Cameron and Shah 2017) show that poor people do not invest in latrines without price subsidies, and many households cite the high cost of construction as their primary barrier.

[2] A latrine may only generate a return in the household's health production function once the surrounding environment becomes sufficiently clean. In the multi-country study with health outcomes data, (Gertler et al. 2015) argues that there communities need to reach a threshold in sanitation coverage before we see impacts on child height.

to invest if a larger fraction of its community are also offered a subsidy. These estimates are statistically precise, and form the basis for our counterfactual simulations to identify subsidy policies that would increase aggregate adoption rates in a community. We explore the effects of widely distributing the subsidy budget versus concentrating larger subsidies to a few; targeting subsidies on the basis of household characteristics such as poverty or social network position; and targeting on the basis of community attributes such as population density. These simulations help us design mechanisms that increase aggregate adoption without changing the subsidy budget outlay.

We run additional randomized experiments to identify behavioral mechanisms that may underlie the strategic complementarity in demand. We find that social norms are relevant, but not in the way that we had expected. Subsidizing socially *marginal* households to invest in latrines produces a positive spillover on others' adoption, whereas adoption by community leaders and socially central households do not influence others as much. We interpret this to mean that *shame* is a key driver of behavior in this setting. When people occupying lower social strata start using a new latrine technology, it becomes shameful for others to continue defecating in the open. Changes in social norms appear to be more "downward-facing" rather than "upward-looking". This insight is useful for the design of social marketing of new products, behaviors, and technologies in sociology (Kim et al. 2015), economics (BenYishay and Mobarak 2018; Akbarpour et al. 2017; Banerjee et al. 2013; Beaman et al. 2018) and marketing (Chan and Misra 1990). Leadership and centrality are popular concepts in these applications, but our results suggest that social influence may work very differently in some settings.

Our analysis demonstrates how two methodologies that are often viewed as competing alternatives – RCTs and structural estimation – can fruitfully serve as complements. RCTs are typically able to convincingly identify behavioral responses to specific treatments, but RCTs alone are not well-suited to analyzing counter-factual policies. Structural models, on the other hand, are frequently well-suited to policy analysis but questions often arise as to the identification of key behavioral parameters. Our paper is an example of bridging the divide (Todd and Wolpin 2006; Kremer et al. 2011; Duflo et al. 2012; Attanasio et al. 2012).

The methodology developed and implemented in this paper is applicable to contexts beyond sanitation. Our general approach is relevant to conducting policy analysis whenever demand inter-linkages are present. As noted at the outset, strategic complementarities are present across many fields of economics and their prevalence extends to a broad spectrum of the social sciences. For example, social norms even guide the willingness to engage in bullying (Paluck et al. 2016) or in militia violence and genocide (Yanagizawa-Drott 2014). They can affect decisions on how much to contribute to public goods or charity (Kessler 2013), whether to purchase health products (Oster and Thornton 2012; Kremer and Miguel 2007), or the type of financial asset chosen (Bursztyn et al. 2014). The effectiveness of policies to either promote or deter such behaviors and actions depend on the nature or size of the behavioral spillovers across individuals. The methods developed in this paper become useful for policy analysis in such settings.

The paper proceeds as follows. In Section 2, we discuss the context and design of our RCT. Section 3 introduces our model of demand and the resulting econometric framework. Section 4 discusses estimation and presents results. In Section 5, we use the estimated structural model to explore several policies to encourage sanitation adoption. Those results highlight the importance of strategic complementarities and in Section 6, we investigate possible mechanisms driving the inter-dependency of adoption decisions. Section 7 concludes.

2 Context and Experimental Design

In this section, we describe the context for our study and the design of our experiment. Additional detail can be found in the online Supplementary Materials to Guiteras et al.

(2015).

2.1 Context

Our study took place in four rural "unions" (the local administrative unit) of Tanore district in northwest Bangladesh. We chose this area primarily because it has higher prevalence of open defecation relative to other rural areas of Bangladesh. For example, in our baseline survey, conducted in 2011, only 39.8% of households owned a hygienic latrine, while 30.8% of adults regularly practiced open defecation. The sample for the experiment and the data collection consisted of the universe of 18,254 households residing in 380 neighborhoods (locally known as "paras") in 107 villages. The neighborhood is a key unit in this study, since it is the level at which we conducted our interventions and at which we study social spillovers. While the neighborhood is not a formal administrative unit, its definition and boundaries are typically commonly understood.

Our baseline survey measured land holdings as a proxy for wealth, because land is the most important (and easily observable) component of wealth in rural Bangladesh. To qualify for the subsidy interventions described below, a household had to fall into the bottom 75% of the distribution of landholdings. The exact cut-off varied from village to village, but was typically about 50 decimals or half an acre of land. 35.1% of households were landless, meaning they had no (agricultural land) holdings beside their homestead. All landless households were eligible for subsidies. These landless households generally possessed a homestead where they could install a latrine. The empirical analysis in this paper will focus on the 12,792 subsidy-eligible households.

2.2 Experimental Design

At an intuitive level, our structural model of demand requires two key inputs. One, which we call the direct effect, is how adoption decisions respond to a change in price, holding everything else constant. The other, which we term the indirect or spillover effect, is how demand responds to a change in the share of one's neighbors that install a latrine, holding price constant. Both the direct and indirect effects enter the own-price elasticity of demand. Concretely, when prices fall, my household's demand may increase for two reasons. The first is the usual direct effect: price enters directly into my household's utility function, so my household is more likely to purchase when the price we face falls, independent of other households' behavior. The second is an indirect effect: if the price other households face also falls, they are more likely to adopt, and their adoption may affect my utility of adoption. We formalize this in Section 3, but the notion of direct versus indirect effects inform our experimental design. Each effect must be convincingly identified.

The RCT has three levels of randomization: village, neighborhood, and household. The basic policy lever we employ is a subsidy for hygienic latrines. At the highest level of randomization, we randomly selected 63 villages for the subsidy treatment, while the remaining 44 received no subsidies.[3] In the 63 subsidy villages, we conducted lotteries that allocated vouchers that gave the voucher-holder a subsidy toward the purchase price of a latrine. The probability of winning a voucher, which we call "saturation" in what follows, was randomized at the neighborhood level. In low-saturation neighborhoods (L, $N_g = 74$), 25% of eligible households won a voucher. In medium-saturation neighborhoods (M, $N_g = 77$), 50% of

[3] The "no subsidy" villages include villages from three categories: pure control villages that received no treatment ($N_v = 22$); "supply only" villages that received only a treatment intended to improve the functioning of sanitation markets through information on supply availability ($N_v = 10$); "Latrine promotion program only" villages that received only a collective demand-stimulation and motivational treatment ($N_v = 12$) described below, without subsidies. Neither of these non-subsidy treatments had economically meaningful or statistically significant effects on demand relative to control (Guiteras et al. 2015), so we combine the three categories into a single "omitted category" in our regressions to increase power. The stimulation/motivation treatment is also sometimes known as "Community Led Total Sanitation", and it did not meaningfully increase latrine adoption in another randomized evaluation in Indonesia either (Cameron and Shah 2017).

households won, while in high-saturation neighborhoods (H, $N_g = 77$), 75% won.[4] Randomizing saturation allows us to identify the role of strategic complementarities in demand.

In each neighborhood, eligible households participated in two independent, public, household-level lotteries. In the first, we randomly allocated vouchers for a 75% discount on sets of latrine parts. Prior to implementing our interventions, we worked with all 11 masons operating in the four sample unions to establish a standardized set of latrine parts with a fixed, unsubsidized price of USD 48. With the voucher, the household could purchase the set of parts for USD 12.[5] Vouchers were linked to households, and we stationed a project employee at each mason's shop to ensure that only winning households could redeem vouchers.[6]

In addition to the price of the latrine components, households also incurred transportation and installation costs. These costs vary by village, and averaged about USD 7-10.[7]

The second lottery was for free corrugated iron sheets, worth about USD 15, to build a roof for the latrine.[8] We refer to this as the "tin lottery." The tin lottery was independent of the

[4]We implemented a simple voucher distribution scheme in half the subsidy villages, and we hit these randomization targets quite precisely in those communities: 24.9%, 50.6%, and 72.9% of households received vouchers in L, M, and H saturation neighborhoods, respectively. In the other half, we implemented an "Early Adopter subsidy" scheme in which 75% of households were targeted with vouchers, but only the first 20% (or 40% or 60%) of households that show up to redeem vouchers in the communities assigned to L (or M or H) saturation are provided the discounts. In practice, these limits on the number of early adopters turned out not to be a binding constraint, and all households that wanted to redeem vouchers were able to. All empirical results reported in this paper remain qualitatively very similar, whether we conduct our analysis on the full sample of neighborhood, or drop the sample of "early adopter" communities.

[5]We pre-negotiated prices and voucher values with masons prior to launching any of our interventions, and masons were not allowed to adjust prices during the course of our project after observing demand conditions. We thereby shut down potential supply-side channels through which each household's adoption decision can affect others (e.g. by changing market prices through economies of scale in production). We therefore restrict our focus to demand spillovers in this paper.

[6]We have investigated the possibility that households sold their vouchers to others in a secondary market, but have found no evidence of such behavior. For example, we will document demand complementarities even within the set of households that won vouchers themselves. This is informative, because such households did not need to buy or borrow a voucher from another household in order to invest.

[7]There were three models available at fixed prices: a single, 3-ring pit (USD 22 unsubsidized, USD 5.5 with a voucher); a single, 5-ring pit (USD 26 unsubsidized, USD 6.5 with a voucher); a dual-pit (USD 48 unsubsidized, USD 12 with a voucher). We focus on the latter because it was by far the most popular.

[8]The additional financial cost to households interested in building walls to complete a privacy shield for the latrine ranged from close to zero for a simple, self-made bamboo structure if the household gathered and cut bamboo on its own, to USD 20 for a bamboo structure made with purchased bamboo and built by a skilled artisan, to as much as USD 85 for a structure with corrugated iron sheets for walls and reinforced by treated wood.

latrine subsidy lottery although to collect the tin, winners had to install a latrine. The two lotteries combined to place eligible households in each neighborhood into one of four cells: won both lotteries, won the voucher but lost the tin lottery, won the tin but lost the voucher lottery, and lost both. In our estimation strategy, these randomly assigned prices identify the own-price effect.

To summarize, the two types of villages (no subsidy vs. subsidy), three saturation levels across subsidy neighborhoods (low, medium and high saturation), and independent household-level lotteries within subsidy neighborhoods (latrine voucher, tin) provide the random variation we exploit to avoid the traditional sources of endogeneity when estimating demand in the presence of strategic complementarities. At a high level, there are two sources of endogeneity. The first arises due to unobserved demand shifters. These will be correlated with price, an included regressor, by inspection of the supply curve. The other source of endogeneity arises due to the interdependent nature of demand and is the focus of Manski's "reflection problem." Crucially, our design generates random variation both in individual household price and in the overall price environment among the household's peers hence addressing both sources of endogeneity.

While the random variation we introduce and exploit for estimation was in the subsidy assignment and the proportion of the neighborhood subsidized, all of the neighborhoods where subsidies were assigned (in low, medium or high saturation) were also provided some information about sanitation behavior prior to the latrine voucher lotteries. We call this public health education and motivation campaign a "Latrine Promotion Program (LPP)".[9] The design of the LPP program was held constant across all neighborhoods that received

[9] LPP was designed after the "Community Led Sanitation Program (CLTS)" popular among sanitation policy professionals around the world, and studied by (Pickering et al. 2015) in Mali and (Gertler et al. 2015) in India and Indonesia. CLTS was invented by our NGO implementing partner in Bangladesh (VERC) before it was replicated in at least 60 other countries by governments and international NGOs and donors such as Plan International, World Bank and UNICEF (da Silva Wells and Sijbesma 2012). The LPP program informed residents of each neighborhood about the dangers of open defecation, and made the community-level problem salient by bringing all neighborhood residents together to discuss the issue. This, combined with the public nature of the sanitation lotteries we ran, made it obvious to each resident who else was receiving vouchers, and would be likely to adopt a new latrine.

subsidies, and the random variation was only in the extent of subsidy saturation.

Before introducing our structural model, we present some key model-free results from the RCT. Figure 3 presents adoption rates across different treatment arms in the experiment. 22% of households in villages where no subsidy was assigned own latrines that we classify as "hygienic" based on direct observation by enumerators, who were trained on criteria that a latrine must meet to be considered hygienic. The middle set of points show the average adoption rates for households that lived in subsidy neighborhoods, were eligible for the subsidy, but lost in the latrine lottery.[10] The three points correspond to the low, medium and high saturation communities. The rightmost set of points show average adoption among households that won the subsidy lottery, again separately by low, medium and high saturation community.

Two important results are apparent in the Figure 3. First, the price that the household faces is a key determinant of adoption: adoption rates among lottery winners are uniformly higher than among lottery losers or households in non-subsidy villages. Second, within both lottery winners and lottery losers, adoption increases as saturation increases. That is, holding the household's own price constant, voucher winners are more likely to convert their voucher into an actual latrine investment if the share of other households in the community receiving subsidies increases. Even lottery losers are more likely to invest in a full-price latrine when a larger share of the community receives subsidies. This suggests that latrine adoption decisions are strategic complements, and the complementarity dominates any psychological price-anchoring effect of not wanting to pay full price when neighbors receive subsidized access.

We also collected data on "*Any* Latrine Ownership" (as opposed to only hygienic), and on "Access to" latrines. Access differs from ownership in that a neighbor or relative may

[10]To keep the exposition simple, here we examine differences in adoption between winners and losers of the latrine subsidy lottery only. The outcome of the tin lottery (which, recall, was independent of the latrine subsidy lottery) is not included in this figure.

choose to share a latrine with you. This figure and the rest of the paper (conservatively) focuses on "ownership" rather than "access" in order to restrict our attention to behavioral demand spillovers, as opposed to sharing of a common resource. Effect sizes are larger for "Any" and "Access" variables. Latrine sharing is uncommon in this setting, except within extended-family compounds.

While these reduced-form RCT results provide strong evidence for price sensitivity and for social spillovers, on their own they would not allow us to address many important policy questions, such as the relative efficacy of widely and thinly dispersing a subsidy budget, versus concentrating on a few. This sort of policy analysis requires predictions about what will happen when the economic environment is different than the particular experimental outcomes in Figure 3, and how the relative magnitudes of the own price effect and indirect social spillover effect interact to drive individual decisions. That requires a model.

3 The Model

We model the utility that a household receives from building a latrine as depending on its own adoption decision as well as on the adoption decisions of other households in its neighborhood. To formalize this notion, we write:

$$U_i = U(a_i, a_{j \neq i}) \tag{1}$$

where U_i is the utility of household i and it depends on its own adoption decision, a_i, and the adoption decisions of other households in the neighborhood, $a_{j \neq i}$. We denote adoption by household i by $a_i = 1$. Latrine usage by other community members is a strategic complement if:

$$\left.\frac{\partial U_1(\cdot)}{\partial a_1}\right|_{a_{j \neq 1}=1} > \left.\frac{\partial U_1(\cdot)}{\partial a_1}\right|_{a_{j \neq 1}=0} \tag{2}$$

In words, the utility household 1 gets from adopting latrine usage is higher when other households are also adopting. We would expect latrine adoption to exhibit strategic complementarities within a neighborhood if social norms were an important determinant of latrine usage.

Conversely, latrine usage by other community members would be a strategic substitute if:

$$\left.\frac{\partial U_1(\cdot)}{\partial a_1}\right|_{a_{j\neq 1}=1} < \left.\frac{\partial U_1(\cdot)}{\partial a_1}\right|_{a_{j\neq 1}=0} \tag{3}$$

This might be the case if a household preferred to free-ride on the public health benefits of latrine usage by other households while practicing open defecation themselves. While the public health benefits of latrine adoption by others are obvious, it is an empirical question as to whether others' adoption decisions are a strategic complement or substitute.

We model the utility a household i residing in neighborhood c receives from adoption ($j = 1$) or not ($j = 0$) as:

$$U_{ijc} = f(z_{ic}, x_c, P_{ijc}, \bar{s}_c, \xi_c, \epsilon_{ijc}) \tag{4}$$

where:

z_{ic} is a vector of observable household-level attributes,

x_c is a vector of observable neighborhood-level attributes,

P_{ijc} is the price of a latrine j faced by household i in neighborhood c,

$\bar{s}_c$ is the share of households purchasing a latrine in neighborhood c,

ξ_c is a neighborhood-level unobservable component of utility; and

ϵ_{ijc} is a household-specific unobservable component of utility, assumed to have a Type 1 extreme value ("logit") distribution.

The utility of not buying a latrine, the outside good (pun intended), is normalized to zero.

The simplest implementation of (4) excludes household-level and neighborhood-level covariates and only includes (the log of) price, the share of neighbors adopting, and a neighborhood-level fixed effect. Utility in this stripped-down model is given by:

$$U_{ijc} = \alpha \ln(P_{ijc}) + \gamma \bar{s}_c + \xi_c + \epsilon_{ijc} \tag{5}$$

The adoption rate within each neighborhood c is given by:

$$\bar{s}_c = \frac{1}{N_c} \sum_{i \in c} \frac{\exp\left(\alpha \ln(P_{ijc}) + \gamma \bar{s}_c + \xi_c\right)}{1 + \exp\left(\cdot\right)} \tag{6}$$

where N_c is the number of households in neighborhood c. Equation (6) illustrates the interdependent nature of demand. The share of households adopting is a function of each household's decision and that household-level decision is itself a function of the share that adopt.

Computing the price semi-elasticity of demand, the change in the adoption share with respect to the percentage change in price, requires the Implicit Function Theorem. The semi-elasticity is given by:

$$\frac{\partial s}{\partial P/P} = \frac{\alpha s_c (1 - s_c)}{1 - \gamma s_c (1 - s_c)} \tag{7}$$

If social spillovers are positive ($\gamma > 0$), then elasticities calculated just using household-level variation in price will understate the true price elasticity. This can be seen by setting γ to zero in Equation (7). Note that this downward bias will exist even if the household price is perfectly random – intuitively, no matter how well-estimated α is, this tells us only about the household's response to its own price and not how the household may respond to the behavior of other households.

When we take (5) to the data, we explore three parameterizations. Our first, and simplest, specification is given by:

$$
\begin{aligned}
U_{ijc} =& \alpha_0 + \alpha_1 p_{ijc} + \delta_c + \epsilon_{ijc} \text{ where} \\
& \delta_c = \gamma_1 \bar{s}_c + \xi_c
\end{aligned}
\tag{8}
$$

In (8), utility is comprised of two parts– a household-level component and a neighborhood level component (δ_c), sometimes referred to as the mean utility. The observable part of the household-level component of utility depends only on $p_{ijc} \equiv \ln(P_{ijc})$, the (log) latrine price. The observable part of the neighborhood-level component of utility depends only on the share of households purchasing a latrine in the neighborhood.

Our second specification allows heterogeneity in the price-responsiveness of households.

$$
\begin{aligned}
U_{ijc} =& \alpha_0 + \alpha_1 p_{ijc} + \alpha_2 L_{ic} + \alpha_3 (p_{ijc} * L_{ic}) + \delta_c + \epsilon_{ijc} \text{ where} \\
& \delta_c = \gamma_1 \bar{s}_c + \xi_c
\end{aligned}
\tag{9}
$$

In (9), we add a household-level covariate, L_{ic}– an indicator variable for whether the household owns land ($L = 1$) or not ($L = 0$)[11] and we interact this covariate with the log of price. This allows the price responsiveness of landless households, which are typically poorer, to differ from that of landed households.

Finally, we introduce another neighborhood-level observable, a measure of the density of households in the neighborhood, D_c:

$$
\begin{aligned}
U_{ijc} =& \alpha_0 + \alpha_1 p_{ijc} + \alpha_2 L_{ic} + \alpha_3 (p_{ijc} * L_{ic}) + \delta_c + \epsilon_{ijc} \text{ where} \\
& \delta_c = \gamma_1 \bar{s}_c + \gamma_2 D_c + \gamma_3 (D_c * \bar{s}_c) + \xi_c
\end{aligned}
\tag{10}
$$

[11] Recall from Section 2 that landless means no land other than the homestead and, importantly, no agricultural land.

In (10), the utility of adopting a latrine varies at the neighborhood level by density and the strategic complementarity also varies with the density of households in a neighborhood.

Note that there are some implicit assumptions in the modeling of the strategic complementarity via $\bar{s}_c$ in all of the specifications above. First, by modeling the share adopting as a neighborhood-level variable, we are assuming that the share of a neighborhood's households is a sufficient statistic for identifying the strategic interaction, and that the identity of those households does not matter. When we explore mechanisms in Section 6, we will revisit this to the extent the data and experimental design allow. Second, this formulation abstracts from the sequencing of adoption decisions and assumes that a household either knows the fraction of its neighbors that adopt or has rational expectations over that share.

4 Estimation and Results

4.1 Estimating the Parameters of the Utility Function

We focus discussion on our simplest specification, (8), since the core issues are present in this basic setup. In (8), a household's utility from adopting a latrine depends on p_{ijc}, the price of the latrine, and $\bar{s}_c$, the share of the neighborhood that adopts. With observational data, both of these variables are likely to be endogenous. Price will typically be correlated with household-level or community-level unobservables (ϵ_{ijc} and ξ_c respectively) by inspection of the supply curve. In our RCT, though, the prices households faced were randomly assigned via the public lotteries so by construction the price a household faces is orthogonal to the unobserved terms in its utility function.

The average adoption rate in a neighborhood, $\bar{s}_c$, is both comprised of and in turn impacts (reflects upon) the household's own adoption decision. This is an application of Manski's reflection problem, and it results in the endogeneity of $\bar{s}$. Furthermore, unobserved components of the neighborhood are likely correlated with the adoption share of that neighborhood so $E(\xi_c, \bar{s}_c) \neq 0$. While we cannot control $\bar{s}_c$ experimentally, we do obtain exogenous variation by randomizing saturation.

Our estimation consists of two steps. The first step is a straightforward household-level binary logit in which the household's adopt / no adopt choice is regressed on the log of price the household faces, $\ln(p_{ijc})$, and a neighborhood-level fixed effect δ_c. As noted above, the latrine prices were fixed and subsidy offers were randomized, which obviates the need to engage in a search for instruments for price and then contend with the challenges that arise with instruments in a logit (MLE) framework. Step one then yields estimates of the coefficient on log price ($\hat{\alpha}_1$) and the neighborhood-level fixed effects ($\hat{\delta}_c$).

The second step of our estimator regresses the estimated fixed effect from step 1, $\hat{\delta}_c$, on the share of the neighborhood that adopts, $\bar{s}_c$. This neighborhood-level regression is a linear instrumental variables regression using the randomized subsidy saturation as our instrument for $\bar{s}_c$. Recall that neighborhoods were randomly assigned one of three different saturation levels: Low, Medium and High. We construct indicator variables reflecting the level of subsidy saturation for each neighborhood and these indicator variables are our instruments for $\bar{s}_c$. By design, the randomly allocated instruments are orthogonal to unobserved neighborhood attributes ξ_c. In the data, we show that the instruments are correlated with adoption share $\bar{s}_c$ and therefore relevant.

We highlight two econometrically based decisions that are implicit in even the simplest of our specifications. First, we have modeled the adoption share as a neighborhood-level variable and not as a household-level variable. If the reference group impacting a household's decision varied by household, as it conceptually could, we would have an instrumental variable in a logit framework and this is econometrically problematic.[12] By modeling the reference group at the neighborhood level, we avoid this problem in a non-contrived way. That is, there are

[12]There is no straightforward analogy in the logit MLE framework to the linear IV regression. See Berry (1994).

good economic reasons to think that the reference group is likely to be the neighborhood in which the household resides. The epidemiological basis for the strategic complementarity is likely neighborhood-based. Plausibly, the role of social norms and learning by observing are also neighborhood-level. Second, we do not include the baseline adoption share, $\bar{s}_0$, in the first step logit. Unlike price, the baseline share is not randomly assigned. If the reason the model does not fit well today is related to the reason it did not fit well last period, this serial correlation may well result in $E(\bar{s}_0, \epsilon_{ijc}) \neq 0$.

In (9), we introduce a household-level covariate into the first step – an indicator variable for whether the household is landless. We interact this variable with log price. The relative price-sensitivity of landless vs. landed households is potentially of interest to a policymaker considering whether to target subsidies on the basis of observable indicators of socio-economic status. From an estimation perspective, this is straightforward. Because price is randomly assigned, the interaction between price and landlessness is exogenous to the extent that landlessness is exogenous to latrine ownership. A shock to the utility of purchasing a latrine is likely exogenous to whether the household owns land.

In (10), we introduce a neighborhood-level covariate into the second step IV regression – the density of households in a neighborhood. Density is potentially relevant to policymakers, for several reasons. First, density is important epidemiologically and sanitation interventions may have greater health effects in dense areas (Hathi et al. 2017). Second, social influence may be more salient in denser areas. On the other hand, latrines are frequently shared, and sharing is easier in dense areas, so density may lead to "congestion" or a negative social spillover (Bayer and Timmins 2005; BenYishay et al. 2016). To compute the density of households in a neighborhood, we first use GPS data to calculate how many households in the neighborhood live within 50 meters of each household. The neighborhood density is then the neighborhood average number of households within 50 meters. While neighborhood density is likely pre-determined, the share of households adopting remains endogenous. Hence, we interact our instruments with our measure of neighborhood density.

4.2 Estimation Results

Estimation results are presented in Table 1. The top panel displays parameter estimates from the first step of the estimation algorithm. This step is a household-level binomial logit. We report all parameter estimates except for the vector of neighborhood fixed effects. The bottom panel gives parameter estimates for the second step of the estimation algorithm. This is a neighborhood-level linear instrumental regression in which the dependent variable is the neighborhood fixed effect and regressors are listed in the rows of the bottom panel. The last two lines of the table report the number of households in the first step and the number of neighborhoods in the second step.

The first column presents estimates of our simplest specification given in Equation (8). The only regressor in the first step is log price and the only regressor in the second step is the share of the the neighborhood purchasing a latrine. A higher price lowers the utility of adoption. The coefficient on share indicates that the strategic complementarity is positive – as more neighbors adopt, the utility of one's own adoption decision increases. In this specification, the price semi-elasticity of demand is given by Equation (7). Evaluated at the mean shares, its value is -0.22 implying that a 10 percent increase in price lowers the share purchasing by 2.2 percentage points.

The second column adds a covariate in the first step logit while the third column also interacts this covariate with price. The covariate is an indicator variable for whether the household is landless. This is a proxy for a household's wealth. We find that landless households receive less utility from adoption at the mean log price but are more price sensitive. The former reflects the lower likelihood of adoption by poor households and the latter is also intuitive.

In column 4, we add our measure of neighborhod density covariate to the second step IV regression. The coefficient on density is positive and precisely measured, indicating that the mean utility of adoption is greater in denser neighborhoods. Column 5 then interacts the density measure with the share adopting to allow for heterogeneity in the degree of strategic complementarity. While the positive point estimate on the interaction term suggests that strategic complementarities are greater in denser neighborhoods, this parameter is not precisely estimated.

We adopt the estimates in column 4 as our base case estimates. This specification allows for heterogeneity at the household-level and includes the covariate, density, at the neighborhood-level. We do not include the neighborhood-level interaction term in our base case because the large and imprecisely estimated coefficient on this term introduces too much noise into the adoption decision model.

5 Policy Analysis

The structural estimates from the model can be used to simulate various policy experiments. We start with a description of our policy simulation methodology and then present simulation results.

5.1 Methodology

The adoption share within each neighborhood c is given by:

$$s_c = \frac{1}{N_c} \sum_{i \in c} \frac{\exp\left(\alpha_0 + \alpha_1 p_{ijc} + \alpha_2 L_{ic} + \alpha_3 (p_{ijc} \times L_{ic}) + \gamma_1 \bar{s}_c + \gamma_2 D_c + \xi_c\right)}{1 + \exp\left(\cdot\right)} \tag{11}$$

When (11) is evaluated at the estimated values of the parameters and the vector ξ_c implied by the equation for mean utility,[13] this equation holds exactly within each neighborhood c. We write this as:

$$s^0 = \Lambda(x^0, z^0, p^0, s^0, \hat{\gamma}, \hat{\alpha}, \hat{\xi}) \tag{12}$$

where Λ is the logit function, the hats indicate estimated parameter values, and the "0" superscript indicates initial values of the data. Note that the initial adoption share in (12) is written as a function of itself.

When prices are perturbed, the likelihood that a household adopts will change. This changes the adoption share for the neighborhood. And with a new neighborhood adoption share, the likelihood that a household in that neighborhood will adopt again changes via the strategic complementarity. We solve for the new equilibrium using a contraction mapping. Denote the original prices and adoption share as (p^0, s^0). Our algorithm is:

1. Start at $s^0 = \Lambda(p^0, s^0, \cdot)$.
2. At new prices, p^1, compute $s^1 = \Lambda(p^1, s^0)$.
3. Compute $s^{(n+1)} = \Lambda(p^1, s^n)$.
4. Repeat until $[s^{(n+1)} - s^n] < \tau$, where τ is the tolerance for conversion.

The new equilibrium forms the basis for the simulation results.

Although we do not prove that our contraction mapping must always converge, extensive numerical experimentation has always resulted in convergence to a new equilibrium. this is true not just at our estimated parameters, but also for parameterizations of the utility function that place relatively much higher weight on the role of the share of neighbors adopting. The consistently observed convergence speaks to the existence of an equilibrium

[13] $\delta_c = \gamma_1 \bar{s}_c + \gamma_2 D_c + \xi_c$.

in our model. This is conceptually distinct from the issue of multiple equilibria. In our context, as in many models with peer effects, there may be multiple equilibria. For example, given the same fundamentals, if the peer effect is sufficiently large, it may be rational for a household both not to adopt when few of their neighbors adopt and to adopt if many of their neighbors adopt.

We use this procedure to conduct counterfactual policy simulations to answer the research questions that we highlighted at the outset:

1. In the presence of strategic complementarities in demand, should a given subsidy budget be widely distributed in small amounts to a large number of people, or distributed more intensively to fewer people? Note that the answer to this question depends on the relative magnitudes of the social spillover effect and the own-price elasticity of demand (which itself depends on the share that adopts). The answer is therefore not obvious without simulating the effects of different subsidy policies using the structural estimates.

2. Similarly, should we focus on a smaller number of communities and intensively subsidize, or should the subsidy budget be spread out in smaller increments to target a larger number of distinct communities?

3. Are there specific types of households that can be targeted with latrine subsidies, who will generate larger spillovers than others, and in the process increase the overall adoption rate per dollar of subsidy budget spent?

4. Are there specific types of communities that can be targeted where the demand spillovers travel more effectively across households?

We can answer the first two questions using our benchmark model without interactions. Answering the third question requires us to control for different household types, and interact those types with log(price) in the first step of structural estimation. Answering the last question requires us to control for variation in a community-level characteristic in the second step of our structural model.

5.2 Model Validation

We begin with a simple model validation exercise. In the villages that received subsidies, we observed in our data how households responded to those subsidies. We use our model to then predict how the control villages would respond to those same subsidies and, because villages were randomly selected for the treatment group, we can then compare the actual response to the subsidies to what the model predicts. We conduct this exercise with our most basic specification, (8), in which the household's decision depends only on the price they face and the share of their neighbors that adopt. This exercise asks how far this very simple model can get us in explaining observed responses to subsidies.

The validation exercise consists of three steps. First we estimate our simplest specification, (8), Second, using the structural parameters from this estimation, we simulate the effects of each level of subsidy saturation in the control villages. For example, we simulate, using the methodology explained in the previous subsection, randomly selecting 25 percent of the eligible households the subsidy for a latrine and giving a randomly selected 50 percent of the households the tin subsidy. We compute the new equilibrium for the control villages. Third, we compare the increase in adoption from the actual subsidies to the simulated increase. We do this for the low, medium, and high saturation levels.

In the control villages, the average adoption rate was 23.2%. In the data, the RCT resulted in latrine adoption rates of 32.4%, 38.6%, and 42.2%, respectively in the low, medium, and high saturation neighborhoods. When we simulate the subsidies in the control villages, the structural model predicts adoption rates of 30%, 34.9%, and 39.8%. We note three observations. First, this is a mostly, but not entirely, out-of-sample prediction exercise. It is mostly out-of-sample because we are simulating the subsidies in villages where none were given. It is not an entirely out-of-sample exercise because the control villages were used in the estimation of the structural model. Second, we have simulated the new equilibria with an extremely simple model in which only price and the share of neighbors adopting enter the household's utility function. Third, this stripped down model did remarkably well at explaining observed behavior. We turn next from model validation to prospective policy analysis.

5.3 Counterfactual Policy Simulations

5.3.1 Varying Subsidy Amount

Our simplest policy simulation explores the effects of varying the subsidy for the latrine. This is a counterfactual policy in the sense that we only experimentally varied the prices households faced at four points (the interaction of the latrine subsidy lottery and the tin or superstructure lottery). However, this variation, combined with the assumption that utility is log-linear in price, allows us to interpolate in our simulation.

In this exercise, we hold constant the fraction of households in each neighborhood offered a subsidy at 50% (the mid-point of our saturation experiments), and simulate the effect of subsidies of 2000 BDT, 3000 BDT and 4000 BDT.[14] Our parameters come from column 4 of Table 1, in which we have estimated Equation 10 including only the levels of density and adoption share in the second step, not their interaction. Figure 4 presents the results, where panels (a), (b) and (c) show the effect of each level of subsidy relative to the predicted level of adoption without any subsidy. In each panel, the "direct effect" (dashed line) shows the effect of the subsidy in the absence of any social multiplier. Mechanically, we change the price faced by each subsidized household and update their implied purchase probability using the estimated price coefficient $\hat{\alpha}$, but in this first step do not update $\bar{s}_c$. To arrive at the

[14] We chose these amounts for the counterfactual policy experiments, so that we are in line with the modal subsidy received by those who redeemed vouchers, which was 2880 BDT.

"total effect" (solid line), we then allow the social multiplier to take effect. Mechanically, we iterate as described in Section 5.1: (1) given the new probability of adoption, we update $\bar{s}_c$ to reflect the new probabilities of adoption; (2) given the updated $\bar{s}_c$, we update household utility using the social multiplier coefficient $\hat{\gamma}_1$; (3) this leads to newly updated household adoption probabilities; (4) we iterate (1)-(3) to convergence.[15]

Panels (a)-(c) of Figure 4 establish a few baseline results. First, unsurprisingly, larger subsidies lead to greater direct effects. Second, indirect effects also increase with the subsidy amount. Third, both effects are hump-shaped in the unsubsidized adoption share.[16] Panel (d), which plots outcomes neighborhood-by-neighborhood, shows that a large majority of neighborhoods are on the upward-sloping portion of this curve. Hence, in our discussion we place more emphasis on the upward-sloping segments of these curves since that is where the vast majority of the data lie.

5.3.2 Targeting Landless Households

Because we have estimated separate own-price elasticities of demand for landless and landed households, we can conduct a second set of counterfactual policy experiments in which we prioritize subsidizing one or the other of the two groups. A standard policy in the sanitation sector is to prioritize the poorest, which is sensible on equity grounds, but may not maximize health gains, especially if takeup is higher among less-poor households and if what matters for health is the overall use of sanitation in the community, rather than the identity of the specific households adopting sanitation (Geruso and Spears 2018).

To explore this tradeoff, we simulate a policy choice where the policymaker has a subsidy

[15] Note that the starting point for these policy simulations is itself the product of a simulation, in that we remove all subsidies from households that received them and then simulate adoption. That is, the initial price vector is the actual prices the households faced, based on their outcomes in the two subsidy lotteries. The new price vector strips subsidies from everyone, imposing a constant price for all households in each neighborhood. Since delivery and installation costs varied by neighborhood, we impose the neighborhood average from our data.

[16] Recall from footnote 15 that the x-axis represents simulated adoption without subsidies, not baseline levels, and that the simulated effect (the height of the curve) is relative to this unsubsidized level of adoption.

budget of 2000 BDT per household in a village and chooses whether to allocate it to landless or landed households. Figure 5 shows the results. In Panel (a), only landless households are subsidized. In Panel (b), only landed households are subsidized. Panel (c) compares the total effects of the two interventions. Panel (d) compares the total effects of the two interventions holding subsidy spending constant, showing that subsidizing landless households is more cost-effective – intuitively, landless households are more price-sensitive, so the subsidy has a larger direct effect on them, while for landed households, being relatively less price-sensitive means that social influence is relatively strong.

5.3.3 Subsidy Amount vs. Subsidy Saturation

Our third experiment considers the policymaker's tradeoff between the magnitude of a subsidy and the saturation level (specifically, the share of households subsidized). We consider first a simple choice among three policies: offering 1000 BDT subsidy to all households in the neighborhood; offering a 2000 BDT subsidy to 50% of households; offering a 4000 BDT subsidy to 25% of households. Figure 6a shows the results: offering a larger subsidy to a smaller share of households has the greatest effect on coverage. However, this simple comparison does not correctly capture the policymaker's tradeoff, since the higher takeup of subsidy vouchers implies a larger program budget. In Figure 6b, we adjust the subsidy amounts so that the resulting program budgets are approximately equal across saturations. The "naive" result of Figure 6a is reversed: the largest total effect, holding program spending fixed, comes from offering a relatively small subsidy to a large share of households. Intuitively, while the social multiplier (γ, the coefficient on $\bar{s}_c$) is important, it is small relative to the own-price effect (α, the coefficient on p).

5.3.4 Subsidy Amount vs. Scope of Program

Next, we consider the policymaker's tradeoff between the magnitude of the subsidy offered and the program's scope – concretely, the number of neighborhoods in which to conduct the subsidy intervention. In Figure 7a, we compare the total effect of three policies: offering a 4000 BDT subsidy in 25% of the neighborhoods; a 2000 BDT subsidy in 50% of the neighborhoods; a 1000 BDT subsidy in all neighborhoods. In all three cases, the subsidy is offered to all households in subsidy neighborhoods. As shown in Figure 7a, in this simple comparison, increasing the subsidy amount at a cost of reducing scope appears to have a greater effect.[17] However, when we adjust the subsidy amounts to equalize costs across the interventions, the result is reversed for most neighborhoods, as seen in Figure 7b.

5.3.5 Targeting on Neighborhood Observables: Density

Finally, the policymaker may target the program based on neighborhood-level observable characteristics. One natural such characteristic is neighborhood density, since it is (a) easy to observe and (b) health effects are plausibly larger in denser areas (Hathi et al. 2017). As shown in the parameter estimates in Table 1, density is positively associated with adoption, both in levels and when interacted with adoption share, although the latter is not statistically significant. In Figure 8, we compare the impacts of intervening in neighborhoods at approximately the 20th, 50th and 80th quantile of the distribution of density. Figure 8a shows results offering a 2000 BDT subsidy to 50% of households, by neighborhood; Figure 8b is similar, but with a 4000 BDT subsidy. As expected given our parameter estimates, targeting densely populated neighborhoods is more cost-effective at increasing coverage.

[17] The figure plots the average effect across *all* neighborhoods, not just subsidy neighborhoods. For example, the effect of the intervention offering a 4000 BDT in 25% of neighborhoods is four times greater *in the subsidy neighborhoods themselves* than indicated by the solid line.

6 Mechanisms

Our analysis thus far has focused on documenting the presence of strategic complementarities in sanitation demand, and exploring their implications for latrine subsidy policies. As noted at the outset, there are multiple channels through which sanitation investment decisions may become strategic complements. There may be an epidemiological link across households in the disease environment, or social norms about open defecation behavior may drive the complementarity, or it may be due to learning spillovers. In this section, we report the results of an additional sub-experiment within our subsidy treatment that was designed to shed light on the specific channels through which peer effects in adoption may operate. We also conduct further heterogeneity analysis to identify possible mechanisms underlying the complementarity.

6.1 Experiment Targeting Subsidies to "Highly Connected Households"

The first approach we employ to uncover mechanisms is to conduct an experiment in which we target subsidies to households that are considered "socially central" in a subset of neighborhoods. Prior to launching any field interventions, we first conducted a complete listing of all households in each village. We re-visited every household with that list of names, to ask them to identify up to four other households in their cluster with whom they interact with most frequently (i.e., members visit each other regularly), and also to identify up to four other households whom they would consult if they needed to resolve a dispute. After aggregating across all households' reports, we assign a "connectedness score" to every household in our sample, which is simply a count of the number of times that household was mentioned by *others* in their cluster in response to these two questions. All households within a cluster are ranked by their connectedness score.

We then randomly selected 123 of the 225 clusters where latrine subsidies were assigned, and bias the subsidy assignment in favor of households that score high on "connectedness". We refer to this sub-treatment as biasing the lottery in favor of "Highly Connected Households", or the *HCH treatment* for short. We did not bias the lottery in favor of HCHs in the other 102 clusters.[18]

The HCH treatment targeted the latrine subsidies toward socially central, connected households relative to the 102 non-HCH clusters. In designing these sub-treatments, our thinking was that if social influence (as opposed to pure technical or epidemiological complementarity about the disease environment) is the primary channel through which a strategic complementarity operates, then demand spillovers should be larger in the HCH treatment clusters. In other words, we should observe that the peer adoption rate has a larger effect on individual latrine purchase decisions in the subset of clusters where socially connected households were targeted with the initial subsidies. Note that such an HCH effect could operate either through a change in norms regarding sanitation behavior, or through learning about the costs and benefits of the sanitation technology from others.

Table 2 reports the second step of the structural estimates when we interact the HCH targeting experiment with the peer adoption rate in the cluster. The interaction term between $\bar{s}$ and HCH-targeting has a negative coefficient, but is statistically insignificant. In other words, targeting the subsidies to socially-central, putatively "influential" households within a neighborhood actually produces a slightly smaller complementarity. Having socially central people play the role of demonstrator is, if anything, *less* useful, surprisingly, to induce others to follow and adopt the new latrine technology.

To understand this surprising result, we present some reduced form experimental tests in Tables 3 and 4 to explore how different sub-groups of households react to HCHs vs. non-

[18]The specific mechanism we employed was to create a Pot 1 in which HCHs were given greater weight, and an identical-looking Pot 2 where they were not. The implementers asked a child from the neighborhood to first choose either Pot 1 or Pot 2, and then other children put their hands in the chosen pot to select the specific households who would receive latrine vouchers in a public lottery

HCHs receiving subsidies.Table 3 shows household-level OLS regressions of the decisions to adopt a hygienic latrine as a function of the voucher experiments (which determines the household's own-price) and the saturation experiments (which determine the average price in the community). Column 1 shows results for the subset of HCH-targeted clusters, and column 2 shows results for the complementary sample of clusters assigned to received subsidies that were not targeted to HCHs. The first three rows show that as expected, lottery outcomes (i.e, own price) matter a lot for latrine adoption decisions. The last two rows show that distributing more vouchers in the community (the source of the "peer effect") is only helpful in increasing individual adoption rates when the subsidies are *not* targeted to socially central, connected households. This is essentially the underlying source of the negative coefficient on $\bar{s}\times$HCH targeting that we observe in the structural estimates presented in Table 2.

Who, specifically, reacts by investing in latrines in clusters where HCHs are not targeted? Table 4 presents results when we split the sample up further by lottery outcomes and examine investment decisions for lucky latrine voucher winners separately from unlucky non-recipients. The only sub-group that becomes significantly more likely to invest when more of their neighbors receive subsidies are ones that resided in the non-HCH neighborhoods, and received vouchers themselves.

These reactions appear consistent with a model of shame in which households may find it shameful to continue practicing open defecation when socially marginal households in their neighborhood are moving away from OD into latrine use. In this view, social influence is more effective when households look down at the behavior of lower-status peers, not when they look up to high-status peers. Defecating in the open becomes acutely shameful when even lower-status peers are now using a more advanced toilet technology. There is not as much shame in continuing OD when higher-status, socially central households pay the expense to adopt the new technology.

To be sure, this concept of shame we are introducing represents ex-post theorizing on our part, after having seen these experimental results from the HCH-targeting. In the next subsection, we try to test and understand these mechanisms better, by exploring heterogeneity in spillover responses as the specific identities of households that receive subsidies changes.

6.2 Household Identities and Specific Lottery Outcomes

The household-level adoption data paired with the network data that we collected at baseline on inter-household connections (in order to design the HCH experiment) provides some opportunities for us to test these new hypotheses about shame and other mechanisms. Prior to launching interventions, we asked each household to name up to four other households in their neighborhood, who they would characterize as:

1. Households that have children that their own children play with,

2. Community leaders whom they would approach to resolve disputes

3. Households from whom they would seek advice about a new product or technology.[19]

The subsequent randomized allocation of latrine vouchers implies that for a specific household resident in a specific neighborhood, by chance, one of their "playmate contacts" may have gotten lucky and won a latrine voucher in the random lottery. For a different household, no playmate contact may have received such a voucher, again by pure chance. This creates household-level variation in our data on the random chance that any specific type of network connection for that household receives latrine vouchers. This allows us to create the following variables for every household in our sample:

1. Proportion of the household's "playmate contacts" (i.e. other households that have children that their own children play with) who won latrine vouchers,

[19] There was also a fourth type of connection listed: "Households that they interact with most frequently." However, we focus on the three types of network connections listed above because each of those is closely linked to a specific mechanism underlying strategic complementarity. In contrast, general interactions may be tied to multiple potential mechanisms, and therefore more difficult to interpret, and not as useful to us for identifying specific channels that underlie strategic complementarity.

2. Proportion of households who this household would approach to resolve disputes that won latrine vouchers, and
3. Proportion of households from whom they would seek advice about a new product or technology that received vouchers.

We calculate these variables as proportions rather than counts because some households may be more outgoing than others, and therefore may have more friends and contacts of all types. The share variables appropriately control for variation in each household's level of friendliness, and only vary based on the random lottery outcomes. Table 5 shows the results of controlling for these three variables in the first step of the structural estimation, and also interacting them with log(price).

The effect of households perceived as local leaders (who resolve community disputes) receiving subsidies on other household's adoption decision is informative about the shame theory we outlined in the previous sub-section. Do households look up to leaders in their technology adoption decisions, or are they more likely to look down towards more marginal members of society, in an urge to stay ahead of them? The negative coefficient on "Pct. Resolve Contacts who won Lottery" in column 1 suggests that leaders are not especially influential in inducing others to adopt new latrines. If the vouchers get allocated to leaders, others in the community are *less* likely to follow through with a purchase. The interaction term suggests that others also become a little less price sensitive (i.e., less reactive to subsidy offers), but this is not a statistically significant effect.

In contrast, the positive coefficient on "Pct. Technical Contacts who won Lottery" in column 2 suggests that a social learning channel may be more relevant. If the household that I rely on for technical advice wins a voucher by chance, then there is more for me to learn, and I become significantly more likely to invest in a latrine. Finally, the effect of "playmate contacts" receiving subsidies on each household's adoption decision may be informative about the epidemiological channel: My children's playmates' families receiving vouchers may change my own marginal return to adoption. My children are now exposed to a cleaner environment, and my own latrine investment now has a better chance of keeping my child healthy. We see in column 3 that the effect of this variable on adoption decisions is essentially nil. The epidemiological channel does not seem as relevant for producing a strategic complementarity in demand as the social learning channel or the shame factor, to the extent that cleanliness of children's play areas capture disease concerns in our research design.

Figure 9 shows the counterfactual policy simulations of targeting subsidies to these specific identities, based on these three regressions. Panel (c) indicates the magnitude of the adoption spillover effect of targeting subsidies to technically competent people (who others rely on for advice about new technologies) is quite large. In a village that starts out at about a 40% toilet usage rate, such targeting increases the community's adoption rate by 7 percentage points. Of course, such network positions of households may not be easily observed by policymakers, so this type of targeting may be difficult to implement. Panel (b) shows that targeting subsidies to leaders also produces a large effect, but in the opposite direction.

In summary, the household level variation provides some suggestive evidence that a social learning spillover is a potential mechanism underlying strategic complementarity. It also provides corroborative evidence that copying the behavior of community leaders is unlikely to be key channel that explains the complementarity that we have documented in this paper. This helps explain why the HCH-targeting experiment described in section 6.1 failed to produce greater adoption of latrines.

7 Conclusion

In this paper, we integrate a structural model of interdependent preferences with a randomized-controlled trial designed to identify the model. By randomizing both the price an individual household faces and the price environment of the household's peer group, we can empirically separate an own-price effect and a social multiplier or complementarity effect. We show that interdependent preferences are an important component of households' decisions, and that estimated price elasticities that do not account for this effect will be biased.

We use our estimated structural model to simulate the effects of counterfactual policies, both at the household level (e.g., targeting on the basis of socioeconomic status) and at the group level (e.g., targeting more densely populated communities). We show that targeting on the basis of landlessness or density increases the per-dollar impact of subsidies on sanitation coverage. We also consider designs of subsidy programs targeting various margins, such as saturation (share of households subsidized vs. amount of subsidy) or scope (number of villages subsidized vs. amount of subsidy). We show that concentrating subsidies does generate greater takeup among those subsidized, but widely dispersing subsidies produces larger spillover effects, so the latter is preferable for a policymaker facing a subsidy budget constraint.

Beyond what we learn about subsidies for sanitation adoption in particular or about subsidies in the presence of strategic complementarities in general, this paper also contributes to a growing literature combining theory with randomization to produce well-identified structural parameters. The estimated model can then be used to prospectively evaluate policies.

References

Akbarpour, M., S. Malladi, and A. Saberi, "Diffusion, Seeding, and the Value of Network Information," 2017.

Akram, A. A., M. Mobarak, and S. Chowdhury, "Effects of Migration on Rural Labor Markets," 2017.

Allcott, H., "Social norms and energy conservation," *Journal of public Economics*, 2011, *95* (9), 1082–1095.

Attanasio, O. P., C. Meghir, and A. Santiago, "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA," *The Review of Economic Studies*, 2012, *79* (1), 37–66.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson, "The diffusion of microfinance," *Science*, 2013, *341* (6144), 1236498.

Bayer, P. and C. Timmins, "On the Equilibrium Properties of Locational Sorting Models," *Journal of Urban Economics*, May 2005, *57* (3), 462–477.

Beaman, L., A. BenYishay, J. Magruder, and A. M. Mobarak, "Can network theory based targeting increase technology adoption?," *Unpublished manuscript*, 2018.

BenYishay, A., A. Fraker, R. Guiteras, G. Palloni, N. B. Shah, S. Shirrell, and P. Wang, "Microcredit and willingness to pay for environmental quality: Evidence from a randomized-controlled trial of finance for sanitation in rural Cambodia," *Journal of Environmental Economics and Management*, 2016, *86*, 121–140.

_ and A. M. Mobarak, "Social Learning and Incentives for Experimentation and Communication," *Review of Economic Studies*, 2018, *86*, to appear.

Berry, S., "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 1994, *25* (2), 242–262.

Bertrand, M., "New perspectives on gender," *Handbook of labor economics*, 2011, *4*, 1543–1590.

Bjorkegren, D., "The adoption of network goods: evidence from the spread of mobile phones in Rwanda," 2015. Working paper.

Bursztyn, L., F. Ederer, B. Ferman, and N. Yuchtman, "Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions," *Econometrica*, 2014, *82* (4), 1273–1301.

Cameron, L. and M. Shah, "Scaling Up Sanitation: Evidence from an RCT in Indonesia," *Unpublished manuscript*, 2017.

Chan, K. K. and S. Misra, "Characteristics of the opinion leader: A new dimension," *Journal of advertising*, 1990, *19* (3), 53–60.

Conley, T. G. and C. R. Udry, “Learning about a new technology: Pineapple in Ghana,” *The American Economic Review*, 2010, *100* (1), 35–69.

da Silva Wells, C. and C. Sijbesma, “Practical innovations for strengthening Community-Led Total Sanitation: selected experience from Asia,” *Development in Practice*, 2012, *22* (3), 417–426.

Duflo, E., R. Hanna, and S. P. Ryan, “Incentives Work: Getting Teachers to Come to School,” *American Economic Review*, June 2012, *102* (4), 1241–1278.

Gautam, S., “Quantifying Welfare Effects in the presence of Externalities: An Ex-ante Evaluation of a Sanitation Intervention,” 2017. Working paper.

_ , “Household (Under) Adoption of Sanitation: Externalities and Borrowing Constraints,” 2018. Working paper.

Gertler, P., M. Shah, M. L. Alzua, L. Cameron, S. Martinez, and S. Patil, “How does health promotion work? Evidence from the dirty business of eliminating open defecation,” Technical Report, National Bureau of Economic Research 2015.

Geruso, M. and D. Spears, “Neighborhood Sanitation and Infant Mortality,” *American Economic Journal: Applied Economics*, 2018.

Guiteras, R., J. Levinsohn, and M. Mobarak, “Encouraging Sanitation Investment in the Developing World: A Cluster-Randomized Trial,” *Science*, May 2015, *22*, 903–906.

Hathi, P., S. Haque, L. Pant, D. Coffey, and D. Spears, “Place and Child Health: The Interaction of Population Density and Sanitation in Developing Countries,” *Demography*, February 2017, *54* (1), 337–360.

Iversen, T. and F. M. Rosenbluth, *Women, work, and politics: The political economy of gender inequality*, Yale University Press, 2010.

J-PAL, P. B., “The Price is Wrong,” 2011.

Kessler, J. B., “Announcements of support and public good provision,” *American Economic Review*, 2013.

Kim, D. A., A. R. Hwong, D. Stafford, D. A. Hughes, A. J. O’Malley, J. H. Fowler, and N. A. Christakis, “Social network targeting to maximise population behaviour change: a cluster randomised controlled trial,” *The Lancet*, 2015, *386* (9989), 145–153.

Kinnan, C., “Distinguishing barriers to insurance in Thai villages,” 2017. Working Paper.

Kremer, M. and E. Miguel, “The illusion of sustainability,” *The Quarterly Journal of Economics*, 2007, *122* (3), 1007–1065.

_ , J. Leino, E. Miguel, and A. P. Zwane, “Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions,” *The Quarterly Journal of Economics*, January 2011, *126* (1), 145–205.

Manski, C., “Identification of Endogenous Social Effects: The Reflection Problem,” *Review of Economic Studies*, 1993, *60* (3), 531–542.

Mara, D., J. Lane, B. Scott, and D. Trouba, "Sanitation and health," *PLoS medicine*, 2010, *7* (11), e1000363.

Mas, A. and E. Moretti, "Peers at work," *The American economic review*, 2009, *99* (1), 112–145.

Meghir, C., A. Mobarak, C. Mommaerts, and M. Morten, "Migration and Consumption Insurance in Bangladesh," in "2015 Meeting Papers" number 190 Society for Economic Dynamics 2015.

Mobarak, A. M. and M. R. Rosenzweig, "Informal risk sharing, index insurance, and risk taking in developing countries," *The American Economic Review*, 2013, *103* (3), 375–380.

Munshi, K., "Networks in the modern economy: Mexican migrants in the US labor market," *The Quarterly Journal of Economics*, 2003, *118* (2), 549–599.

Oster, E. and R. Thornton, "Determinants of technology adoption: Peer effects in menstrual cup take-up," *Journal of the European Economic Association*, 2012, *10* (6), 1263–1293.

Paluck, E. L., H. Shepherd, and P. M. Aronow, "Changing climates of conflict: A social network experiment in 56 schools," *Proceedings of the National Academy of Sciences*, 2016, *113* (3), 566–571.

Pickering, A. J., H. Djebbari, C. Lopez, M. Coulibaly, and M. L. Alzua, "Effect of a community-led sanitation intervention on child diarrhoea and child growth in rural Mali: a cluster-randomised controlled trial," *The Lancet global health*, 2015, *3* (11), e701–e711.

Prüss-Ustün, A. et al., "Burden of disease from inadequate water, sanitation and hygiene in low-and middle-income settings: a retrospective analysis of data from 145 countries," *Tropical Medicine & International Health*, 2014, *19* (8), 894–905.

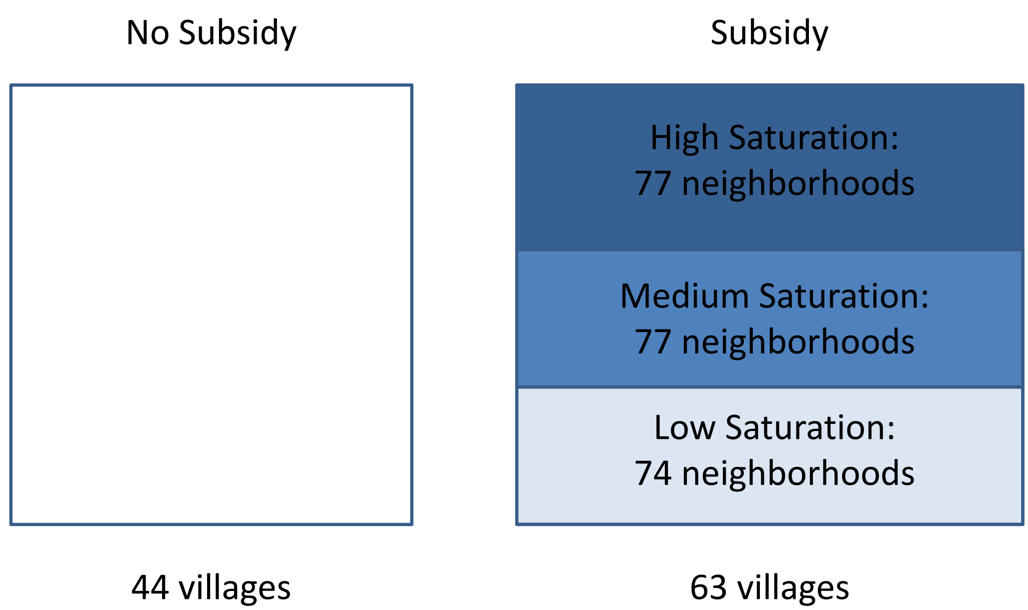
Spears, D., "Effects of rural sanitation on infant mortality and human capital: evidence from India's Total Sanitation Campaign," *Washington, DC: Princeton University*, 2012.

Todd, P. E. and K. I. Wolpin, "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, December 2006, *96* (5), 1384–1417.

WHO, Unicef et al., *Progress on drinking water and sanitation: 2014 update*, World Health Organization, 2014.

Yanagizawa-Drott, D., "Propaganda and Conflict: Evidence from the Rwandan genocide," *The Quarterly Journal of Economics*, 2014, *129* (4), 1947–1994.

Figure 1: Experimental Design - Village and Neighborhood Level Treatments



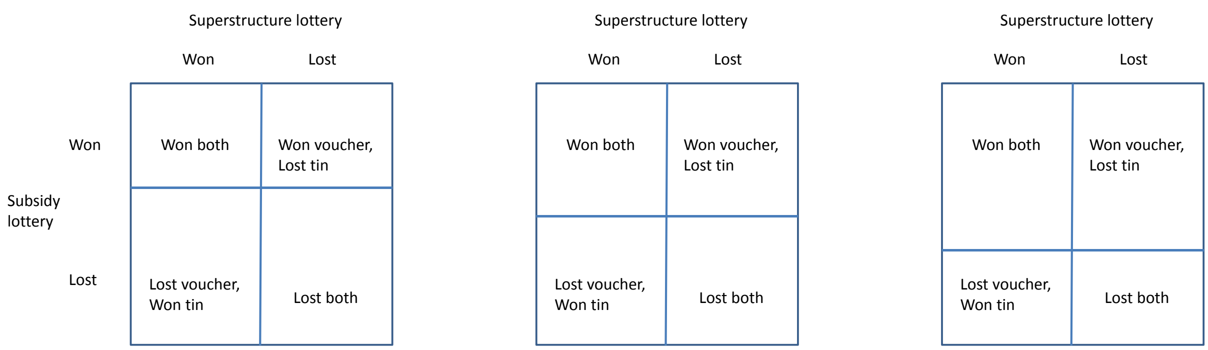
Notes: This figure shows two levels of randomization, one at the village level and another at the neighborhood level. Of the 107 villages in our sample, 63 villages were given subsidies (right), and 44 villages were not given subsidies (left). The second level of randomization (indicated by the three different shades of blue) is at the neighborhood level where 25%, 50%, or 75% of households received the subsidy (denoted as Low, Medium, and High saturation respectively). "No subsidy" includes Control, Supply Only and Latrine Promotion Program Only villages as described in the text.

Figure 2: Combining Household and Neighborhood Level Experimental Design

(a) Low Saturation Neighborhood

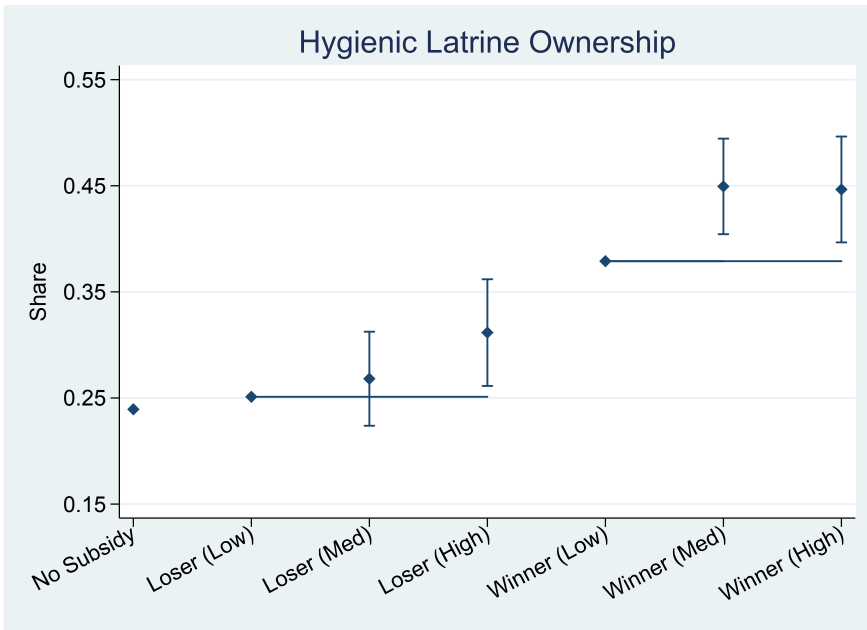
(b) Medium Saturation Neighborhood

(c) High Saturation Neighborhood



Notes: These figures show the randomized treatments assigned at the household level: independent lotteries for (1) a latrine subsidy voucher and (2) sheets of corrugated iron for build a roof for the latrine (tin lottery). The neighborhood-level saturation of the latrine subsidy is indicated by the positioning of the horizontal divider between winners and losers of the latrine subsidy voucher lottery. Approximately 25%, 50% and 75% of eligible households won in Low-, Medium- and High-saturation neighborhoods, respectively. The share of households winning in the tin lottery was constant at 50% in all Subsidy neighborhoods.

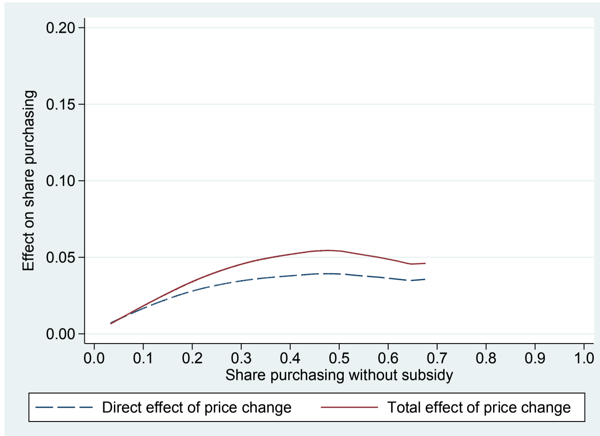
Figure 3: Raw Experimental Results on Demand Spillovers



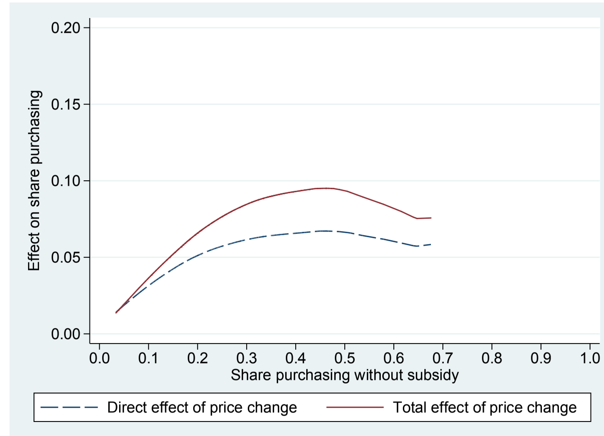
Notes: Adapted from Guiteras et al. (2015), Figure 2b. The top row of the x-axis refers to the neighborhood saturation level (Low, Medium, High). The bottom row of the x-axis refers to the household's outcome in the latrine subsidy voucher lottery. The left-most group, then, represents households in villages where no subsidies were offered. The middle group represents households in the Low, Medium, and High neighborhoods who did not win a latrine subsidy voucher. The right-most group are the households in the Low, Medium, and High saturation neighborhoods who did win. We display 95% confidence intervals for the strategic complementarity effect: adoption rates in Medium and High saturation neighborhoods relative to adoption rates in Low saturation neighborhoods.

Figure 4: Counterfactual policy: subsidize 50% of households, vary subsidy amount

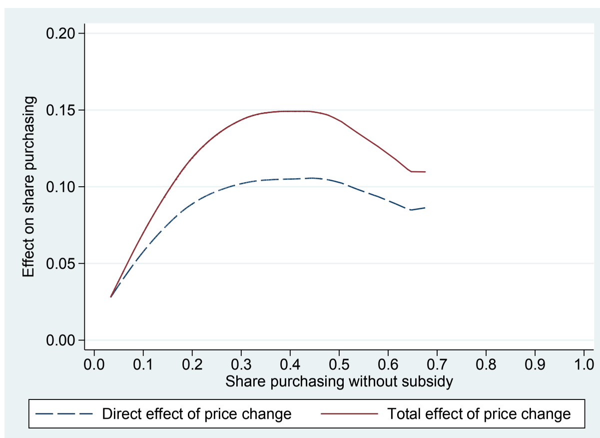
(a) 2000 BDT



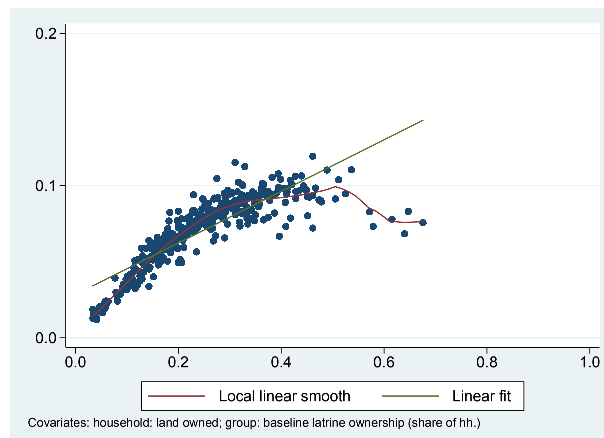
(b) 3000 BDT



(c) 4000 BDT



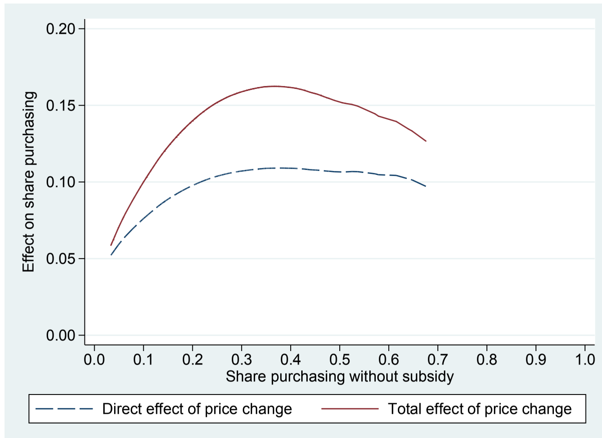
(d) 3000 BDT; Neighborhood-Level Outcomes



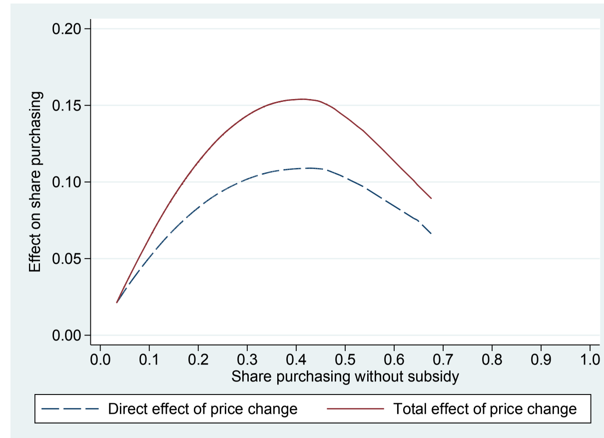
Notes: Panels (a)-(c) plot direct (dashed line) and total (solid line) effects of subsidizing 50% of households by neighborhood at the amount given (2000 BDT, 3000 BDT, 4000 BDT). Panel (d) repeats the exercise of Panel (b) but additionally plots outcomes neighborhood-by-neighborhood, showing that most neighborhoods are in the range where adoption is increasing. The x-axis represents simulated adoption without subsidies, and the height of the curve represents the simulated effect relative to this unsubsidized level of adoption.

Figure 5: Counterfactual policy: prioritize landless vs. landed households

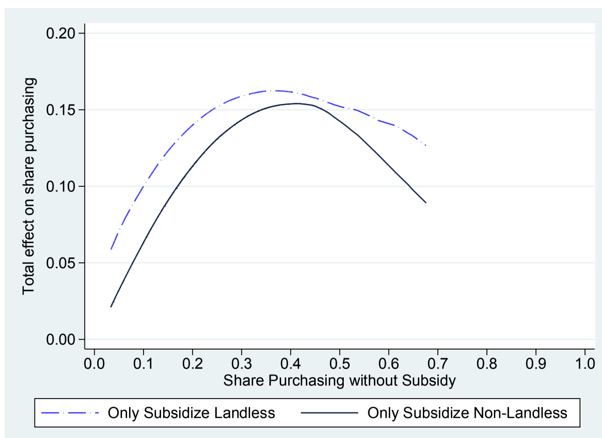
(a) Subsidize Landless Only



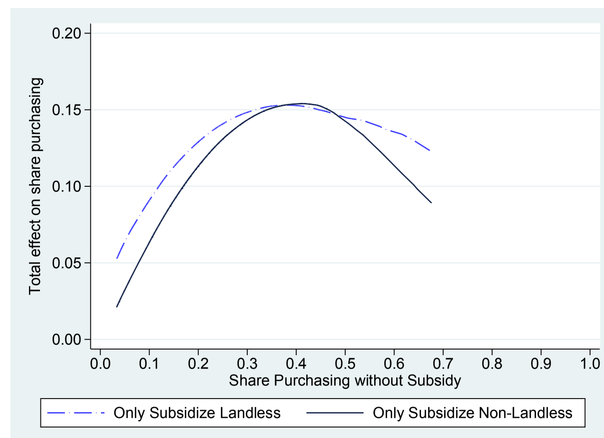
(b) Subsidize Landed Only



(c) Comparison



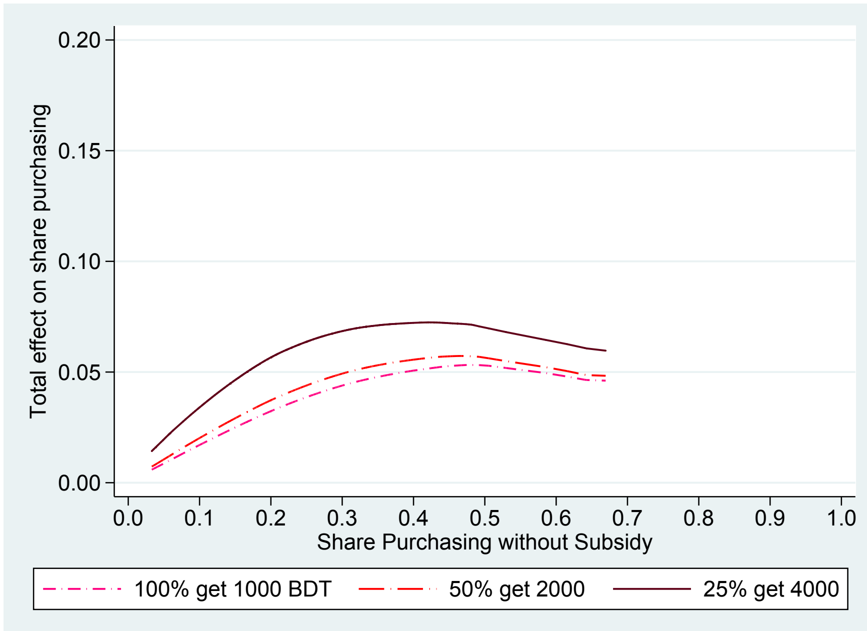
(d) Budget-neutral Comparison



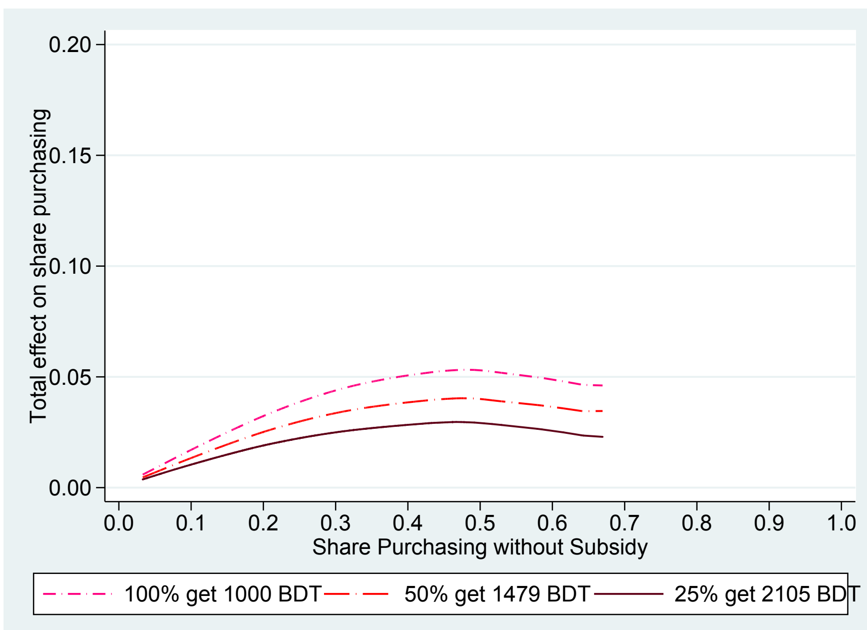
Notes: These figures show the effect of an intervention with a neighborhood budget of 2000 BDT times the number of households, including landed and landless, in the village. In Panel (a), only landless households are subsidized. In Panel (b) only landed households are subsidized. Panel (c) compares the total effects of the two interventions where the dashed line represents the intervention that targets the landless, while the solid line represents the intervention that targets the non-landless. Panel (d) compares the total effects of the two interventions holding subsidy spending constant.

Figure 6: Intensive (Price) Margin vs. Extensive (Saturation) Margin Within Neighborhoods

(a) Naïve experiment: Budget varies



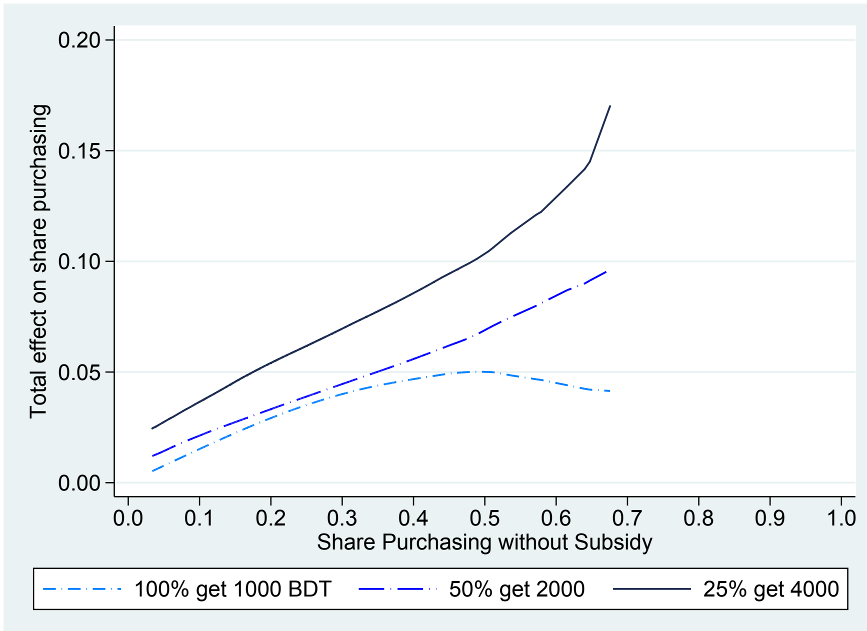
(b) Budget-neutral experiment



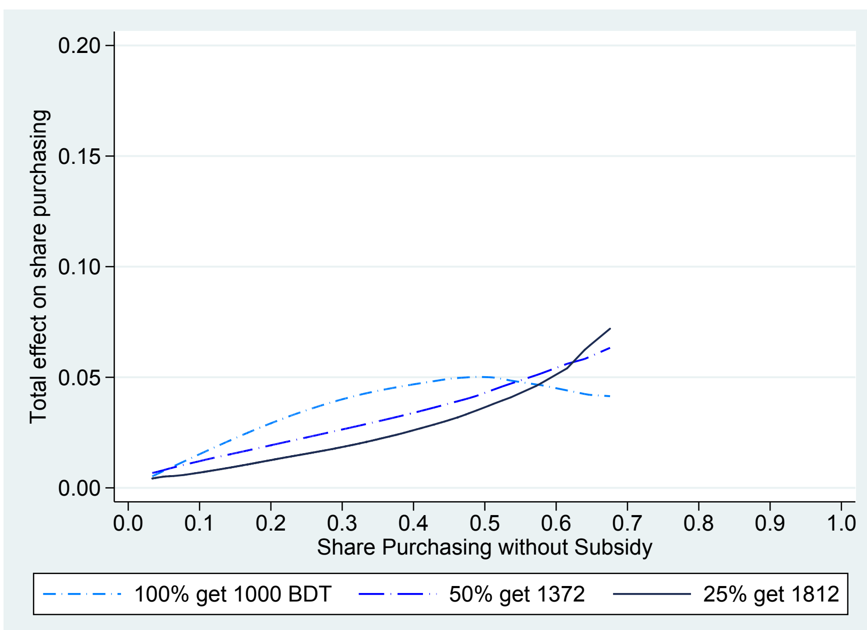
Notes: These graphs compare the total effects of interventions trading off subsidy amount against scope (the share of neighborhoods where subsidies are offered). The top panel (6a) compares: offering a 1000 BDT subsidy in 100% of neighborhoods (short-dashed line); 2000 BDT in 50% (long-dashed line); 4000 BDT subsidy in 25% (solid line). The bottom panel (6b) adjusts the subsidy amounts so that the cost per neighborhood is constant across the three interventions. In all cases, all households in subsidy neighborhoods are offered subsidies.

Figure 7: Intensive (Price) Margin vs. Extensive (Scope) Margin Across Neighborhoods

(a) Naïve experiment: Budget varies



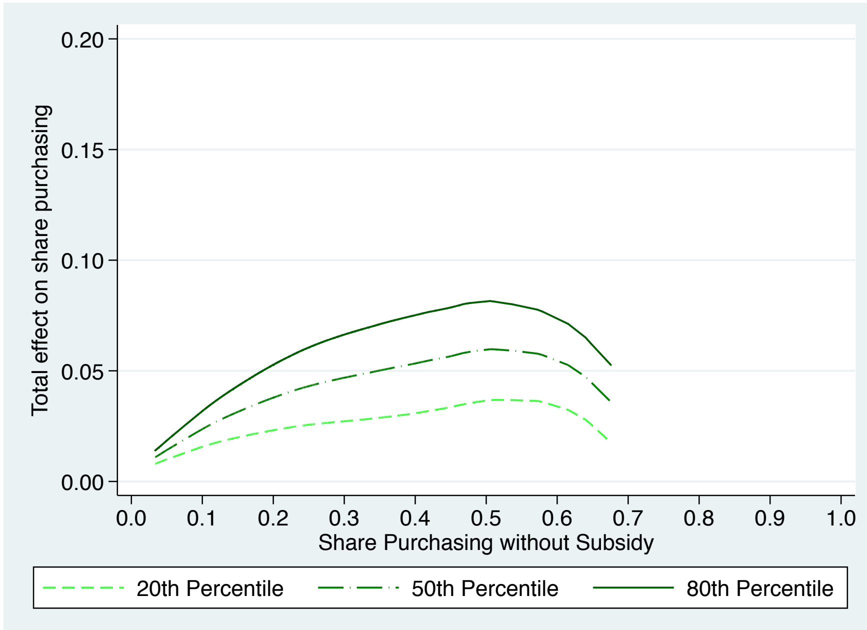
(b) Budget-neutral Experiment



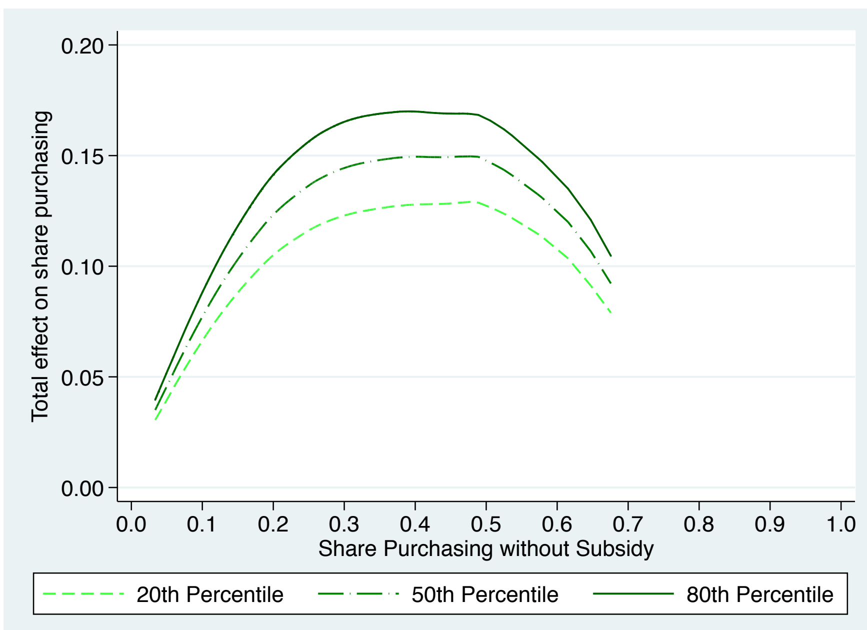
Notes: These graphs compare the total effects of interventions trading off subsidy amount against scope (the share of neighborhoods where subsidies are offered). The top panel (7a) compares: offering a 1000 BDT subsidy in 100% of neighborhoods (short-dashed line); 2000 BDT in 50% (long-dashed line); 4000 BDT subsidy in 25% (solid line). The bottom panel (7b) adjusts the subsidy amounts so that the cost per neighborhood is constant across the three interventions. In all cases, all households in subsidy neighborhoods are offered subsidies.

Figure 8: Targeting Densely Populated Neighborhoods

(a) Subsidy: 2000 BDT per household



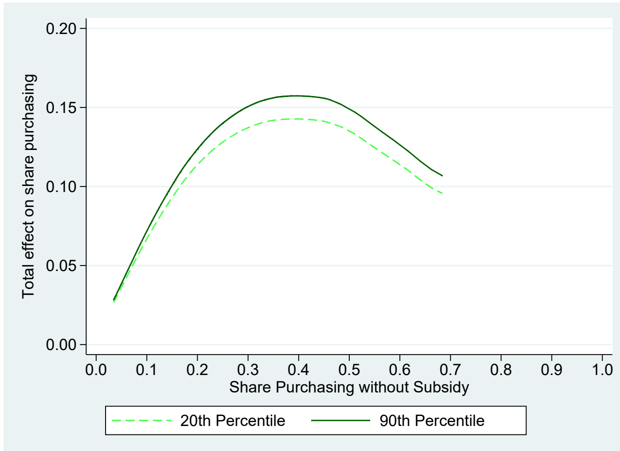
(b) Subsidy: 4000 BDT per household



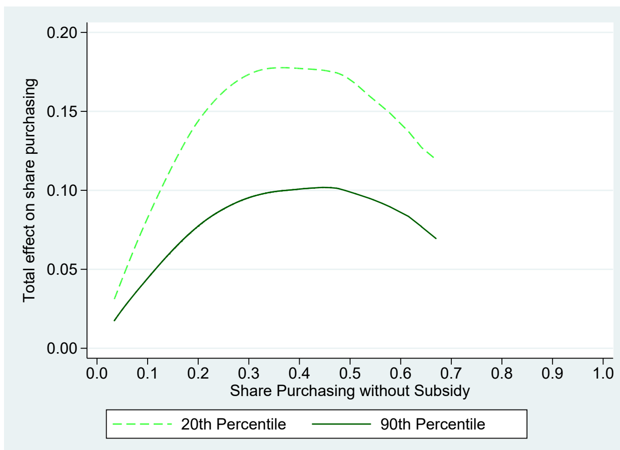
Notes: These graphs compare the effects of interventions targeting neighborhoods of the 20th quantile (short-dashed line), the 50th quantile (long-dashed line), and the 80th quantile (solid line) of the density distribution. The top panel (8a) shows results offering a 2000 BDT subsidy to 50% of households, by neighborhood; The bottom panel (8b) is similar, but with a 4000 BDT subsidy.

Figure 9: Targeting Subsidies to Households with Specific Network Positions

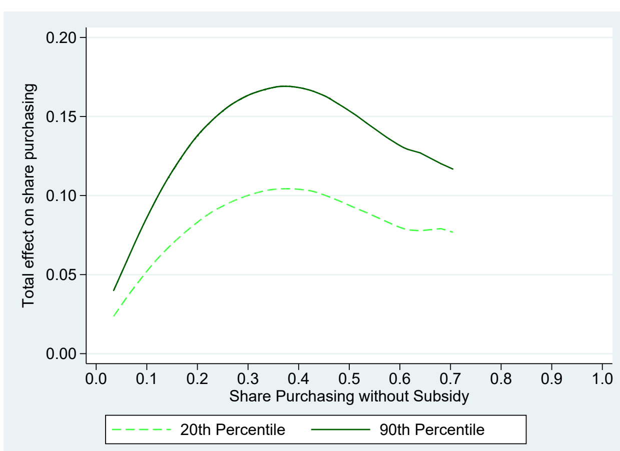
(a) Playmate Contacts, 4000 BDT



(b) Resolve Contacts, 4000 BDT



(c) Technical Contacts, 4000 BDT



Notes: These graphs compare the effects of interventions targeted at highly connected households. The two lines represent the 20th, and 90th percentile of the nearby-neighbor distribution – that is, the proportion of household's contacts who won latrine vouchers. The top panel (9a) shows the effects of targeting playmate contacts; the middle panel (9b) shows the effects of targeting conflict resolution contacts; and the bottom panel (9c) shows the effects when targeting technical contacts. The subsidy amount (4000 BDT) is constant across the three graphs.

Table 1: Parameter estimates

Household-Level Logit	(1) No Landless	(2) w/ Landless	(3) w/ Landless Int.	(4) w/ Density	(5) w/Density Int.
ln(Price)	-0.797***	-0.819***	-0.745***	-0.745***	-0.745***
	(0.062)	(0.062)	(0.072)	(0.072)	(0.072)
Landless		-0.821***	-0.835***	-0.835***	-0.835***
		(0.052)	(0.054)	(0.054)	(0.054)
Landless*ln(Price)			-0.180**	-0.180**	-0.180**
			(0.089)	(0.089)	(0.089)
Community-Level IV					
Share	1.242***	1.311***	1.299***	1.391***	1.276***
	(0.435)	(0.417)	(0.418)	(0.396)	(0.444)
ln(Density)				0.188*	0.204**
				(0.097)	(0.100)
ln(Density)*Share					0.882
					(1.094)
N (households)	12,792	12,792	12,792	12,792	12,792
N (neighborhoods)	369	369	369	369	369

Notes: The top panel reports the results of the first-step household-level logit estimates where the dependent variable is one if the household purchased a latrine. The regressors are as indicated, plus neighborhood fixed effects (not reported). The bottom panel reports the results of the neighborhood-level linear instrumental regression in which the dependent variable is the neighborhood fixed effect. Columns (3)-(5) are identical but correspond to different second-step IV regressions in the bottom panel. In the top panel, log price is centered so the coefficient on Landless in columns (3)-(5) represents the difference in utility between landless and landed at the mean log price of 8.26. The 11 neighborhoods where all households adopted are dropped in the logit estimation. Standard errors are clustered at the neighborhood level. * p<0.10, ** p<0.05, *** p<0.01.

Table 2: Second step neighborhood-level IV
Dependent variable: neighborhood-level fixed effect from first step

	(1) No Landless Int.	(2) (1) w/ Int.	(3) (2) w/ HCH
Share of peers adopting	1.307***	1.293***	1.373**
	(0.417)	(0.418)	(0.550)
Interact peer adoption & HCH			-0.103
			(0.824)
Randomized HCH binary			-0.013
			(0.062)
First-stage F stat.	33.6	33.6	17.0
Num. neighborhoods	368	368	368

Notes: The table reports results of the second stage IV estimates where the dependent variable is the village level fixed effect from the first stage logit. Columns (2) and (3) have first-step interactions of landlessness and price. (1) does not. (3) randomly assigns HCH=1 to half of non-subsidized neighborhoods pairwise by baseline adoption. Share of peers adopting recentered (mean zero). The mean share of peers adopting is 0.333. Standard errors are clustered at the neighborhood level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 3: Reduced-form household-level OLS
Dependent variable: Ownership of hygienic latrine

	(1) HCH Treatment Only	(2) Non-HCH Treatment Only
Only won latrine	0.111***	0.072***
	(0.025)	(0.024)
Only won tin	0.023	-0.021
	(0.028)	(0.028)
Won both	0.213***	0.172***
	(0.023)	(0.026)
Subsidy Med	0.048	0.058*
	(0.031)	(0.034)
Subsidy High	0.040	0.100**
	(0.032)	(0.040)
Mean ownership %, excluded group	24	32
Num. of neighborhoods	123	102
Num. of households	4,266	3,362

Notes: The table presents results from household-level OLS regressions where the samples are split into households in the HCH treatment only, column (1), and a sub-sample of households in the non-HCH treatment group, column (2). Regressors indicate whether the household won the latrine, tin, or both, as well as whether they were in medium intensity neighborhoods or high intensity neighborhoods. The excluded group are lottery-losers in low-subsidy paras. * $p<0.10$, ** $p<0.05$, *** $p<0.01$..

Table 4: Reduced-form household-level OLS
Dependent variable: Ownership of hygienic latrine
by latrine-lottery winning status

	(1) Won Lottery HCH Group	(2) Lost Lottery HCH Group	(3) Won Lottery Non-HCH Group	(4) Lost Lottery Non-HCH Group
Subsidy Med	0.052	0.038	0.105***	-0.010
	(0.040)	(0.032)	(0.040)	(0.039)
Subsidy High	0.036	0.034	0.127***	0.067
	(0.039)	(0.035)	(0.043)	(0.048)
Mean excluded ownership %	42	27	36	29
Num. of neighborhoods	123	123	102	102
Num. of households	2,760	1,506	2,130	1,232

Notes: The table presents results of household-level OLS regressions where the sample is split based on lottery outcome, and HCH group. Column (1) is the subsample of households that won the lottery, and are in the HCH group; column (2) are those who lost the lottery, and are in the HCH group; column (3) are households who won the lottery and are in the non-HCH group; and column (4) are households who lost the lottery, and are in the non-HCH gorup. The excluded group are households in low-subsidy paras. * p<0.10, ** p<0.05, *** p<0.01..

Table 5: First-step household-level fixed effects logit
Dependent variable: Ownership of hygienic latrine

	(1) Resolve	(2) Technical	(3) Playmate
Price (log)	-0.821***	-0.785***	-0.814***
	(0.075)	(0.091)	(0.077)
Price (log) × % of Resolve contacts who won lottery	0.090		
	(0.162)		
% of Resolve contacts who won lottery	-0.203*		
	(0.112)		
Price (log) × % of Technical contacts who won lottery		-0.016	
		(0.140)	
% of Technical contacts who won lottery		0.234**	
		(0.115)	
Price (log) × % of Playmate contacts who won lottery			0.064
			(0.142)
% of Playmate contacts who won lottery			0.052
			(0.092)
Num. of neighborhoods	369	369	369
Num. of households	12,824	12,824	12,824

Notes: The table reports the results of the first-step logit estimates where the dependent variable is one if the household owns a hygienic latrine. The regressors are the share of neighborhood contacts that randomly received a voucher. Column (1) includes the share of contacts the household would go to for conflict resolution; column (2) includes the share of contacts the household would go to for technical advice; and column (3) includes the share of contacts that have children that their own children play with. The price variable is recentered with the mean at zero; the true mean of log price is 8.26. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.