

Social Image and the Social Multiplier: Experimental Evidence from Community Deworming in Kenya*

Edward Jee[†]

Anne Karing

Karim Naguib

September 2025

Abstract

Moving services farther typically reduces take-up of preventive care. Yet when take-up is observable, requiring greater effort can raise the social image return to participation, dampening the distance penalty. We study this interaction in a community deworming campaign across 144 Kenyan communities. We randomly vary distance to treatment sites—close or far—and cross-randomize public signals that increase the observability of deworming. Reduced form results show that, without signals, moving from close to far lowers take-up by 16 percentage points; with signals, the close–far gap is 7 percentage points smaller. We estimate a structural model that separates the private cost and social image channels and accounts for the fact that, absent signals, observability falls with distance. The estimates imply that, as distance increases and participation falls, being seen deworming becomes more informative about motivation. With signals, social image returns rise with distance and partly offset private costs, so take-up declines only 80–90% as much as it would if social image played no role. Without signals, declining observability reduces social image returns and take-up declines 10–30% more. These dynamics matter for optimal site placement: a policymaker who accounts for social image can achieve the same deworming coverage with 6% fewer sites when actions are highly observable.

Keywords: social image, social multiplier, signaling, observability, travel distance

JEL codes: C93, D82, I12, I18, O12

*We thank Ned Augenblick, Stefano DellaVigna, Rachel Glennerster, Edward Miguel, and seminar participants at Berkeley, Bocconi, Chicago, Georgetown, Harvard, MIT, NYU, Pompeu Fabra, Stanford and the NBER Summer Institute for their valuable comments and discussions. Arthur Baker provided outstanding research assistance in Kenya. The field experiment would not have been possible without the invaluable support and collaboration of the Ministry of Health of Kenya, Evidence Action, and REMIT. Funding was provided by the Children’s Investment Fund Foundation. The experiment and data collection were approved by the UC Berkeley IRB and the Kenya Medical Research Institute.

[†]Jee: University of Chicago (edjee@uchicago.edu); Karing: University of Chicago (akaring@uchicago.edu); Naguib: AstraZeneca (karimn2.0@gmail.com)

1 Introduction

Governments routinely steer behavior by altering private costs. Yet behavior is also shaped by social image concerns. When participation is observable, shifts in private costs can change what actions signal to others, attenuating or amplifying behavioral responses—a social multiplier in the sense of [Benabou and Tirole \(2025\)](#).

Understanding these forces in real-world settings is essential for policy design. Consider moving a community health campaign closer to a village: lower travel costs draw in people who were previously deterred by distance. If the new participants are perceived as less motivated, the reputational return to attendance may fall because observers infer that ‘it must not take much dedication to attend now’. At the other extreme, when services are placed very near—for example, a blood drive in one’s workplace—participation may become nearly universal, and abstention may carry sharper stigma. In both cases, distance shifts equilibrium participation and therefore the inferences observers draw from take-up. Unlike work on crowding out from monetary incentives (e.g., [Ariely et al. 2009](#)), we study how non-monetary cost changes shift who participates, thereby changing what participation signals.¹

Despite the intuitive appeal of such multipliers, direct evidence on their magnitude remains scarce. This paper provides causal, field-based evidence on how economic incentives and social image interact, and quantifies the implied multiplier for a large-scale program. We ask: How do travel distance and social image concerns, operating through observability, jointly determine equilibrium take-up in a preventive health campaign? What is the size of the social multiplier, and what does it imply for allocating public resources across space?

We collaborate with the Government of Kenya on a community deworming program that targeted roughly 200,000 adults. Treatment is free and offered at centrally located sites. Deworming is well known and socially valued: 68 percent of adults say they would look down on someone who does not attend, and 93 percent say they would praise someone who does. Yet, as with many preventive health behaviors, take-up is modest even though treatment can be purchased at clinics and pharmacies for a nominal fee. Because deworming reduces disease transmission, under-adoption imposes social costs.

A central logistical choice for any policymaker is how many treatment sites to open and how far apart to place them. Distance affects private travel costs, but it can also affect social image payoffs. We adapt [Benabou and Tirole \(2025\)](#)’s concept of a social multiplier

¹In [Ariely et al. \(2009\)](#), monetary incentives can “crowd out” social image returns by creating ambiguity over motives: did someone act to do good or to do well? For example, a generous tax credit for electric vehicles may lead observers to infer financial rather than environmental motives. Here, the signal itself remains clear; what changes is its interpretation because the composition of participants changes. A salient EV subsidy could attract less environmentally motivated adopters, diluting the composition. Conversely, when participation becomes nearly universal, abstention may carry greater stigma, potentially amplifying the policy’s impact beyond a standard private cost–benefit prediction.

to study how distance and observability jointly influence behavior, focusing on a core theoretical insight: an action’s cost affects both selection into the action and observers’ beliefs about those who act. Exogenous shifts in cost that are easily understood and common knowledge are required to measure this feedback loop. Distance is well suited: when a treatment site is farther, attending is visibly more effortful if the act can be observed. If effort is less observable, the social image return may also fall.

The program’s geographic scope allows us to randomize both distance and observability at the community level in a factorial design. Operational constraints—sites are hosted at schools and only a limited number can open—mean not all communities can be near a site. First, we randomly assign 144 communities to be Close (0–1.25 km) or Far (1.25–2.5 km) from their assigned treatment site, inducing an average 1 km difference in travel distance. Second, we vary observability with low-cost public signals provided upon deworming: a colorful bracelet or an indelible ink mark. The bracelet is highly visible and novel, whereas the ink mark is familiar from elections and nearly costless to implement. Absent such signals, observability may co-move with distance, so distance effects alone cannot separate private travel-cost effects from responses driven by social image concerns. To control for any private consumption value of the bracelet, we include a Calendar arm that provides a low-visibility wall calendar upon deworming. We also include a pure Control arm where no item is given. Comparing Bracelet to Calendar isolates signaling value, while the Calendar arm also detects any distance-related heterogeneity in responses to material incentives. This design separates signaling from explanations based solely on private rewards.

Signals might also remind people about deworming or convey information about peers’ take-up. To probe these channels, we send SMS messages to a random subsample of 2,635 adults, directly reminding them of the campaign and reporting the current share of local take-up. Because we randomize messages across signaling and Control communities, we can test whether reminders or social learning—rather than signaling—drive the effects.

We begin by documenting how distance and incentives affect the observability of deworming decisions. Deworming is highly observable in campaign settings: in Control communities, respondents report knowing 69 percent of peers’ deworming status. Observability falls by 12 percentage points ($p = 0.056$) when comparing Close and Far Control communities, consistent with adults learning about others’ actions mainly through direct encounters or by seeing them en route to treatment. As distance increases, such observation becomes more difficult. Introducing signals reverses this decline: in Far communities, knowledge of peers’ status rises by 22 percentage points with the bracelet signal ($p < 0.001$) and by 15 percentage points with ink ($p = 0.007$). In Close communities, the same signals raise observability by only 3–5 percentage points ($p = 0.357$ – 0.644). This pattern is consistent with the idea that signals that are costlier to achieve are more informative about motivation, giving individuals a stronger incentive to display them and

thereby increasing observability.

We then measure effects on actual take-up by directly monitoring 9,805 adults at treatment sites. Three reduced form findings emerge. First, absent signals, a 1 km increase in distance lowers take-up by 16 percentage points ($p < 0.001$). Surveyed beliefs track this pattern: participants predict a 12 percentage point decline when distance doubles ($p = 0.002$). Second, adults value the opportunity to signal with a bracelet: take-up rises by 7.5 percentage points relative to Control ($p = 0.008$), from 33 to roughly 41 percent. Ink has no discernible effect and yields an imprecisely estimated -3 percentage points ($p = 0.35$), plausibly reflecting messiness or its association with election monitoring. Take-up in Bracelet communities exceeds that in Calendar communities by 5 percentage points ($p = 0.048$), indicating that the bracelet's signaling value, rather than its consumption value, drives behavior. SMS messages leave these patterns unchanged, ruling out reminders or information as stand-alone drivers. Third, pooling Bracelet and Ink, the signal effect—relative to Calendar and Control—is nearly twice as large in Far as in Close communities (6.8 percentage points, $p = 0.083$). Additional evidence favors a social image mechanism over a scarcity-only story. In Control, a willingness-to-pay elicitation indicates that calendars are privately valued more than bracelets. Among non-dewormers in the Bracelet and Calendar arms, gift choices at endline reveal a prevalence penalty: participants avoid the locally common item in Calendar and Bracelet–Close communities. In Bracelet–Far, the share choosing a bracelet is statistically indistinguishable from that in Control, inconsistent with the prediction that the bracelet's private value should rise purely because fewer people have it.

Reduced form comparisons by distance cannot identify the multiplier: because deworming is observable even in Control, distance effects conflate private travel costs with distance-induced shifts in social image returns. We therefore develop and estimate a structural model, building on [Benabou and Tirole \(2025\)](#). Individuals choose whether to deworm to maximize utility with (i) private benefits net of travel cost and (ii) a social image payoff that depends on observability and observers' inferences about motivation. The model generates an endogenous multiplier: distance shifts both the composition of participants and the observability of their actions, thereby altering social image returns. The multiplier can mitigate or amplify the direct cost effect depending on how observability and inference covary with distance.

Identification rests on three randomized sources. First, the Control-arm WTP exercise identifies the difference in private consumption value between the bracelet and the calendar. Second, the cross-randomization of distance and observability, together with survey data on knowledge of others' deworming status, identifies levels of observability and their slope with respect to distance. Third, the distance-by-observability gradient in take-up identifies responsiveness to travel costs in each arm; combined with the first two ingredients, this pins down the social image return to being observed deworming and thus

the multiplier. Recovering the multiplier requires experimentation at scale and across multiple distances: single-distance pilots confound private costs with social image, miss distance-induced changes in observability and equilibrium inference, and cannot recover the sign or magnitude.

The estimated model reproduces the reduced form patterns and separates the direct effects of incentives and distance from equilibrium changes in social image returns that operate through type inference and observability. We formally define the social multiplier as the ratio of the distance effect on take-up when participation is observable to the distance effect when it is not (the private-cost effect). With signals, the multiplier lies between 0.8 and 0.9 (mitigation); without signals, between 1.1 and 1.3 (amplification). If participation were unobservable or social image concerns were absent, the multiplier would equal one. Mechanically, greater distance lowers take-up by raising travel costs. With signals, observability remains high as distance increases and participants become more selected, so being observed deworming is a stronger signal of health conscientiousness; the social image payoff therefore rises with distance and partly offsets travel costs. Without signals, observability falls with distance faster than inference strengthens, so the social image payoff declines, reinforcing the private cost effect.

In the final part of the paper, we combine the structural estimates with geographic data on 144 study communities and 1,451 feasible sites to solve a site-allocation problem: choose the minimal number and optimal locations of treatment sites needed to achieve a target deworming rate. We contrast two scenarios. A policymaker who calibrates demand on short-range data and treats image payoffs as distance-invariant would fund 100 sites when observability declines with distance and 96 when observability remains high. A policymaker who incorporates the multiplier—recognizing that distance shifts both the composition of participants and observability—would allocate 107 sites in the first scenario but only 90 in the second, a 6–7 percent change relative to the distance-invariant benchmark. This exercise demonstrates why separating observability from equilibrium beliefs matters: low take-up at distance can reflect low private value relative to travel cost, or low observability that depresses image payoffs. The remedies differ. When low private value is the binding constraint, moving sites closer is appropriate. When observability is low, raising observability at a given distance increases image payoffs and take-up without adding sites.

This paper, to our knowledge, provides the first causal field estimate of a social multiplier generated by social image concerns. By cross-randomizing distance and observability and leveraging beliefs and willingness-to-pay data, we separately identify (i) the private travel cost channel, (ii) observability, and (iii) belief-based equilibrium inference, and use these to estimate the multiplier in the sense of [Benabou and Tirole \(2025\)](#). Existing field studies largely hold costs fixed and capture static image effects ([Bursztyn and Jensen 2017](#); [Bursztyn et al. 2018](#); [Chandrasekhar et al. 2018](#); [DellaVigna et al. 2017](#); [Karing](#)

2024; Breza and Chandrasekhar 2019). Related observational work examines how policies interact with norms (e.g., Jia and Persson 2019; Besley et al. 2023) but cannot disentangle material incentives, observability, and beliefs. Our experiment isolates all three, quantifies the multiplier, and shows that a low-cost visibility lever can flip the multiplier from amplification to mitigation, with direct implications for spatial design. We also build on a literature emphasizing that social interaction effects are endogenous rather than fixed, beginning with Glaeser et al. (2003), and provide experimental evidence on how observability that policy can change affects them.

A second contribution is to work on contextual factors that shape social image and public good provision (Kessler 2017; Perez-Truglia and Cruces 2017; Karing 2024). We complement Ariely et al. (2009), who show that image concerns interact with monetary incentives, by varying a non-monetary cost—distance. This offers two advantages. First, it shifts the composition of participants without confounding reputational inferences with financial payoffs—a distinct mechanism central to crowd-out theories. Second, because distance varies continuously, we can trace how marginal cost changes influence both participation and the image return to participation, enabling a quantitative estimate of the multiplier. This distinction is not only conceptual but also policy relevant: governments rarely pay people to adopt preventive health behaviors, but they routinely decide where to place services.

Finally, we contribute to a nascent literature that blends field experiments with structural modeling to inform counterfactual design (Todd and Wolpin 2006; Kaboski and Townsend 2011; Duflo et al. 2012; Lagakos et al. 2023). We integrate structural estimates with geocoded data on 144 communities and 1,451 feasible sites to compute policy counterfactuals, solving a siting problem subject to a coverage target. This addresses a methodological question—how to account for social image in policy design—and shows that doing so can alter optimal placement, allowing programs to achieve the same coverage with fewer sites.

The remainder of the paper proceeds as follows. Section 2 describes the empirical setting. Section 3 details the experimental design, and Section 4 reports reduced form results. Section 5 develops the theoretical framework and derives predictions and Section 6 estimates the model. Section 7 uses the estimates to solve the policymaker’s siting problem. We preregistered a pre-analysis plan prior to data collection (AEARCTR-0001643); Section 3.7 summarizes the plan and the post-PAP extensions, and Appendix Table E1 maps preregistered items to reported analyses.

2 The Setting

Intestinal worms impose a substantial health burden on children and adults in many low-income countries. The World Health Organization estimates that roughly 24 per-

cent of the world’s population is infected with soil-transmitted helminths (2023). Severe infections can cause abdominal pain, iron deficiency, anemia, malnutrition, and stunting (Hotez et al. 2008). School-based deworming programs have sharply reduced worm prevalence among children (Miguel and Kremer 2004), yet adults remain an important reservoir that sustains reinfection. While the social benefits of community-wide deworming are potentially high, the private health gains for adults are modest (Chan 1997; Ahuja et al. 2015).

Community deworming offers a useful setting to study how economic incentives interact with social image concerns in health. First, deworming is already an established behavior. Since 2009, the Kenyan National School-Based Deworming Program has treated more than 5 million children in high-endemic areas—including our study counties (Evidence Action 2017). In our baseline survey, 78 percent of adults were aware of deworming treatment and 49 percent knew it should be taken every 3–12 months (Table B1, Panel C). When asked who is at risk of worm infection, 94 percent cited children and 70 percent named adults as well.

Second, deworming is subject to strong social judgment. At baseline, 95 percent of adults said they would praise someone who took the free treatment, whereas 70 percent said they would look down on someone who did not (Table B1, Panel D). Figure A1 shows that these perceptions are comparable to attitudes toward childhood immunization and the avoidance of open defecation. Qualitative endline responses demonstrate that adults view deworming as a signal of desirable character traits, stating that deworming “shows I care about my health” or that “I am responsible when it comes to health,” linking compliance to personal responsibility. Others emphasize civic virtue and leadership—e.g., “I must take it as a leader to set an example to others,” “They know I am always vocal in community projects,” and “She knows I am all about health in the community.” Such remarks indicate that deworming publicly signals concern for both one’s own well-being and that of the community. Fewer than 2 percent worry that taking the pills might signal dirtiness or illness, so social image payoff is overwhelmingly positive (Table B1, Panel C). These social judgments are not contingent on knowledge of transmission: the shares who say they would praise someone who deworms and look down on someone who abstains are statistically indistinguishable across communities with above- and below-median externality knowledge and along a continuous knowledge index of externality knowledge (Table B2). Only 37 percent of adults understand that worms can be transmitted between household members or neighbors. Accordingly, throughout the paper we interpret the social image motive as signaling health conscientiousness and personal responsibility rather than prosociality.²

Third, adults underinvest in deworming despite low private cost. A full course is available in pharmacies for roughly US\$0.50–2.00, yet only 38 percent of adults report

²The precise interpretation is not critical, as in both cases the behavior yields a social image return.

deworming in the previous year, although 69 percent have done so at least once (Table B1, Panel C). In high-prevalence areas adults are targeted through mass-drug-administration campaigns or door-to-door visits by Community Health Volunteers (CHVs), but no permanent program provides free treatment.³

Community Deworming Program. In partnership with the Kenyan Ministry of Health, we implemented a new community deworming campaign that offered free treatment to approximately 200,000 adults in Busia, Siaya, and Kakamega counties, where helminths are endemic. The campaign operated for 12 consecutive days, 8 a.m.–5 p.m., at central locations staffed by trained CHVs.⁴ This large-scale implementation created the logistical scope to randomize both travel distance and the observability of participation, forming the core of our experimental design described in Section 3.

3 Experimental Design

We cross-randomize, at the community level, both (i) the travel distance to the nearest point of treatment (PoT) and (ii) a set of incentives that alter how observable the act of deworming is to others. This design isolates how changes in private cost of deworming interact with social image returns—defined as the utility difference that arises from how others perceive an individual who does versus does not deworm. Public signals (bracelet, ink) provide exogenous variation in observability that helps identify the social image component. Travel distance shifts time cost and thus the private incentive to participate. Greater distance depresses take-up mechanically, but it also changes who is willing to incur the cost. If only highly health-conscious or civic-minded adults are willing to travel, deworming becomes a stronger positive signal, while non-participation becomes less informative about motivation. Depending on the relative strength of these two forces, social image returns may rise or fall, thereby amplifying or dampening the direct effect of distance on take-up.

3.1 Distance Treatment

Within each county, we randomly assigned 144 communities to either a *Close* or *Far* travel condition. In the Close condition, the centroid of the community lay 0–1.25 km from the designated point of treatment (PoT); in the Far condition, it lay 1.25–2.5 km away. Random assignment produced an average distance difference of 1.02 km and doubled the median walking time from 15 to 30 minutes (Table B3).⁵ Minor adjustments in

³CHVs distribute drugs during periodic Mass Drug Administration (MDA) rounds, which typically last one to two days.

⁴Implementation took place in two waves in 2016: early–mid October (Busia and Siaya) and late October–early November (Kakamega).

⁵Ninety percent of adults report walking to the PoT; the remainder travel mainly by bicycle.

site placement and household dispersion mean that a small share of Close households walked more than 1.25 km and vice versa; Figure A2 shows the full distribution. We next detail the process through which communities were assigned to Close or Far treatment conditions, subject to geographic and logistical constraints on site selection.

1. Cluster Selection and Assignment. We used a geographically constrained randomization procedure to identify suitable clusters and assign them to distance treatments. We began by drawing 158 clusters from a pool of 1,451 primary schools, using an acceptance–rejection algorithm to minimize spatial overlap among selected clusters.⁶ Each cluster was defined by a primary school—serving as a proxy for the PoT—and its catchment area. After cluster selection, we randomly assigned each cluster to either the Close or Far treatment group.

2. Anchor Schools and Target Communities. For each cluster, we designated an “anchor school” that satisfied the assigned distance category. For Close clusters, we used the originally selected or a nearby school; for Far clusters, we chose an alternative primary school to ensure sufficient geographic separation. Enumerators then surveyed local communities near the anchor school, and we randomly selected one as the study’s target community.⁷ In all but two clusters, the assigned PoT was indeed the closest deworming site to the selected community, ensuring strong compliance with the intended distance assignment. Appendix C provides additional details on the site selection and randomization process.

Distance as a Continuous Measure. Beyond the binary classification, we also use a continuous measure of distance in both the reduced form and structural analyses of the social multiplier. Specifically, we calculate the distance from each community’s centroid to its assigned PoT—both because this is common knowledge within the community and because it directly enters individuals’ utility as a travel cost.⁸

A potential concern is that the acceptance–rejection algorithm may have introduced systematic differences in the distribution of distances across communities, violating the assumption of random assignment. For example, more remote communities may have had fewer nearby primary schools available, resulting in higher expected distances under random assignment. To address this concern, we apply Borusyak and Hull (2023) to compute, for each community, the distribution of potential distances under the assignment mechanism. We re-run the randomization algorithm 100 times to generate counterfactual PoT placements and associated distance distributions. A two-sample Kolmogorov–Smirnov test shows no statistically significant difference in the counterfactual

⁶We initially targeted 150 clusters, drawing 8 additional clusters as fallbacks (see AEA RCT Registry). Ultimately, 144 clusters were implemented due to practical constraints.

⁷If the initially selected community was too small, an additional nearby community was included.

⁸Robustness checks using household-level distance are reported in Appendix B.

distance distributions for Close versus Far clusters (Figure C1), indicating that potential distances were balanced across arms.

Throughout our analysis, we include expected distance as a control variable to address any residual imbalance. Finally, we verify that the continuous distance measure is uncorrelated with baseline characteristics using a permutation test and find no statistically significant association with demographic covariates such as gender or age (Figure D1). Taken together, these checks support the internal validity of our design and justify analyzing distance as if it were randomly assigned—both as a binary and continuous treatment.

3.2 Signaling Treatments

We implemented two public signaling incentives and one active control, randomized at the Point of Treatment (PoT) level and cross-randomized with the distance treatment, yielding four incentive conditions:^{9,10}

- **Bracelet (39 PoTs; Close 21, Far 18).** Adults received a green bracelet upon deworming, saying “Treat worms: improve the health of your community” (Figure A4).
- **Ink (36 PoTs; Close 21, Far 15).** An indelible green ink mark was applied to an adult’s thumb upon deworming, similar to the ink used in elections.¹¹
- **Calendar (35 PoTs; Close 19, Far 16).** Adults received a simple wall calendar (Figure A5).¹²
- **Control (34 PoTs; Close 19, Far 15).** No incentive was given upon deworming.

The Bracelet and Ink treatments were designed to increase the observability of deworming decisions and thereby enhance existing social image concerns. However, signaling incentives may also confer private benefits (e.g., the bracelet’s decorative appeal) or private costs (e.g., dislike of the ink mark). The Calendar arm serves as an active control for the bracelet, matching private consumption value while minimizing public visibility. While the calendar is less visible outside the home, it serves as a personal reminder of participation and thus helps distinguish social image from self-signaling (self-image). Importantly, because incentives were randomized at the community level, their meaning was

⁹We piloted the incentives to identify acceptable colors and an active control item with comparable consumption value to the bracelet—small, useful, and scalable for implementation. Green was selected for both bracelets and ink because it avoids association with political parties and was favored by adults during piloting.

¹⁰The cross-randomization yielded 64 Far and 80 Close clusters overall. Because distance assignment was implemented within county and subject to geographic feasibility constraints at the anchor school stage, exact 1:1 balance by distance was not guaranteed. The realized split is consistent with chance under a 50–50 assignment (two-sided binomial $p = 0.21$), and per-arm counts are likewise within sampling variation (all $p > 0.40$). Identification derives from the randomized assignment itself, and our analyses incorporate county strata and a continuous distance measure.

¹¹Our implementing partner was particularly interested in testing ink due to its near-zero cost. Bracelets cost approximately US\$0.20 each.

¹²Wall calendars are popular decorative items in Kenya and cost approximately US\$0.50 each.

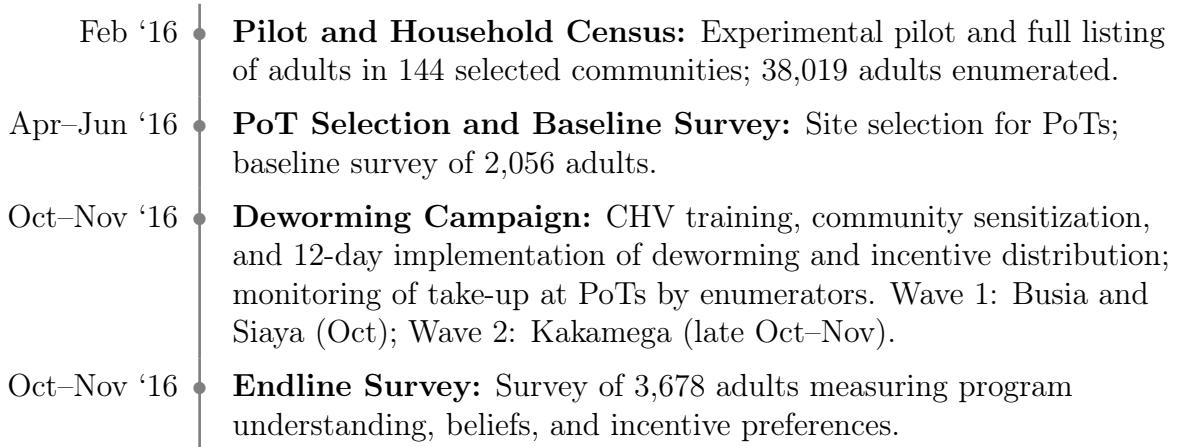


Figure 1: Timeline of the Experiment and Data Collection Activities

common knowledge: all adults in a given area knew which item was offered and what it signified.

3.3 Information Treatment

One week before the launch of the community deworming program, Community Health Volunteers (CHVs) and research staff visited each of the 144 selected communities to publicize the upcoming campaign. CHVs are trusted local health workers who live in the communities they serve and support a range of public health efforts, including Kenya’s school-based deworming program. During the community visits, CHVs: (i) informed households of the specific dates, location, and incentives for treatment; (ii) explained that regular deworming—even in the absence of symptoms—benefits both children and adults; and (iii) emphasized that deworming reduces worm transmission within the community, thereby protecting others (Figure A6 lists the full message script). In addition, CHVs distributed flyers and displayed posters showing the treatment dates, location, and relevant incentives (Figures A7). These materials served as visual reminders and intended to establish common knowledge about the program.

3.4 Experiment Timeline and Data

Our analysis uses several data sources, including administrative data on deworming take-up and survey data collected before and after the deworming campaign.

1. *Household Census.* We conducted a census of all adults (aged 18 or older) in the 144 selected communities. Enumerators visited each household to record geographic coordinates and basic demographics, and to determine eligibility for deworming

(e.g., pregnancy status). The census listing yielded 38,019 adults, from which we drew random subsamples for take-up monitoring, baseline and endline surveys, and the SMS treatment.

2. *Baseline Survey.* We randomly selected 15 households per community and surveyed one adult per household (2,056 individuals total) about the private and social benefits of deworming, prior deworming experience, and social norms (see Section 2).¹³
3. *Monitored Deworming Sample.* We randomly selected 12,827 adults from the census to be monitored for take-up at points of treatment (PoTs). Of these, 9,805 constitute the main analysis sample; 2,635 were assigned to the SMS Social Info arm and 387 to the Reminder Only arm. Enumerators used tablets pre-loaded with census data to confirm identity and record deworming status.
4. *Endline Survey.* We surveyed 3,678 adults 2–14 days after the campaign.¹⁴ The survey assessed implementation fidelity and understanding of the incentive and SMS treatments, and is our primary source for three endline outcomes: (i) beliefs about the observability of deworming for specific individuals the respondent knows; (ii) predictions about community take-up, elicited for a random sample of 10 community members; and (iii) externality knowledge (that worm infections can spread between individuals).¹⁵

To verify the SMS intervention, we administered an SMS verification module to a subsample of respondents assigned to receive a Social Info or Reminder Only SMS ($N = 1,019$). Our main endline analysis focuses on the SMS-control group to obtain clean estimates of the incentive-by-distance effects ($N = 2,659$). Among SMS controls, take-up was directly monitored at the point of treatment for $N = 2,312$ individuals. Unless otherwise noted, we therefore report endline results for this SMS-control, monitored sample. Results are robust to the inclusion of the non-monitored SMS controls ($N = 347$).

Gift Choice and WTP modules. At the end of the endline, respondents were offered a calendar or bracelet as a thank-you gift—identical to those distributed in the respective incentive arms—and asked to choose between them (“gift choice”). We use these choices to study private valuation of the two items. In the Control arm only, we additionally implemented a willingness-to-pay (WTP) exercise: after making

¹³We intended to survey 2,160 respondents and were able to contact 2,151 individuals, of whom 2,069 consented. Of those, 13 did not complete the survey and are excluded from the analysis. We find no statistically significant differences in survey response and completion rates across treatment arms.

¹⁴We intended to survey 3,750 adults. Of the 3,747 contacted, 3,702 consented; 24 did not complete the survey. Response and completion rates are balanced across treatment arms.

¹⁵The beliefs/observability module was randomized at the respondent level; its sample is $N = 1,627$. Of these, 1,035 respondents belong to the monitored sample; 36 could not name any acquaintances, leaving 999 respondents for whom we observe both beliefs and take-up outcomes.

the initial gift choice, respondents were offered a randomly drawn cash payment (US\$0–1) to switch items. Sixteen individuals declined both items and are excluded from the WTP estimation, leaving $N = 998$ choices used in the WTP analysis.

3.5 Randomization Checks

Table B1 reports control group means and standard errors for each baseline covariate, together with pairwise differences between treatment and control groups and the associated p -values, presented separately for the Far and Close conditions. The final column shows F -tests of joint significance across all treatments. All p -values exceed 0.10 except for three covariates: (i) Distance to PoT, by design; (ii) Know medicine treats worms, driven by higher knowledge in the Close–Ink arm; and (iii) Dewormed in the past year, which is higher in the Far–Bracelet and Far–Ink arms. Within the Far condition, communities assigned to Ink are on average 344 meters farther from the PoT than Far–Control communities ($p = 0.02$). To address this imbalance, we include continuous community-to-PoT distance as a control in robustness specifications of the reduced form analysis; the structural model conditions directly on continuous distance.

Overall, baseline characteristics are well balanced. Out of 200 comparisons in Table B1, 16 differ at the 10 percent level and 7 at the 5 percent level—fewer than the 20 and 10 differences expected by chance. Because our identification strategy exploits variation both within distance conditions (across incentive arms) and within incentive arms (across distance), Table B4 reports balance across distance within each incentive arm. Of 104 such comparisons, 12 differ at the 10 percent level and 5 at the 5 percent level, again consistent with chance.

3.6 Implementation Fidelity and Understanding of Signals

Endline survey evidence confirms that the deworming campaign was implemented as intended (Table B1, Panel E). Eighty-three percent of respondents recalled a visit from a CHV prior to the launch.

Table B5 shows that each incentive—ink, bracelet, and calendar—was distributed effectively at the PoT, with more than 95 percent of respondents in the respective arms reporting receipt of the assigned item. Nearly all respondents (94–96 percent) correctly associated the ink mark and bracelet with deworming, and 90 percent recognized the calendar as indicating that an individual had attended treatment. Because indelible ink fades within three days on skin and roughly two weeks on nails, only 14 percent of respondents displayed a visible ink mark at endline. By contrast, 81 percent still possessed their bracelet and 95 percent retained the calendar; 46 percent were still wearing the bracelet. Consequently, the bracelet was the most publicly observable signal: 95 percent of respondents reported seeing it in their community. Given the calendar’s high reten-

tion yet limited public visibility, it serves as an appropriate active control comparison for isolating any self-signaling (self-image) component of the bracelet incentive.

3.7 Preregistration and Analytical Extensions

Before data collection we preregistered a plan on the AEA RCT Registry (AEARC-TR-0001643) covering the treatment arms (bracelet, ink, calendar, control) cross-cut by individual SMS treatments, the primary outcome (adult deworming) and secondary belief/knowledge outcomes, and the set of reduced-form comparisons and distance heterogeneity. We implemented all arms and report every preregistered comparison; no unregistered reduced form tests are presented as primary results. After observing the reduced form patterns, we added two analytical extensions: (i) a structural adaptation of [Benabou and Tirole \(2025\)](#) that uses the experiment’s randomized variation in observability and distance to quantify the social multiplier component of take-up, where the structural step decomposes the distance gradient into private cost and social image components; and (ii) a policy allocation exercise that applies those estimates to a feasible siting problem to choose point of treatment locations that meet coverage targets. These extensions are clearly labeled as post-PAP and are intended to interpret—not replace—the preregistered evidence. Appendix Table [E1](#) maps preregistered items to the corresponding analyses in the paper.

4 Reduced Form Estimates

4.1 Empirical Strategy

We estimate the effects of distance and incentive treatments on (i) the observability of deworming and (ii) individual take-up. For individual i in community c we estimate:

$$Y_{ic} = \alpha_s + \sum \beta_z \text{treat}_{z,ic} + \sum \gamma_z \text{treat}_{z,ic} \times \text{Far}_{ic} + X_{ic} + \varepsilon_{ic}$$

where Y_{ic} is either (a) the share of peers whose deworming status respondent i reports knowing or (b) a binary indicator equal to 1 if i dewormed. $\text{treat}_{z,ic}$ denotes indicators for the Control, Calendar, Ink, and Bracelet arms, where z indexes each arm. Far_{ic} equals 1 for communities assigned to the Far distance condition (1.25–2.5 km) and 0 otherwise. The interaction terms capture differential incentive effects at greater distance. X_{ic} includes baseline covariates selected via LASSO (sex and age) and the community’s expected distance to the PoT computed using the method of [Borusyak and Hull \(2023\)](#). α_s denotes strata fixed effects, and ε_{ic} is an idiosyncratic error. Standard errors are cluster-bootstrap estimates at the community level.

4.1.1 Effects of Distance and Signals on the Observability of Deworming

We first examine how travel distance and signaling incentives jointly affect whether an individual's deworming status is known within the community.

At endline, respondents were shown a random subset of ten adults from the same community and asked, conditional on recognition, "Do you think this person came for deworming?". We code "Yes" and "No" as 1 (status known) and "Don't know" as 0 (status unknown). Observability is the respondent-level share of recognized peers whose status is known; it captures the ease with which deworming can be observed or inferred through direct sight, communication, or close familiarity.

Deworming decisions are highly observable in a campaign setting: Control group respondents know the status of 69.6 percent of recognized peers (Table 1). Knowledge is higher in Close communities (75.1 percent) than in Far communities (63.1 percent, $p = 0.056$). Endline qualitative data point to two mechanisms (Figure A8): (i) greater proximity to PoTs increases direct encounters and conversation; (ii) lower take-up in Far communities reduces opportunities for observation. Absent physically seeing a peer traveling to the PoT or present there, it is hard to discern whether the peer failed to deworm or whether the deworming simply went unobserved.

Signaling incentives raise observability, especially in Far communities. The Bracelet and Ink treatments increase knowledge of peers' status by 12.6 percentage points ($p = 0.001$) and 8.1 points ($p = 0.046$), respectively. The treatment effects are sharply heterogeneous: for the Bracelet, observability rises by 22 percentage points in Far communities ($p < 0.001$) but by only 4.8 points in Close communities ($p = 0.357$); for the Ink, the corresponding increases are 14.8 points in Far communities ($p = 0.007$) and 2.6 points in Close communities ($p = 0.644$). Because obtaining the signal is costlier in Far communities, the signal likely conveys stronger information about an individual's motivation and is more actively displayed, resulting in larger gains in observability.

As expected, the Calendar has a limited effect on observability, increasing it by 3.5 percentage points ($p = 0.382$). The difference between the Calendar and Bracelet arms is significant ($p < 0.05$) in all comparisons except for Close communities ($p = 0.826$).

Combining the Bracelet and Ink arms (Any Signal) and comparing them with the Calendar and Control arms (No Signal), we find that signals offset the negative distance gradient in observability by 15 percentage points ($p = 0.002$). Table B6 shows that the results are robust when distance enters the specification as a continuous variable.

4.1.2 Effects of Distance and Signals on Deworming Take-up

Table 2 reports average treatment effects on the share of adults who dewormed. Even in a large, community-wide campaign, take-up is far from universal: only 33.8 percent of adults in the Control arm attended treatment. Consistent with the idea that travel

Table 1: Effects of Signals and Distance on the Observability of Deworming

Dependent variable: Observability	Reduced Form			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.126** (0.038)	0.048 (0.052)	0.221*** (0.053)	0.173** (0.072)
Calendar	0.035 (0.04)	0.039 (0.056)	0.03 (0.057)	-0.009 (0.081)
Ink	0.081** (0.041)	0.026 (0.057)	0.148** (0.055)	0.122 (0.077)
Control mean	0.696 (0.032)	0.751 (0.045)	0.631 (0.045)	-0.12* (0.063)
Observations	999	999	999	999
H_0 : Any Signal = No Signal, p -value	0.001	0.624	<0.001	0.002
H_0 : Bracelet = Calendar, p -value	0.002	0.826	<0.001	0.004

Notes: Each column reports OLS estimates of the respondent-level observability measure—the fraction of recognized community members whose deworming status the respondent reports knowing. “Control” is the mean in the Control arm. Rows for Bracelet, Ink, and Calendar give treatment effects relative to Control. Column (4) reports the difference in those effects between Far and Close communities. We pool Bracelet and Ink as “Any Signal” and Control and Calendar as “No Signal” to test H_0 : Any Signal = No Signal. The second test, H_0 : Bracelet=Calendar, compares the bracelet and active control incentive. All regressions include strata fixed effects and the LASSO-selected controls—sex, age, and expected distance to the PoT. Bootstrapped standard errors, clustered at the community level, are reported in parentheses. Table B7 reports similar results for the sample that includes individuals who received the SMS treatment and individuals who were not monitored at treatment locations and their deworming status is therefore not observed. Table B8 presents analogous results for respondents’ beliefs about how observable their own deworming decisions are to other community members. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

distance imposes a non-trivial cost, 41 percent of adults in Close communities dewormed, whereas 24.8 percent did so in Far communities—a decline of 16.2 percentage points ($p < 0.001$). Respondents’ beliefs move in tandem: in the Control arm they expect 12.1 percentage points fewer adults to deworm in Far than in Close communities (Table B9).

Providing a bracelet increases take-up by 7.5 percentage points relative to Control ($p = 0.008$). Ink has no detectable effect and even shows a negative point estimate (-3.0 pp, $p = 0.35$), plausibly because respondents dislike its messiness or associate it with voting.¹⁶ The Calendar yields a modest and insignificant 2.7 percentage point increase ($p = 0.334$). The Bracelet–Calendar gap of 4.8 percentage points ($p = 0.048$) provides a naïve estimate of the social image component. Calendar effects are identical in Close and Far communities (-0.0 pp, $p = 0.959$), implying that the 16 pp distance gradient persists when no public signal is offered. Bracelet and Ink appear more effective in Far communities—treatment effects are 7.5 pp (Bracelet, $p = 0.218$) and 5.7 pp (Ink, $p = 0.363$) larger than in Close communities—but the interaction terms are imprecisely

¹⁶Indelible ink marks are commonly used in Kenyan elections. Kenya’s presidential election was held in 2017 amid allegations of voter fraud.

Table 2: Effects of Signals and Distance on Deworming Take-up

Dependent variable: Take-up	Reduced Form			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.075** (0.029)	0.042 (0.037)	0.117** (0.047)	0.075 (0.061)
Calendar	0.027 (0.028)	0.028 (0.038)	0.025 (0.041)	-0.003 (0.056)
Ink	-0.03 (0.032)	-0.055 (0.037)	0.002 (0.052)	0.057 (0.063)
Control mean	0.338 (0.023)	0.41 (0.027)	0.248 (0.038)	-0.162*** (0.046)
Observations	9,805	9,805	9,805	9,805
H_0 : Any Signal = No Signal, p -value	0.632	0.384	0.13	0.083
H_0 : Bracelet = Calendar, p -value	0.048	0.708	0.004	0.101

Notes: Each column reports OLS estimates of the respondent-level take-up indicator—equal to 1 if the adult dewormed during the campaign. The estimating equation saturates the binary distance indicator (Far vs. Close) with dummy variables for the four incentive arms (Control, Calendar, Ink, and Bracelet). “Control” is the mean take-up rate in the Control arm; rows for Bracelet, Ink, and Calendar give treatment effects relative to Control. Columns (1)–(4) correspond to the combined sample, Close communities, Far communities, and the Far–Close difference, respectively. We pool Bracelet and Ink as “Any Signal” and Control and Calendar as “No Signal” to test H_0 : Any Signal = No Signal. The second test, H_0 : Bracelet = Calendar, compares the bracelet to the active control incentive. All regressions include strata fixed effects and the LASSO-selected controls—sex, age, and expected distance to the PoT. Bootstrapped standard errors, clustered at the community level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

estimated given only 17–20 clusters per distance-incentive cell. Pooling Bracelet and Ink (Any Signal) and comparing them with Calendar and Control (No Signal) reduces the distance penalty by 6.8 pp, reaching marginal significance ($p = 0.083$).

Estimates are similar when we replace the binary distance indicator with continuous centroid–PoT distance (Table B10). Bracelet raises take-up by 6 pp in Close communities ($p = 0.065$), and the Ink interaction with distance becomes larger (7.6 pp, $p = 0.067$). The difference-in-difference test of whether signals are differentially effective at greater distances is significant at the 5 percent level ($p = 0.049$).

To visualize these interactions in the raw data, Figure A9 plots community-level take-up against the centroid-to-site distance with arm specific penalized splines. Three features stand out. First, take-up declines with distance in all arms. Second, the decline is markedly steeper in the lower-visibility arms (Control/Calendar) than when deworming is publicly signaled (Bracelet/Ink): the fitted curves diverge with distance, so the signal arms exhibit a smaller distance penalty. Third, curvature differs by arm—Control shows an accelerating decline at longer distances, whereas Bracelet flattens beyond roughly one kilometer.¹⁷ These non-parametric patterns mirror the regression results and motivate

¹⁷The curvature patterns are less visually pronounced in the Ink and Calendar arms which is unsur-

the structural decomposition that follows.¹⁸

Taken together, the reduced form evidence supports the mechanism that social image returns rise with distance: as fewer adults deworm, the observability premium from signaling increases and offsets part of the higher travel cost.¹⁹

4.2 Alternative Mechanisms and Robustness

4.2.1 Reminder and Social Learning Effects

Because adults could seek treatment on any of the twelve consecutive campaign days, the bracelet and ink might also operate as reminders or facilitate social learning by providing information about community take-up. To gauge the importance of these alternative mechanisms, we implemented an SMS treatment among adults who owned a phone (75 percent of the sample). We randomly selected 20 adults per incentive community and 12 adults per Control community ($N = 2,635$) to receive Social Info SMS that both reminded them of free deworming and reported the fraction of adults in their community who had already participated.^{20,21} We limited recruitment to fewer than 10 percent of each community to reduce spillovers and avoid shifting beliefs about aggregate take-up. Treatment recipients received their first message the evening before the campaign began and one message every other day thereafter. The control group for the SMS treatment consists of 7,363 phone owners who received no texts.

Figure A10 compares incentive effects for respondents who received Social Info SMS (red bars) with those who did not (blue bars). Across the Bracelet, Ink, and Calendar arms, point estimates and significance levels are broadly similar in the SMS and non-SMS subsamples. If the Bracelet or Ink operated primarily as reminders or through social learning, their effects should diminish once equivalent information is delivered by SMS; instead, the red and blue confidence intervals overlap in every case, and no pairwise difference is statistically significant at the 5 percent level. Hence, the impact of signaling

prising given non-parametric estimators are relatively data-intensive.

¹⁸With larger samples per distance cell one could non-parametrically compare arm-specific distance curves to form a reduced form analogue of the multiplier for a signal arm. This would still leave the Control-arm multiplier unidentified and would not separate observability from inference, because both co-move with distance. Our model in Section 5 uses structure only to the extent needed to map the measured observability and take-up responses into the two components of the social image term.

¹⁹Appendix Table B11 explores heterogeneity by the limited individual and community covariates available (age, gender, phone ownership, prior deworming, attitudes, and externality knowledge). Power is low, but adults aged 40 and older appear especially averse to Ink. We find no statistically significant community-level moderators, likely due to limited variation.

²⁰Message: “Free deworming now at [location]. [No/Few/Almost half/Half/More than half/Almost all/All] of your village came—that is [X] in 10 adults. Reply 111 to stop texts. All texts are free.”

²¹In addition, we selected 387 phone owners in Control communities to receive a Reminder Only text with identical wording to the Social Info message but without the social information clause: “Free deworming now at [location]. Reply 111 to stop texts. All texts are free.” Estimates for the Reminder Only arm appear in Table B12. They are statistically indistinguishable from those in the Social Info arm, suggesting that the SMS served as a general reminder to deworm.

is unlikely to be driven by these alternative channels.

4.2.2 Externality-Based Motives

A second concern is that the larger signal effects in Far communities reflect externality motives rather than social image. In principle, an adult might reason, “If few neighbors deworm, I should do so to protect myself. If many neighbors deworm, I can free-ride”. Three features of the data argue against this explanation.

First, there is at most weak evidence that public signals increased knowledge of epidemiological externalities, and the pattern does not match the distance gradient. At baseline, about 37 percent of adults correctly reported that worm infections can spread to others, with no Close–Far difference (Table B4, Panel C). At endline, pooling Bracelet and Ink against Calendar and Control yields a small average increase (Any Signal vs. No Signal, $p = 0.10$) but no Far–Close differential ($p = 0.85$; Table B13). Bracelet coefficients are small and imprecise (Close: +0.4 pp, $p = 0.947$; Far: +6.3 pp, $p = 0.214$), and Bracelet and Calendar are statistically indistinguishable overall ($p = 0.785$). Ink shows an increase in Close communities (+10.7 pp, $p = 0.066$) and a smaller, imprecise coefficient in Far (+5.1 pp, $p = 0.414$); their difference is not significant (−5.6 pp, $p = 0.505$). Calendar effects are similarly small (Close: +3.0 pp, $p = 0.627$; Far: +1.2 pp, $p = 0.781$). Because knowledge gains are not larger in Far—negligible for Bracelet and, if anything, concentrated in Close for Ink—differential learning about externalities is unlikely to account for the larger Far-than-Close take-up effects in either signal arm.

Second, the signals do not cause respondents to revise downward their beliefs about community take-up. Under standard free-riding logic, perceiving lower coverage should increase one’s own demand for deworming, while higher expected coverage should reduce it. Instead, respondents in the signal arms expect higher—rather than lower—community coverage than in Control, in both distance conditions (Table B9). In Control communities, respondents overestimate participation (predicting 57.3% versus an actual 33.8%). Relative to Control, beliefs are 9.7 percentage points higher in Bracelet and 1.7 points higher in Ink. Thus, greater observability from the signals did not lead adults to infer that fewer peers dewormed, making a free-riding mechanism an unlikely explanation for the observed treatment effects.

Third, the distance interaction appears only in arms with a public signal. Calendar and Control show the full 16-percentage-point drop in take-up between Close and Far communities, whereas Bracelet and Ink attenuate this gradient. Externality motives should apply equally across arms, whether or not a bracelet or ink mark is offered. Accordingly, the distance effect attenuates only when deworming becomes more observable—consistent with social image concerns. Taken together, these patterns indicate that the stronger signal effects in Far communities are not driven by externality calculations.

4.3 Private Valuation: Evidence from Gift Choice and WTP

An alternative to the signaling interpretation is that bracelets have a higher private value than calendars and that this valuation could be larger in Far communities.

We first assess any difference in private monetary valuation between the two items by implementing an endline willingness-to-pay (WTP) exercise in the Control arm. Using 998 endline choices, respondents first chose between a bracelet and a calendar as a thank-you gift and were then offered a randomly drawn cash payment (US\$0–1) to switch; this random-offer design identifies a monetary indifference point between the items. Seventy-three percent preferred the calendar to the bracelet, and the implied mean value gap is US\$0.47 in favor of the calendar (Table [B14](#)).

Because the Control-arm WTP exercise samples only a few respondents per community, it does not capture any social utility arising from item prevalence. To test for prevalence-related value, we examine endline gift choices among adults in Bracelet and Calendar communities who did not deworm and therefore did not already possess either item. In these communities the prevalence of each item is common knowledge, so gift choice should reflect any incremental private value associated with local prevalence. Table [B15](#) shows a prevalence penalty—non-holders avoid the prevalent item—in Calendar and in Bracelet (Close), but not in Bracelet (Far); this pattern does not support a scarcity-only increase in bracelets’ private value with distance and instead aligns with higher social image value in Far.

In Control communities, preferences are stable across distance: 23.5 percent choose the bracelet (Close: 24.9 percent; Far: 22.2 percent). Preferences are similarly stable in the Ink arm, consistent with the fact that neither item is prevalent in Ink. Thus, where ownership of either item is not widespread, most respondents prefer the calendar and this preference does not vary by distance.

When an item is already prevalent, the demand of those who do not yet have it falls. In the Calendar arm, the probability of choosing the bracelet rises by 15.9 percentage points (pp) ($p < 0.01$) relative to Control, with similar effects in Close (+19.9 pp, $p < 0.01$) and Far (+12.2 pp, $p < 0.05$). Symmetrically, in the Bracelet arm the probability of choosing the bracelet falls by 6.6 pp ($p < 0.05$) overall, driven entirely by Close communities (−13.6 pp, $p < 0.01$), with no detectable change in Far (−0.2 pp). The distance interaction is significant for Bracelet (Far–Close = 13.4 pp, $p < 0.05$).

If bracelets had higher private value in Far communities, non-dewormers living in Bracelet communities should be more likely to choose a bracelet at endline; instead, their choices in Far are statistically indistinguishable from Control. The evidence is consistent with a signaling interpretation: in Far communities the social image value of bracelets offsets any prevalence fatigue, leaving private demand no higher than in Control.

4.3.1 Substitution

A residual concern is that our findings could be driven by differences in individuals' marginal rates of substitution between private incentives and travel cost. In particular, if the bracelet's private value trades off differently against walking distance in Far versus Close communities, we might observe larger bracelet effects at greater distances even without social image being at play. Appendix F formalizes this argument: if heterogeneous substitution alone explained the distance interaction, the willingness to pay (WTP) gap between the bracelet and the calendar should widen with distance, because the calendar's own effect is essentially flat across distance. Empirically, we find no such pattern—preferences for the bracelet relative to the calendar are invariant to distance (Table B15). Therefore, heterogeneous substitution alone cannot account for the stronger signal effects in Far communities.

4.4 Summary

The reduced form evidence delivers two central findings. First, the distance assignment creates exogenous variation in take-up and observability: moving a community from Close to Far lowers deworming by about 16 percentage points and, by reducing direct encounters, makes deworming less observable. Second, the signaling arms increase observability and raise participation relative to the no signaling arms. The Far–Close drop in take-up is meaningfully smaller with public signals than without them, indicating that signals blunt the distance penalty. Alternative mechanisms—salience, scarcity of incentives, or social learning—cannot jointly explain both the average increase in take-up and the disproportionate effect of signals at greater distances.

Taken together, these results point to two opposing forces. Greater distance increases travel effort and reduces opportunities for others to observe participation, while public signals restore observability and thereby attenuate the distance effect on take-up. Because distance and observability can push in either direction—and their net effect differs between signaling and no signaling arms—we formalize these patterns using a simple framework adapted from Benabou and Tirole (2025): changes in distance shift who deworms and how observable their actions are, making the distance response smaller under high observability and larger under decreasing observability.

5 Theoretical Framework

We adapt the signaling framework of Benabou and Tirole (2025) for two purposes. First, it delivers predictions for how cost and observability jointly interact to determine deworming take-up, guiding our interpretation of the reduced form evidence in Section 4. Second, it provides the structure for Section 6, where we quantify the *social multiplier*—the pro-

portional responsiveness of take-up to distance when social image concerns are present, relative to a private cost benchmark.

5.1 Deworming with Social Image Concerns

Setup. Individual i chooses whether to deworm, $y_i \in \{0, 1\}$, weighing net private benefits with reputational gains. Utility is given by

$$U_i = (b - d + v_i + u_i)y_i + \mu(d)E_{-i}[V|y_i]$$

where $b - d$ is the net private payoff, combining the health (and material incentive) benefit b with a distance-induced travel cost $d \geq 0$. Intrinsic motivation v_i —an individual’s health conscientiousness or civic-mindedness—is privately known to them and unobservable to others. The idiosyncratic shock u_i captures taste variations or one-off costs and is independent of v_i . Since both variables are latent, we define the composite type $w_i \equiv v_i + u_i$. $\mu(d)$ is the weight placed on social image and may itself depend on distance, while $E_{-i}[V | y_i]$ is the expectation that observers form about v_i after seeing action y_i . Following [Benabou and Tirole \(2025\)](#), $\mu(d)$ bundles two elements: the physical observability of the act and the individual’s valuation of social image (their parameter λ). Observability is empirically much lower in Far than in Close communities, so we allow $\mu(d)$ to decline with distance. Conversely, if performing a costlier act makes the signal more valuable and actively displayed, $\mu(d)$ could rise with d ; the model accommodates either pattern.

Equilibrium. Following [Bénabou and Tirole \(2006\)](#); [Benabou and Tirole \(2025\)](#), a unique equilibrium exists under observability where each individual’s choice y_i depends on the equilibrium actions of all other individuals. Equilibrium behavior is characterized by a cut-off type $w^*(d)$ such that an individual with $w_i > w^*(d)$ deworms and one with $w_i \leq w^*(d)$ abstains. The cut-off satisfies the fixed-point equation:²²

$$w^*(d) - d + b + \mu(d)\Delta(w^*(d)) = 0$$

where $\Delta(w^*(d)) = E[V|w > w^*(d)] - E[V|w \leq w^*(d)]$ is the difference in expected intrinsic motivation between those who deworm and those who abstain. Hence the social image return is the product of observability and type inference, $\mu(d)\Delta(w^*(d))$. Given the cut-off, the equilibrium participation rate at distance d is

$$\bar{y}(d) = 1 - F_w(w^*(d)) \tag{1}$$

²²The cut-off $w^*(d)$ depends on the primitives $(b, d, \mu(d))$. Our comparative static results focus on d , treating the other parameters as fixed.

where F_w is the cumulative distribution function of w .

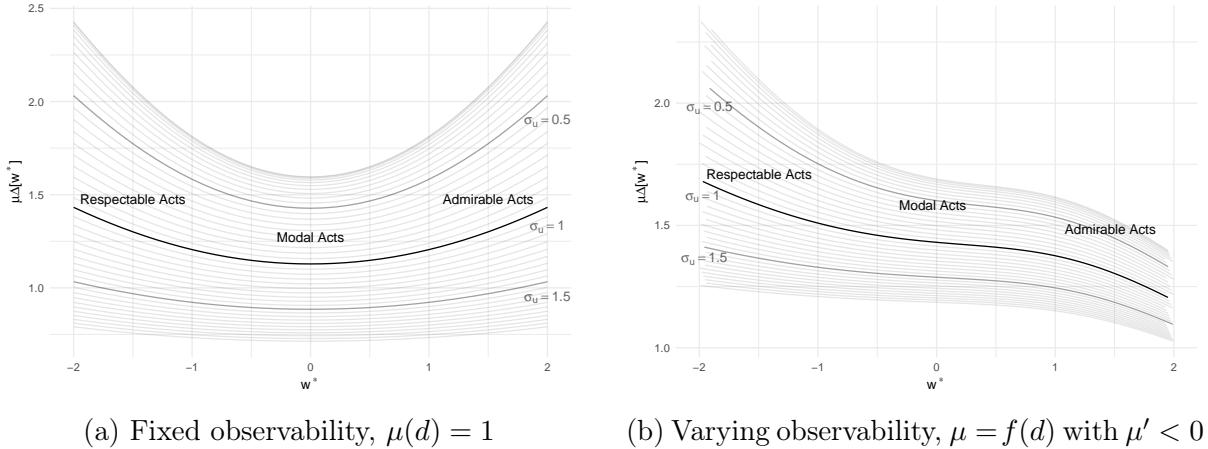


Figure 2: Equilibrium Social Image Returns $\mu(d) \Delta(w^*)$

Notes: Each panel plots $\mu(d) \Delta(w^*)$ as the equilibrium cut-off type w^* varies. As the variance of the idiosyncratic shock σ_u rises, observers find it harder to infer v_i from actions, which reduces social image returns. Panel (a) holds $\mu(d) = 1$ and evaluates $\mu(d) \Delta(w^*)$ on a grid of w^* and σ_u values. Panel (b) lets $\mu(d)$ decline with distance while using the same computation; the latent type is drawn from $w \sim N(0, 1 + \sigma_u)$.

Social Image Returns and Cut-off Type. Figure 2a illustrates how social image returns, $\mu(d) \Delta(w^*(d))$, change as the cut-off type $w^*(d)$ —and thus the participation rate—varies while $\mu(d)$ is held fixed. When $w^*(d)$ is near zero (“modal” acts) an individual’s decision conveys little information about their type, so Δ is small. For “respectable” acts ($w^* \ll 0$) almost everyone deworms except the very lowest types, producing strong type inferences and increasing Δ . For “admirable” acts ($w^* \gg 0$) only the most motivated deworm, again raising Δ . Because deworming is a minority behavior in our data, the relevant region is likely the admirable range.

Allowing for idiosyncratic shocks u_i blurs type inferences. When the variance σ_u is small (lighter curves at the top of Figure 2a), the shock contributes little in expectation and actions closely reveal v_i . As σ_u rises, observers cannot tell whether an individual deworms because of intrinsic motivation or because of a favorable shock. Social image returns fall at every cut-off w^* , and the U-shaped returns curve progressively flattens.²³

Figure 2b allows observability to fall with distance, $\mu'(d) < 0$. If the decline in $\mu(d)$ dominates the rise in Δ , the U-shape can invert, producing a downwards sloping pattern: at very large d , only the most motivated individuals deworm, yet their virtue is largely hidden because few witness the act, and the social image return $\mu(d) \Delta(w^*(d))$ becomes small.

²³The curve flattens rather than merely shifts downward. Holding v fixed, an increase in σ_u has its greatest impact in the tails of the w^* distribution, where a larger share of individuals act because of noise rather than high motivation.

The Social Multiplier. A social multiplier different from one arises because the cut-off type—and hence social image returns—shifts when distance changes. To see this, differentiate the equilibrium participation rate in (1) with respect to d :

$$\frac{\partial \bar{y}}{\partial d} = -f_w(w^*(d)) \underbrace{\frac{1 - \mu'(d)\Delta(w^*(d))}{1 + \mu(d)\Delta'(w^*(d))}}_{\partial w^*/\partial d}.$$

Two endogenous channels accompany the direct cost term $-f_w(w^*)$:

1. Type inference effect $\Delta'(\cdot)$: A higher cost filters out less motivated individuals; if this makes deworming appear more virtuous so that $\Delta'(\cdot) > 0$, the social image payoff rises with distance and partly mitigates the direct cost effect. By contrast, if a marginal increase in distance pushes the activity toward the modal range so that $\Delta'(\cdot) < 0$, the social image payoff falls and amplifies the direct cost effect.
2. Observability effect $\mu'(\cdot)$: Greater distance may reduce the observability of actions, $\mu(d)$, and thus amplify the cost effect.

Whether distance mitigates or amplifies its own direct impact depends on $\partial w^*/\partial d$. If the cut-off shifts more than one-for-one ($\partial w^*/\partial d > 1$), amplification occurs; if it shifts less than one-for-one ($\partial w^*/\partial d < 1$), the cost effect is mitigated.

Existence of a Social Multiplier. A social multiplier is present whenever the cut-off does not adjust one-for-one with distance, that is $\partial w^*/\partial d \neq 1$. Two special cases instead yield $\partial w^*/\partial d = 1$, so the multiplier equals one.

The first is a knife-edge case: if the marginal decrease in observability exactly offsets the marginal increase in type inference, $-\mu'(d)\Delta(w^*(d)) = \mu(d)\Delta'(w^*(d))$, the two forces cancel, so the cut-off shifts one-for-one with distance.

The second case is when social image returns are locally flat. Formally, this requires: $\mu'(d) = 0$ and $\Delta'(w^*(d)) = 0$. Because $\mu'(d) = 0$ whenever observability does not vary with distance, the critical restriction is $\Delta' = 0$. Recall $\Delta(w^*) = E[V | w > w^*] - E[V | w \leq w^*]$, the difference between the two “centers of mass” defined by the cut-off w^* . In a non-uniform density f_w , moving the cut-off changes those centers unequally, so Δ typically changes. Δ can be locally flat only if f_w is (approximately) uniform in a neighborhood of the cut-off, causing both centers to shift in parallel and leave their gap unchanged. A fully uniform density gives $\Delta(w^*) \equiv c$ for all w^* ; a symmetric density does so only at $w^* = 0$ where marginal gains in “honor” equal marginal losses in “stigma”.

Figure 2a illustrates the uniform density limit: as the variance of the idiosyncratic shock σ_u becomes large, the composite type $w = v + u$ becomes dominated by the noise term,

and its probability density flattens toward a uniform distribution. Consequently $\Delta(w^*)$ approaches a constant, $\Delta'(w^*) \rightarrow 0$, and—since $\mu'(d) = 0$ —the social multiplier converges to one. Introducing the shock u_i nests both the benchmark of Benabou and Tirole (2025) and the constant return case, leaving the model agnostic about the multiplier’s magnitude. In Section 6 we estimate σ_u directly. If the data were to imply a very large value, Δ' would be small and the implied multiplier close to one (i.e., negligible).

5.2 Predictions

To connect the model to the reduced form evidence in Section 4, we state three theoretical predictions.

Prediction 1. Direct Cost Effect: *An increase in travel distance d raises the private cost of deworming, thereby reducing aggregate take-up:*²⁴

$$\frac{\partial \bar{y}}{\partial d} = -f_w(w^*(d)) \underbrace{\frac{1 - \mu' \Delta(w^*(d))}{1 + \mu \Delta'(w^*(d))}}_{\partial w^*/\partial d} < 0$$

Consistent with this prediction, take-up is 16 percentage points lower in Far than in Close communities.

Prediction 2. Direct Observability Effect: *Holding distance fixed, higher observability μ raises social image returns and increases take-up:*

$$\frac{\partial \bar{y}}{\partial \mu} = f_w(w^*(d)) \underbrace{\frac{\Delta(w^*(d))}{1 + \mu \Delta'(w^*(d))}}_{\partial w^*/\partial \mu} > 0$$

Consistent with this, the high-visibility Bracelet arm yields markedly higher participation than the low-visibility Calendar or Control arms.

Prediction 3. Indirect Effect of Distance - Amplification versus Mitigation: *If the semi-elasticity of the type-inference term exceeds that of observability,*

$$\frac{\partial \Delta / \partial d}{\Delta} > -\frac{\partial \mu / \partial d}{\mu},$$

then a marginal increase in distance is mitigated:

$$\frac{\partial \bar{y}}{\partial d} = -f_w(w^*(d)) \frac{1 - \mu'(d) \Delta(w^*(d))}{1 + \mu(d) \Delta'(w^*(d))} > -f_w(w^*(d)).$$

²⁴Following Benabou and Tirole (2025), we assume that the multiplier is positive – this holds true provided social image returns are not too large $1 - \mu' \Delta(w^*(d)) > 0$. Figure A11 verifies this condition for all distances we consider, using estimates from our structural model.

Conversely, if the semi-elasticity of observability dominates,

$$-\frac{\partial \mu / \partial d}{\mu} > \frac{\partial \Delta / \partial d}{\Delta},$$

the distance effect is amplified:

$$\frac{\partial \bar{y}}{\partial d} = -f_w(w^*(d)) \frac{1 - \mu'(d) \Delta(w^*(d))}{1 + \mu(d) \Delta'(w^*(d))} < -f_w(w^*(d)).$$

Our experimental results show that the negative effect of distance on deworming take-up is substantially smaller when observability is high—Bracelet or Ink—than when it is low, as in the Calendar or Control arms.

6 Structural Estimates

This section presents analyses that were not prespecified in the preregistered plan (AEARC-TR-0001643; see Section 3.7 and Appendix Table E1). These are theory driven extensions intended to interpret—rather than replace—the preregistered reduced form estimates.

We take the model from Section 5 to the data to obtain explicit estimates of the social multiplier. The analysis proceeds in four steps. First, we map the theory into three interlocking submodels: (i) a deworming submodel fit to 9,805 monitored take-up observations; (ii) an observability submodel estimated on belief observations from 999 respondents; and (iii) a willingness to pay (WTP) submodel based on 998 valuation choices. Because the take-up data alone cannot identify social image and preference parameters, the deworming submodel takes those parameters from the observability and WTP submodels. We combine the three components in a joint likelihood so that the resulting posterior must simultaneously rationalize all observed outcomes. Second, we show that the model implied average treatment effects replicate the reduced form evidence. Third, we decompose each treatment effect into its social image and private valuation components. Finally, we report the social multiplier estimates for each treatment arm.

Deworming Take-up Model. We parameterize utility using indicator variables for the incentive arms, $\text{treat}_{z,i}$, and each individual's community distance to the treatment site, d_i :

$$U_i = (\beta_z \text{treat}_{z,i} - \delta d_i + w_i)y_i + \mu(\text{treat}_{z,i}, d_i)E_{-i}[V|y_i]. \quad (2)$$

Here β_z captures the private value of incentive $z \in \{\text{Control}, \text{Ink}, \text{Calendar}, \text{Bracelet}\}$; $\delta > 0$ is the marginal disutility of distance; and $\mu(\cdot)$ is the weight on social image. We write $w_i = v_i + u_i$ and assume $w_i \sim \mathcal{N}(0, \sigma_u^2 + 1)$, normalizing $\text{Var}(v_i) = 1$.²⁵

²⁵Although error variances are typically unidentified in latent-index models, σ_u enters the conditional

Following Benabou and Tirole (2025), we factor observability as

$$\mu(\text{treat}_{z,i}, d_i) = \lambda \cdot p_{\text{Observed}}(\text{treat}_{z,i}, d_i),$$

where $p_{\text{Observed}}(\cdot)$ is the probability that an individual's deworming status is known in the community. We estimate p_{Observed} by a logit model using the endline beliefs data; λ is a scalar converting observability into utility.

The equilibrium cut-off satisfies the fixed-point equation

$$w^*(\text{treat}_{z,i}, d_i) - \delta d_i + \beta_z \text{treat}_{z,i} + \mu(\text{treat}_{z,i}, d_i) \Delta(w^*(\text{treat}_{z,i}, d_i)) = 0. \quad (3)$$

Since $w^*(\text{treat}_{z,i}, d_i)$ is the type exactly indifferent between deworming and abstaining, the probability that individual i deworms under treatment z at distance d_i is

$$P(y_i = 1 | \text{treat}_{z,i}, d_i) = 1 - \Phi\left(\frac{w^*(\text{treat}_{z,i}, d_i)}{\sqrt{\sigma_u^2 + 1}}\right),$$

where $\Phi(\cdot)$ denotes the standard normal c.d.f. Accordingly, the individual likelihood contribution is

$$\mathcal{L}(\beta_z, \delta, \lambda, \sigma_u | y_i, \text{treat}_{z,i}, d_i) = \left[1 - \Phi\left(\frac{w^*(\text{treat}_{z,i}, d_i)}{\sqrt{\sigma_u^2 + 1}}\right)\right]^{y_i} \times \left[\Phi\left(\frac{w^*(\text{treat}_{z,i}, d_i)}{\sqrt{\sigma_u^2 + 1}}\right)\right]^{1-y_i}$$

with $w^* = w^*(\text{treat}_{z,i}, d_i)$ implicitly defined by equation (3). Thus, the structural parameters $(\beta, \delta, \lambda, \sigma_u)$ enter the likelihood only through their impact on the cut-off w^* .

Observability Model. We model the probability that an individual's deworming status is known within the community as a logit function of incentive arm and travel distance:

$$p_{\text{Observed}}(\text{treat}_{z,i}, d_i) = \text{logit}^{-1}(\beta_z^O \text{treat}_{z,i} + \gamma_z^O \text{treat}_{z,i} \times d_i)$$

At endline each respondent i identifies r_i recognizable adults and reports knowing the deworming status of $k_i \leq r_i$ of them. Conditional on treatments, $(\text{treat}_{z,i}, d_i)$, we treat k_i as binomial:

$$\mathcal{L}(p_{\text{Observed}} | k_i, r_i) = \binom{r_i}{k_i} \times p_{\text{Observed}}(\text{treat}_{z,i}, d_i)^{k_i} \times (1 - p_{\text{Observed}}(\text{treat}_{z,i}, d_i))^{r_i - k_i}.$$

mean through $\Delta(\cdot)$ permitting identification here.

The logit specification guarantees $p_{\text{Observed}} \in [0, 1]$, a property that disciplines counterfactual simulations in the structural model. The binomial likelihood naturally up-weights respondents who recognize more community members (r_i large) and hence provide more information about observability.

Preference (WTP) Model. We use the bracelet–calendar preference experiment to discipline the private value parameters of the two incentives. The elicitation had two stages. First, each respondent chose a gift—bracelet or calendar. Second, the enumerator offered a random cash amount $m_i \in [0, 1]$ (USD) to switch to the alternative gift.

We adopt a random utility framework and denote the latent utilities of a calendar and a bracelet by $U_{c,i}$ and $U_{b,i}$, respectively. Because the survey reveals no systematic difference in private valuations for bracelet and calendar across distance (Table B15), we model the difference in private value with a single parameter ψ :

$$\begin{aligned} U_{c,i} &= \psi + \varepsilon_{c,i} \\ U_{b,i} &= \varepsilon_{b,i} \end{aligned}$$

where $\varepsilon_{c,i} - \varepsilon_{b,i} \sim \mathcal{N}(0, 1)$. Hence the probability of initially choosing a calendar is

$$P(c_i = 1) = \Phi(\psi).$$

Conditional on that choice, the respondents accepts the cash offer m_i and switches ($s_i = 1$) if $m_i > \psi$. Thus

$$P(c_i = 1, s_i = 1) = \Phi(\psi) \times \Phi(m_i - \psi),$$

and analogous expressions hold for the other three outcomes, yielding a two-stage probit likelihood detailed in Appendix G. In the joint model we link the survey-based preference parameter ψ to the private benefit coefficients in (2) through

$$\beta_{\text{bracelet}} = \beta_{\text{Calendar}} + \rho\psi,$$

where ρ maps the \$1 valuation interval into utility units comparable with distance.

Joint Likelihood. Let θ denote the full vector of parameters. The joint likelihood can be written as the product of the three submodel likelihoods:

$$\begin{aligned} \mathcal{L}(\theta | y_i, \text{treat}_{z,i}, d_i, r_i, k_i, c_i, s_i) &= \\ \mathcal{L}(\beta_z, \delta, \lambda, \rho, \sigma_u | \beta_z^O, \gamma_z^O, \psi; y_i, \text{treat}_{z,i}, d_i) \times \mathcal{L}(\beta_z^O, \gamma_z^O | r_i, k_i) \times \mathcal{L}(\psi | c_i, s_i) \end{aligned} \quad (4)$$

so that parameter draws must simultaneously rationalize take-up, observability beliefs, and gift choices.

Identification. We now outline the sources of identification, working from the preference and observability submodels back to the deworming parameters.

Private valuation ψ . The two-stage willingness to pay (WTP) experiment in Control communities identifies the difference in purely private utility between the calendar and the bracelet. Respondents' initial gift choices, together with their switching decisions at random cash offers, pin down ψ free of any signaling value.

Observability parameters β_z^O and γ_z^O . Holding distance fixed, cross-incentive arm variation in reported observability identifies β_z^O . Within an incentive arm, variation across distance identifies γ_z^O .²⁶

Noise parameter σ_u . Variation in take-up across distance and incentive arms, together with $v_i \sim \mathcal{N}(0, 1)$ and $u_i \sim \mathcal{N}(0, \sigma_u^2)$ independent (so $w_i \equiv v_i + u_i \sim \mathcal{N}(0, 1 + \sigma_u^2)$), identifies σ_u . For the normal mixture, the type inference term is

$$\Delta(w) = \frac{\phi(x)}{\sqrt{1 + \sigma_u^2} \Phi(x) [1 - \Phi(x)]}, \quad x \equiv \frac{w}{\sqrt{1 + \sigma_u^2}},$$

where ϕ and Φ denote the standard normal pdf and cdf.²⁷ Figure 2 plots $\Delta(w^*(d))$ for several values of σ_u . When σ_u is small, $\Delta(\cdot)$ varies sharply with distance, effectively acting as a non-linear control function in the take-up probit and facilitating the separation of the social image component from private benefits. As $\sigma_u \rightarrow \infty$, the noise term dominates w , the density flattens, and $\Delta(w^*)$ approaches a constant; consequently the social multiplier tends to one, and the social image and private components become observationally equivalent. Under the fixed-point condition $1 - \mu'(d) \Delta(w^*(d)) > 0$ (Benabou and Tirole 2025), the equilibrium is unique for all finite σ_u .

Remaining parameters. Given $\psi, \sigma_u, \beta_z^O, \gamma_z^O$, the cost coefficient δ is identified by distance-induced shifts in take-up, λ by differences in social image returns across observability levels, and the private benefit coefficients by differences in take-up across treatment arms.

Estimation. Our Bayesian procedure has two steps. First, we draw a parameter vector θ from its prior, compute the implied objects such as p_{Observed} and the WTP parameter ψ , solve the fixed-point equation in (3), and construct the vector of cut-off types $w^*(d)$ for every community distance d . We then evaluate, in turn, the three conditional likelihoods—preference, observability, and deworming—and multiply them to obtain the joint likelihood in (4) at that draw. Second, we update θ by sampling from the joint posterior

²⁶Because the full structural model conditions on p_{obs} , it is invariant to rescaling of (β_z^O, γ_z^O) ; we normalize the Control arm to zero.

²⁷A derivation of this expression, and an analogous expression with bounded v , appears in Appendix L.

with Hamiltonian Monte Carlo, a gradient-based Markov-chain algorithm implemented in Stan ([Stan Development Team 2023](#)).

Weakly informative priors (Appendix H) regularize the sampler and steer it away from regions where the fixed-point routine fails to converge, eliminating the need for ad-hoc box constraints common to numerical optimization. The Bayesian framework also lets us propagate parameter uncertainty directly into the policy exercise of Section 7: at each posterior draw we solve the planner’s allocation problem and thus obtain a full posterior distribution over optimal site allocations.

Before conditioning on the data, we examine the prior predictive distribution: we draw parameter vectors from the priors, simulate implied outcomes (e.g., the social multiplier), and inspect whether these simulations align with economic intuition. Grey bands in subsequent figures display these prior predictive ranges; for instance, our prior over the social multiplier is centered on one (non-existence) and deliberately diffuse.

6.1 Structural Results

6.1.1 Average Treatment Effects

Deworming Take-up and Observability. Tables B17 and B18 report the posterior means from the structural model. These estimates closely track the reduced form estimates in Tables B10 and B6, which use the continuous distance measure.²⁸

Take-up. Distance reduces participation in the control arm by 15 percentage points, compared to 17.3 points in the reduced form. The bracelet raises take-up by 7.3 points (reduced form: 8.5). Both Ink and Bracelet have larger effects in Far than in Close communities, with Far–Close differences of 3.9 (95% CrI 1.3–6.7) and 4.7 (2.4–7.5) points, slightly below the reduced form gaps of 5.6 and 7.6 percentage points. For Calendar the Far–Close difference is essentially zero (0 points, CrI –0.01–0.03), so the bracelet–calendar double difference remains sizable at 3.1 points (reduced form: 4.7).

Observability. In the control group, distance lowers knowledge of peers’ deworming status by 9 points (reduced form: 12.2). Bracelet increases observability by 9.3 points (95% CrI 5.4–13.2). Ink and Bracelet again yield larger gains in Far than in Close communities—11.5 and 9.9 points, respectively—while the calendar differential is only 3.6 points. The bracelet–calendar Far–Close double difference is positive (8 points, CrI 3.5–13.8) and, as in the reduced form (17.3 points), excludes zero.

Overall, the structural estimates are slightly more conservative but more precise than their reduced form counterparts, reflecting the additional structure imposed by the model.

Social Image Returns versus Private Benefits. Table B19 decomposes each incentive’s effect into a social image component and a private benefit component via two

²⁸Table B16 reports posterior means and 95% credible intervals for the full parameter vector.

counterfactual exercises. To isolate social image returns we, fix the private utility term at its Control value. To isolate private benefits, we hold the social image term at the Control value.

Bracelet and Ink generate the largest social image returns, raising take-up by 2.9 percentage points (95% CrI 1.2–4.9) and 1.8 points (0.4–3.6), respectively, when pooling across distance conditions. Calendar’s social image effect is essentially zero (0.4 points, –0.8–1.7). The Far–Close differentials in social image returns are 3.3 points (1.2–6.2) for Bracelet and 2.9 points (1.1–5.4) for Ink, versus only 1.1 points (–0.0–2.7) for Calendar, consistent with signaling incentives having greater reputational value at farther distances.

Finally, we estimate that Ink carries a sizable negative private valuation: –4.4 points (–7.3 to –1.5). This dislike of thumb-marking, combined with higher travel cost in Far communities, pushes the cut-off type upward; only individuals with very high intrinsic motivation v deworm when Ink is the signal.²⁹

6.1.2 The Social Multiplier

We assess whether the negative effect of distance on take-up is mitigated or amplified when social image concerns are taken into account. Figure 3 reports the estimated social multiplier, $\tilde{SM}(\text{treat}_z, d)$, for each incentive arm—Control, Ink, Bracelet, and Calendar—and its decomposition into type inference and observability components (see the figure notes for details).

Mitigation. When $\tilde{SM}(\text{treat}_z, d) < 1$, distance has a smaller proportional impact on take-up than under the no-observability benchmark (dashed black line). For Ink the multiplier is 0.90 at $d = 0.75$ km, so a 10% increase in travel cost yields roughly a 9% local decline in participation.³⁰ Bracelet shows a similar pattern: $\tilde{SM} = 0.93$ at $d = 0$ km, falling to 0.82 at $d = 2.5$ km. Consistent with the Bénabou–Tirole framework, mitigation for Ink is driven mainly by a steeper type inference schedule—i.e., a larger slope of $\Delta(w^*)$ (higher Δ') with little change in observability—whereas for Bracelet an increase in observability $\mu(d)$ accounts for a sizable share of the reduction.

Amplification. In the Control and Calendar arms, $\tilde{SM}(\text{treat}_z, d) > 1$: distance therefore has a larger proportional impact on take-up than under the no-observability benchmark. Although the type inference term remains positive ($\Delta'(w^*) > 0$), the decline in observability dominates ($\mu'(d) < 0$). Evaluated at $d = 0$ km, the distance semi-elasticity is about 1.2× the private cost effect—i.e., a one-kilometer increase in distance induces roughly a 20% larger decline in take-up than in the no-observability benchmark.

In sum, greater distance raises social image returns in the Bracelet and Ink arms,

²⁹Although the negative private valuation of ink was not anticipated ex-ante, it is analytically useful: together with distance cost, it shifts the cut-off w^* further to the right, strengthening identification of the social image component.

³⁰In the probit, the marginal effect is $\partial \bar{y} / \partial d = f_w(w^*) \delta \tilde{SM}$, so 0.90 is a local proportional effect.

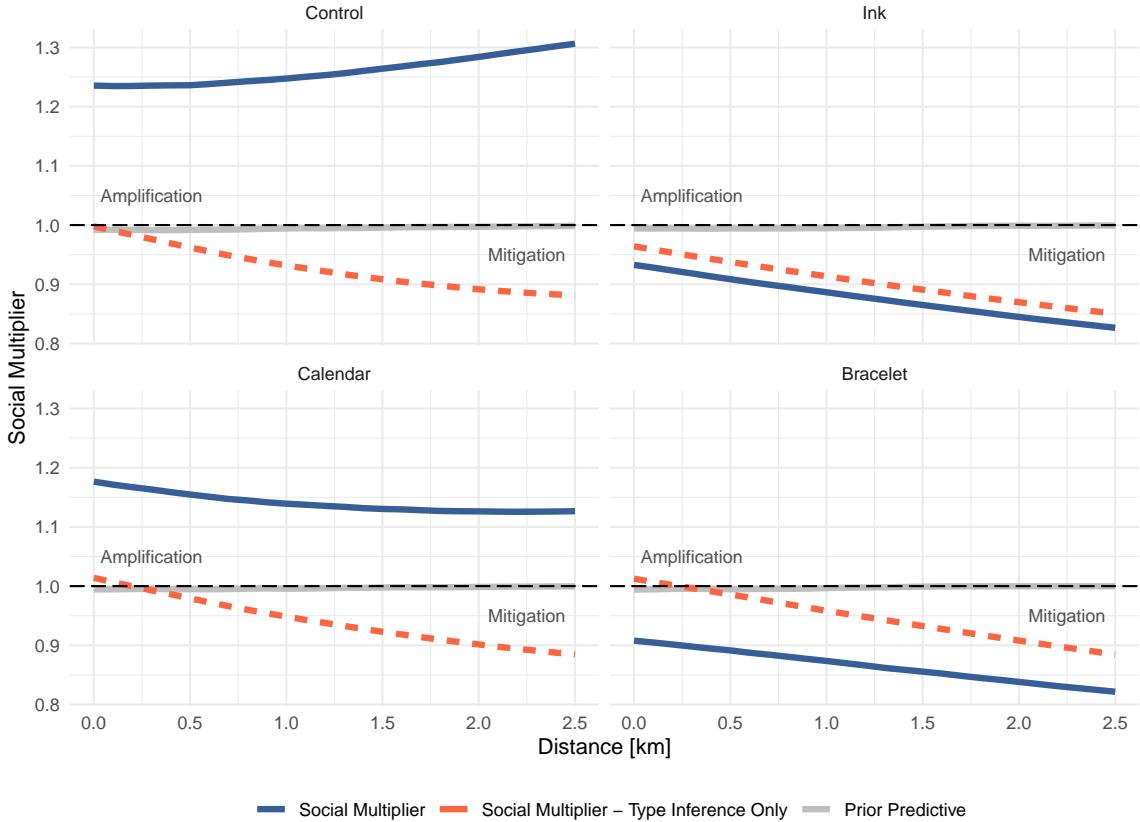


Figure 3: Estimated Social Multiplier and its Decomposition

Notes: The figure plots the social multiplier implied by the structural model and decomposes it into its two components. The solid blue curve is the total multiplier. The red dotted curve holds observability fixed and therefore isolates the type inference term $\Delta'(w^*(\text{treat}_z, d))$; the vertical gap between the blue and red curves equals the contribution of changes in observability $\mu(\text{treat}_z, d)$. The dashed black line depicts the no-observability counterfactual, $\mu(\text{treat}_z, d) = 0$, under which the multiplier equals one. The gray curve shows the prior-predictive median, included to demonstrate that the posterior shape is driven by the data rather than by the choice of priors or functional form assumptions. All curves represent median posterior estimates from the Bayesian structural model.

making individuals less sensitive to travel cost; by contrast, declining observability in the Control and Calendar arms makes them more sensitive, thereby amplifying the distance penalty.

6.1.3 Robustness

Our main specification measures travel cost by the distance from the community centroid to the treatment site and assumes that adults form expectations on that publicly observable metric. To gauge sensitivity to alternative information sets we estimate three variants (online Appendix I).

First, we keep the centroid distance when computing the equilibrium cut-off type but let the private cost in each household’s utility depend on its own household-to-site distance. Second, we adopt a full information model in which every household is assumed to know the exact distance of every other household; algebraically the structure is unchanged, but the fixed point is now solved for all 9,805 adults rather than at the community level. Third, we replicate the main specification after excluding spatially dispersed clusters—those whose mean squared household distance from the centroid exceeds 0.5 km^2 —where the centroid may be a poor proxy for individual cost.

Results are reported in Tables I1–I3. Across all three variants the estimated effects for Bracelet, Calendar, and the Control arm are essentially unchanged. The Ink estimates are somewhat less stable: when social image returns are computed at the centroid but costs are household specific, the Ink effect is negative even in Far communities (though still less negative than in Close ones); under full information the Ink effect remains negative and becomes flat across distance. Dropping dispersed clusters brings the Ink coefficients back in line with the main specification. Taken together, these robustness checks show that the core findings—mitigation for high visibility incentives and amplification for low visibility incentives—are insensitive to alternative cost measures or information sets.

Finally, online Appendix J investigates the implications of alternative distributional assumptions for the composite type, $w = v + u$. Simulations with bimodal densities lead to similar qualitative conclusions regarding the social multiplier as the Gaussian, suggesting results are not driven by the normality assumption.

7 Spatial Allocation with a Social Multiplier

We now show how the estimated structural parameters, together with geocoded data on communities and potential points of treatment (PoTs), can guide a policymaker’s spatial allocation decisions. In particular, we ask how awareness of the social multiplier changes (i) the minimum number of PoTs needed to reach a coverage target and (ii) the locations

of those PoTs.³¹

7.1 Policymaker's Optimization Problem

We model a policymaker who minimizes the number of funded PoTs subject to attaining an expected deworming rate at least as high as the experimental control mean, $\bar{T}_{\text{control}}^{\exp}$.³²

With prevalence data unavailable—and village density fairly homogeneous (78 percent of communities lie within one standard deviation of the mean spatial density, 95 percent within two; Table B20)—the policymaker adopts a uniform coverage target.³³ Because policymakers typically fix such a target, the “fewest-PoTs-for-target-coverage” formulation both mirrors practice and avoids specifying a PoT production function; any resources saved by not opening additional sites can be diverted to other health programs. Maximizing coverage for a fixed PoT budget would coincide with this solution only under additional regularity conditions—an analogue to strong duality in linear or convex optimization—but those conditions fail for the nonlinear take-up function we estimate. We therefore adopt the minimization problem as the benchmark.

There are $n = 144$ communities (i) and $m = 1,451$ candidate PoTs (j). Let $x_{ij} \in \{0, 1\}$ indicate whether community i is served by PoT j and $y_j \in \{0, 1\}$ whether PoT j is funded. For a given parameter draw θ the planner solves the integer program

$$\begin{aligned} \min_{x,y} \quad & \sum_{j=1}^m y_j \\ \text{s.t.} \quad & \sum_{i=1}^n \sum_{j=1}^m T(\text{treat}_z, d_{ij}; \theta) x_{ij} \geq \bar{T}_{\text{control}}^{\exp}, \\ & \sum_{j=1}^m x_{ij} = 1, \quad x_{ij} \leq y_j, \quad x_{ij} d_{ij} \leq D, \quad \forall i, j. \end{aligned} \tag{5}$$

Here $T(\text{treat}_z, d_{ij}; \theta)$ is the expected take-up under treatment z at community-PoT distance d_{ij} for parameter draw θ . We set $D = 3.5$ km to cap total travel distance.³⁴ The target $\bar{T}_{\text{control}}^{\exp}$ is the expected coverage that would arise under the experiment’s distance assignments if every village were in the Control arm:

³¹See Section 3.7 and Appendix Table E1 for the preregistration summary and mapping. This allocation exercise was not prespecified and is presented as a theory driven extension that applies the estimated parameters to a feasible siting problem.

³²Formally, $\bar{T}_{\text{control}}^{\exp} = E_{\Theta}[\sum_{i=1}^n T(\text{control}, d_{ij}; \theta) x_{ij}^{\text{experiment}}]$, the posterior expectation of take-up implied by the experiment’s distance assignments ($x_{ij}^{\text{experiment}}$), with the expectation taken over parameter draws $\theta \sim \Theta$.

³³If prevalence data were available, the policymaker could weight communities by disease burden.

³⁴The cap balances two goals: (i) limiting extrapolation beyond the 2.5 km range observed in the experiment, and (ii) allowing enough scope to pool nearby villages under a single PoT. Extending D to 10 km changes results negligibly (Appendix Table K1); the posterior assigns virtually no density to positive demand beyond 5 km.

$$\bar{T}_{\text{control}}^{\exp} = E_{\Theta} \left[\sum_{i=1}^n T(\text{control}, d_{ij}; \theta) x_{ij}^{\text{experiment}} \right].$$

Because θ is treated as a random variable, we solve problem (5) for 200 posterior draws and aggregate the resulting solutions to obtain the policymaker's posterior distribution of the optimal PoT count $\gamma(\theta)$. Solving at every draw propagates estimation uncertainty that a single plug-in estimate would understate.

7.1.1 Counterfactual Simulations

We decompose expected take-up into counterfactual components by holding either the private benefit or social image index fixed at Control values. For exposition, we label these contributions ‘private benefit’, $B(\cdot)$, and ‘social image return’, $\mu\Delta$:

$$T(\text{treat}_z, d_{ij}; \theta) = \underbrace{B(\text{treat}_z, d_{ij})}_{\text{private benefit}} + \underbrace{\mu(\text{treat}_z, d_{ij}) \Delta(w^*(\text{treat}_z, d_{ij}))}_{\text{social image return}}.$$

Holding the private benefit fixed, we vary the social image term to assess how alternative signaling environments affect the number—and spatial pattern—of PoTs required to achieve the coverage target.

Knowledge of the Social Multiplier

Figure 4 contrasts the demand curves a policymaker would perceive under different beliefs about the interplay between distance and social image returns. Table B21 reports the corresponding numbers of PoTs.

We first consider a policymaker who recognizes that social norms matter ($\mu > 0$) but—having calibrated demand only at short distance (0.5 km)—assumes that social image returns are constant in d . The resulting demand curve (blue line) leads her to allocate 100 PoTs in the Control scenario (Panel B, row 1). When observability is raised with bracelets, she would open 96 PoTs (row 2).

The yellow and red lines correct this misspecification by incorporating the endogenous change in image returns. Under Control observability (yellow), observability declines with distance, demand falls more steeply, and amplification sets in; the policymaker now needs 107 PoTs, 7% more than under the constant returns belief (row 3). Under Bracelet observability (red), image returns rise with distance, flattening the demand curve (mitigation); the required number of PoTs drops to 90, about 6% fewer than in the constant-returns case (row 4). Recognizing the social multiplier therefore shifts resources toward settings where social image concerns render demand more inelastic.

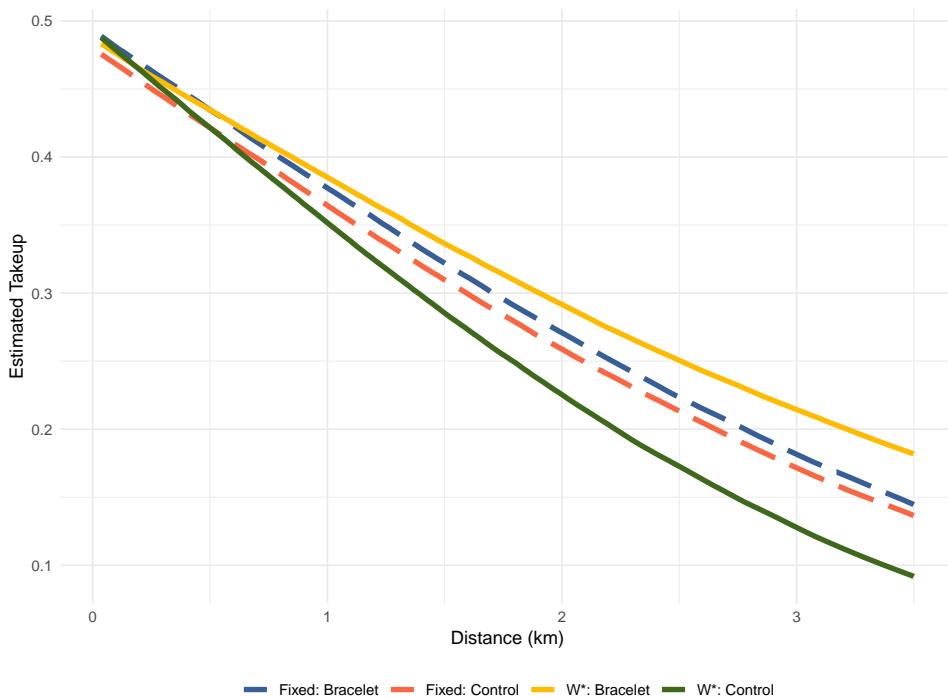


Figure 4: Demand under Alternative Assumptions about Social Image Returns

Notes: The figure plots deworming take-up as a function of travel distance under three informational assumptions. Blue—the policymaker holds social image returns fixed at their 0.5 km value, ignoring any distance dependence. Yellow—Control observability with distance-dependent image returns; demand therefore falls more steeply with distance (amplification). Red—Bracelet observability with distance-dependent image returns; rising image returns flatten the curve (mitigation). All curves correspond to posterior median demand and are evaluated at the Control private benefit parameters to isolate the role of social image effects.

Spatial Allocation of PoTs

Figure 5 and Figure A12 compare PoT locations and community travel distances under three scenarios: (i) the experiment’s original design, (ii) the policymaker’s optimum with Control observability, and (iii) the optimum with Bracelet observability.

Experimental benchmark. Treating every community as if it were in the Control arm (left panel of Figure A12) and keeping the original Close/Far assignment yields an expected deworming rate of 32.9 percent, an average walking distance of 1.20 km, and a distance distribution that is distinctly bimodal (green curve in Figure 5).

Optimal allocation with Control observability. Solving the integer program (5) for each posterior draw and reporting the median solution (center panel) links every village to its nearest feasible PoT. Whenever the 3.5 km cap allows, several villages are pooled at a single site, so the number of funded PoTs falls from 144 to 107. Mean travel distance increases only modestly, to 1.27 km (yellow curve). Because social image returns decay with distance, the policymaker counterbalances this by placing many PoTs relatively close to the communities they serve.

Optimal allocation with Bracelet observability. Under Bracelet observability (right panel), the social multiplier mitigates travel cost, allowing the same coverage target with only 90 PoTs—about 16 percent fewer than under Control—while the mean travel distance rises to 1.68 km (purple curve). Fewer communities remain within 0.5 km, and many are now around 3 km, raising average distance by 11 percent without lowering take-up. The comparison illustrates how a policymaker can leverage this mitigation to design health services more efficiently.

Need for Experimentation at Scale

Finally, consider a policymaker who is unaware of the social multiplier and therefore sets $\mu = 0$. This situation mirrors a pilot that distributes travel vouchers privately, revealing only how take-up falls with travel distance. Because the incentive is not publicly observable, the pilot cannot shift community beliefs or generate the equilibrium feedback captured by the social multiplier and would only reveal private cost sensitivity. Relying on such data, the policymaker would view demand as highly elastic: even if all 144 candidate PoTs were opened and each community were assigned to its nearest site (average distance 0.58 km), projected take-up is only 17.4 percent (Table B21, Panel B, row 5). In practice, once observability and social inferences are accounted for, coverage can be substantially higher with fewer sites.

These results underscore that identifying—and ultimately leveraging—social image motives requires experimentation at scale, with incentives made sufficiently public to shift equilibrium beliefs.

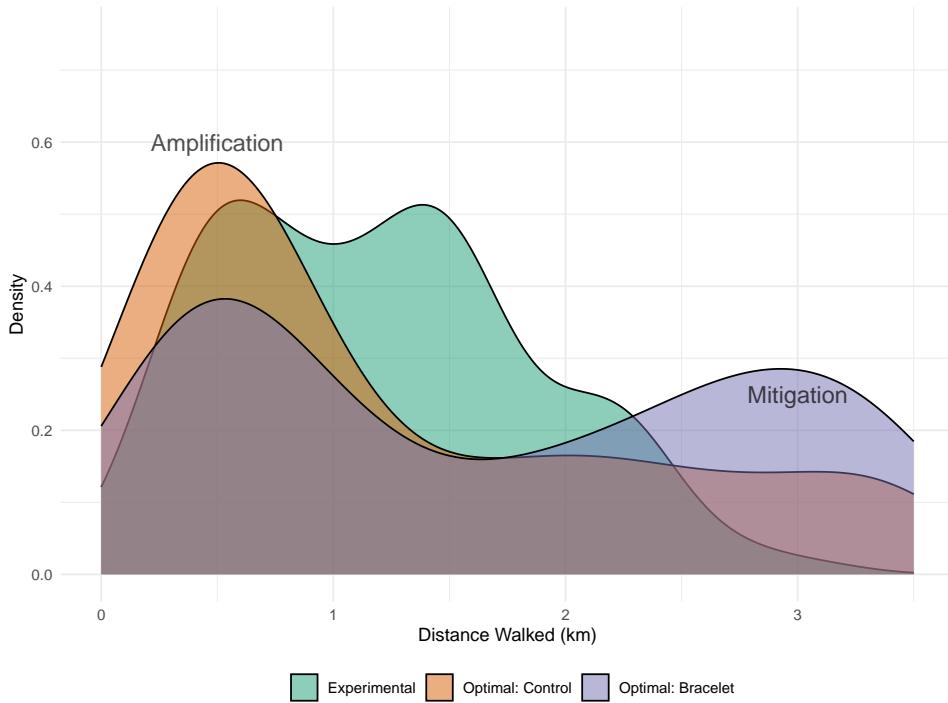


Figure 5: Distribution of Community–PoT Distances under Alternative Allocations

Notes: Each curve plots the posterior median density of community–to-PoT distances under one of three scenarios. Green — experimental assignment (randomized Close /Far design, Control incentive). Yellow — policymaker’s optimal allocation when observability follows the Control schedule; the objective is to meet the coverage target with the fewest PoTs. Purple — policymaker’s optimal allocation under Bracelet observability, solved with the same objective. Densities are based on the policymaker’s allocation at the posterior median of the structural parameters.

8 Conclusion

Leveraging social image concerns is receiving increasing attention among policymakers as a strategy to enhance the provision of public goods and promote socially desirable behaviors. This study demonstrates that economic incentives do not operate in isolation: private costs and social image concerns interact in equilibrium, with implications for policy design.

We draw three lessons. First, a routinely used non-monetary lever, distance, interacts with social image in ways that materially change demand for preventive care. As expected, moving sites farther reduces take-up, but when actions are publicly observable the associated image payoff can compensate for extra walking. In practice, higher observability softens the travel cost penalty, whereas lower observability amplifies it.

Second, observability can be directly shaped by policy. Simple, low-cost signals make take-up more visible, especially where baseline take-up is low, and can flip the social multiplier from amplification (distance further depresses demand) to mitigation (distance's negative effect is dampened). In our setting, increasing the observability of actions broadened awareness of others' take-up and attenuated the sensitivity of take-up to travel distance.

Third, recognizing these interactions reshapes optimal spatial design. A policymaker who calibrates demand at short range and treats image benefits as distance-invariant will misallocate resources. Accounting for the social multiplier shows that public observability can substitute for proximity: with public signals, the same coverage can be achieved with fewer, more widely spaced sites.

Several open questions remain. First, our data capture only short-run responses: whether social image multipliers persist, fade, or grow as communities learn and the novelty of the signal dissipates is unknown. Second, larger experiments that randomize distances across a greater number of clusters would permit non-parametric estimation of the multiplier, relaxing the functional form assumptions we impose here. Third, a full welfare analysis would weigh cost savings against additional travel time and any stigma borne by non-participants. Future work should quantify these distributional implications.

Taken together, the results constitute a proof of concept: while broader external validity remains to be established, they highlight that overlooking how private costs reshape social image incentives can lead to misallocation and leave untapped opportunities for using public observability to expand program reach at lower fiscal cost.

References

- Roland Benabou and Jean Tirole. Laws and norms. *Journal of Political Economy*, 2025. doi: 10.1086/738343. URL <https://doi.org/10.1086/738343>.
- Dan Ariely, Anat Bracha, and Stephan Meier. Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American Economic Review*, 99(1):544–55, March 2009. doi: 10.1257/aer.99.1.544. URL <https://www.aeaweb.org/articles?id=10.1257/aer.99.1.544>.
- Leonardo Bursztyn and Robert Jensen. Social Image and Economic Behavior in the Field: Identifying, Understanding, and Shaping Social Pressure. *Annual Review of Economics*, 9(1):131–153, 2017. doi: 10.1146/annurev-economics-063016-103625. URL <https://doi.org/10.1146/annurev-economics-063016-103625>. _eprint: <https://doi.org/10.1146/annurev-economics-063016-103625>.
- Leonardo Bursztyn, Bruno Ferman, Stefano Fiorin, Martin Kanz, and Gautam Rao. Status Goods: Experimental Evidence from Platinum Credit Cards*. *The Quarterly Journal of Economics*, 133(3):1561–1595, August 2018. ISSN 0033-5533. doi: 10.1093/qje/qjx048. URL <https://doi.org/10.1093/qje/qjx048>.
- Arun G. Chandrasekhar, Cynthia Kinnan, and Horacio Larreguy. Social Networks as Contract Enforcement: Evidence from a Lab Experiment in the Field. *American Economic Journal: Applied Economics*, 10(4):43–78, October 2018. ISSN 1945-7782. doi: 10.1257/app.20150057. URL <https://www.aeaweb.org/articles?id=10.1257/app.20150057>.
- Stefano DellaVigna, John A. List, Ulrike Malmendier, and Gautam Rao. Voting to Tell Others. *The Review of Economic Studies*, 84(1):143–181, January 2017. ISSN 0034-6527. doi: 10.1093/restud/rdw056. URL <https://doi.org/10.1093/restud/rdw056>.
- Anne Karing. Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone*. *The Quarterly Journal of Economics*, page qjae025, 08 2024. ISSN 0033-5533. doi: 10.1093/qje/qjae025. URL <https://doi.org/10.1093/qje/qjae025>.
- Emily Breza and Arun G. Chandrasekhar. Social networks, reputation, and commitment: evidence from a savings monitors experiment. 87(1), 2019. ISSN 175-216.
- Ruixue Jia and Torsten Persson. Individual vs. social motives in identity choice: Theory and evidence from china. Technical report, National Bureau of Economic Research, 2019.
- Timothy Besley, Anders Jensen, and Torsten Persson. Norms, enforcement, and tax evasion. *Review of Economics and Statistics*, 105(4):998–1007, 2023.

Edward L. Glaeser, Jose A. Scheinkman, and Bruce I. Sacerdote. The social multiplier. *Journal of the European Economic Association*, 1(2/3):345–353, 2003. ISSN 15424766, 15424774. URL <http://www.jstor.org/stable/40005184>.

Judd B. Kessler. Announcements of Support and Public Good Provision. *American Economic Review*, 107(12):3760–3787, December 2017. ISSN 0002-8282. doi: 10.1257/aer.20130711. URL <https://www.aeaweb.org/articles?id=10.1257/aer.20130711>.

Ricardo Perez-Truglia and Guillermo Cruces. Partisan Interactions: Evidence from a Field Experiment in the United States. *Journal of Political Economy*, 125(4):1208–1243, August 2017. ISSN 0022-3808. doi: 10.1086/692711. URL <https://www.journals.uchicago.edu/doi/full/10.1086/692711>. Publisher: The University of Chicago Press.

Petra E. Todd and Kenneth 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*, 96(5):1384–1417, December 2006. doi: 10.1257/aer.96.5.1384. URL <https://www.aeaweb.org/articles?id=10.1257/aer.96.5.1384>.

Joseph P. Kaboski and Robert M. Townsend. A structural evaluation of a large-scale quasi-experimental microfinance initiative. *Econometrica*, 79(5):1357–1406, 2011. doi: <https://doi.org/10.3982/ECTA7079>. URL <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA7079>.

Esther Duflo, Rema Hanna, and Stephen P. Ryan. Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4):1241–78, June 2012. doi: 10.1257/aer.102.4.1241. URL <https://www.aeaweb.org/articles?id=10.1257/aer.102.4.1241>.

David Lagakos, Ahmed Mushfiq Mobarak, and Michael E. Waugh. The welfare effects of encouraging rural–urban migration. *Econometrica*, 91(3):803–837, 2023. doi: <https://doi.org/10.3982/ECTA15962>. URL <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA15962>.

WHO. Soil-transmitted helminth infections. fact sheet. 2023.

Peter J Hotez, Paul J Brindley, Jeffrey M Bethony, Charles H King, Edward J Pearce, Julie Jacobson, et al. Helminth infections: the great neglected tropical diseases. *The Journal of clinical investigation*, 118(4):1311–1321, 2008.

Edward Miguel and Michael Kremer. Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217, 2004.

M-S. Chan. The global burden of intestinal nematode infections — fifty years on. *Parasitology Today*, 13(11):438–443, 1997. ISSN 0169-4758. doi: [https://doi.org/10.1016/S0169-4758\(97\)01144-7](https://doi.org/10.1016/S0169-4758(97)01144-7). URL <https://www.sciencedirect.com/science/article/pii/S0169475897011447>.

Amrita Ahuja, Sarah Baird, Joan Hamory Hicks, Michael Kremer, Edward Miguel, and Shawn Powers. When should governments subsidize health? the case of mass deworming. *The World Bank Economic Review*, 29(suppl_1):S9–S24, 2015.

Evidence Action. Kenya national school-based deworming programme. ciff end-of-project report. 2017.

Kirill Borusyak and Peter Hull. Nonrandom exposure to exogenous shocks. *Econometrica*, 91(6):2155–2185, 2023.

Roland Bénabou and Jean Tirole. Incentives and Prosocial Behavior. *American Economic Review*, 96(5):1652–1678, December 2006. ISSN 0002-8282. doi: 10.1257/aer.96.5.1652. URL <https://www.aeaweb.org/articles?id=10.1257/aer.96.5.1652>.

Stan Development Team. Stan Reference Manual, 2023.

A Online Appendix

A Appendix Figures

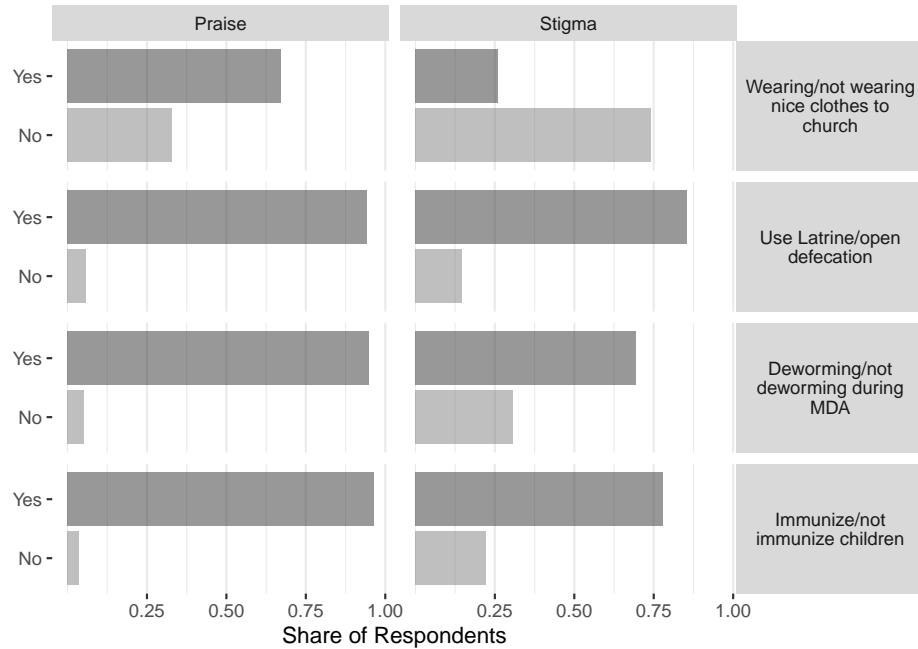


Figure A1: Social Image Concerns at Baseline

Notes: The figure reports baseline attitudes toward four behaviors. For each vignette respondents stated whether they would *praise* someone who performs the behavior and whether they would *look down on* someone who does not. All 2,056 baseline adults answered the deworming and childhood-immunization vignettes; half were randomly assigned an open-defecation vignette and half a church-attendance vignette.

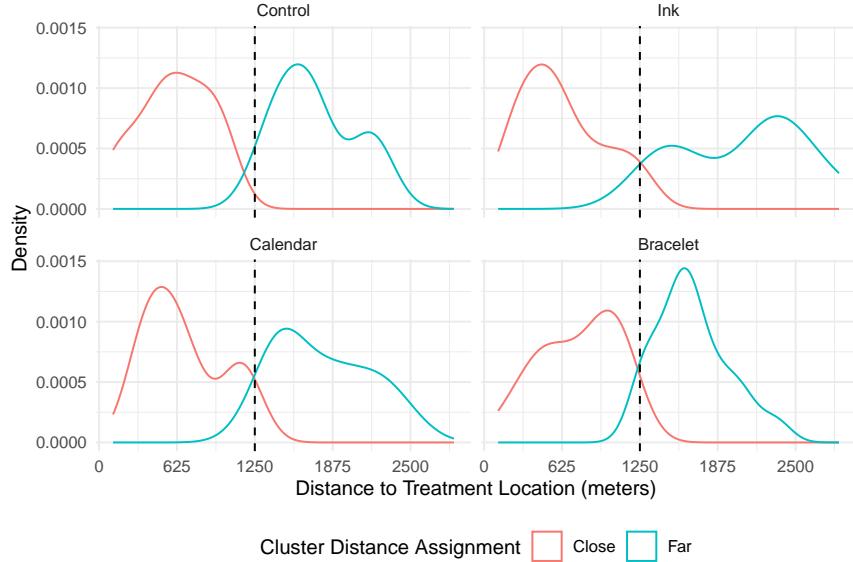


Figure A2: Random Distance to Treatment Location

Notes: Each panel plots the kernel density of household-level distance (in meters) from the centroid of the community to its assigned point of treatment. Communities were randomly assigned to a Close (0–1.25 km, red) or Far (1.25–2.5 km, blue) travel condition; the vertical dashed line marks the 1.25 km cut-off. Panels correspond to the four signaling arms (Control, Ink, Calendar, Bracelet). Across all 144 study communities, the randomization generated an average Close–Far distance difference of 1.02 km. Minor overlap of the two distributions reflects household dispersion and small, logistics driven adjustments in site placement.

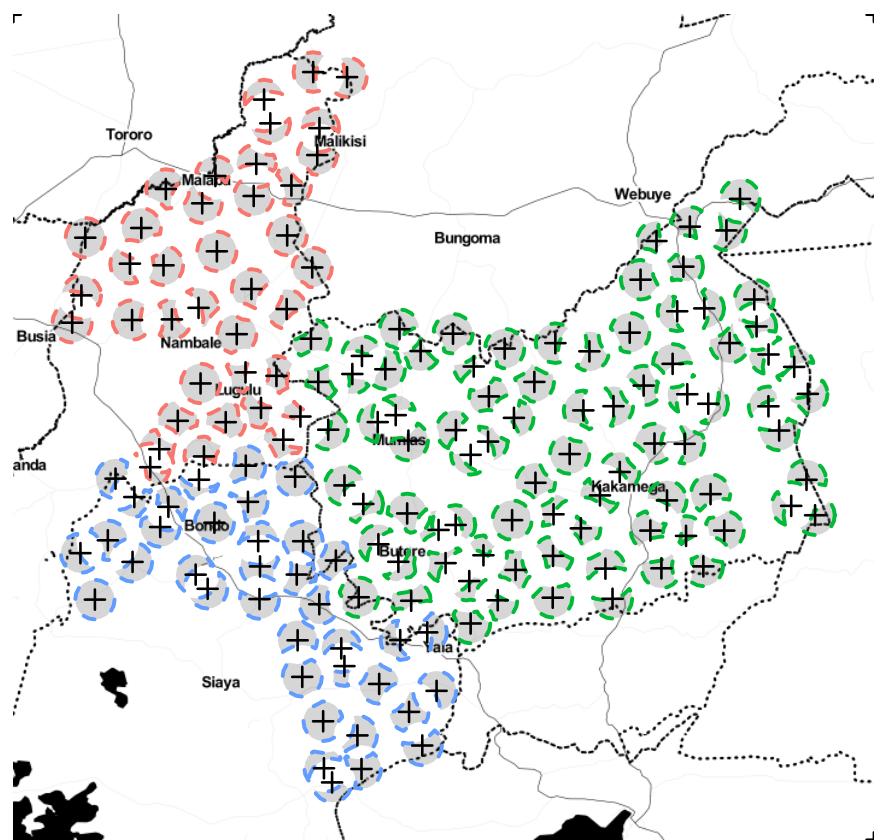


Figure A3: Site Selection

Notes: This map shows the selected deworming points of treatment (black crosses) and the 2.5-km catchment circles that define the set of communities eligible for assignment to each point of treatment.



Figure A4: Bracelet

Notes: Green silicone bracelet distributed by CHVs at the point of treatment. The Swahili text reads “Tibu minyoo: boresha afya ya jamii yako”, translated as “Treat worms: improve the health of your community”



Figure A5: Calendar

Notes: Wall calendar distributed at the point of treatment, used as an active control for public observability. The calendar contains no deworming message.

1. Deworming is not only for children because everyone is at risk of being infected by worms or is infected but does not know.
2. Taking deworming tablets is like using a mosquito net to prevent Malaria or washing hands before eating to avoid diarrhea. You do not have to be sick or experience symptoms in order for you to get dewormed.
3. It is important to take deworming tablets every six 6 months to ensure that your body is always free of worms.
4. The government is providing free deworming tablets and all adults are encouraged to deworm themselves.
5. Deworming all adults will keep our community free from worms and those who do not deworm themselves shall put the entire community at risk, especially towards our children.
6. Remind your family members and neighbors to turn up for the free deworming medication on _____ at _____.
7. You will receive _____ for deworming yourself as a symbol of your passion towards improving the health of the members in your family and the community.

Figure A6: Information Script

Notes: This figure reproduces the standard script that Community Health Volunteers (CHVs) read when visiting households one week before the campaign. The script (i) explains that everyone is at risk of worm infection, (ii) emphasizes the community benefits of universal treatment, (iii) notes that the medication is provided free of charge by the government, (iv) specifies the treatment date and location—filled in on site—and (v) inserts the appropriate incentive for the cluster (bracelet, calendar, ink mark, or none). The language and sequencing were held constant across all 144 communities.



Figure A7: Flyers

Notes: This figure displays the four one-page flyers that CHVs handed out during the pre-deworming information visits. Each flyer—printed in Swahili—announces the free adult deworming campaign, leaves blank fields for the locally customized site (“*Wapi*”) and date (“*Lini*”), and, except in the Control arm, depicts the incentive offered upon deworming (bracelet, calendar, or ink mark). Apart from the incentive panel, wording and layout are identical across flyers to preserve comparability.

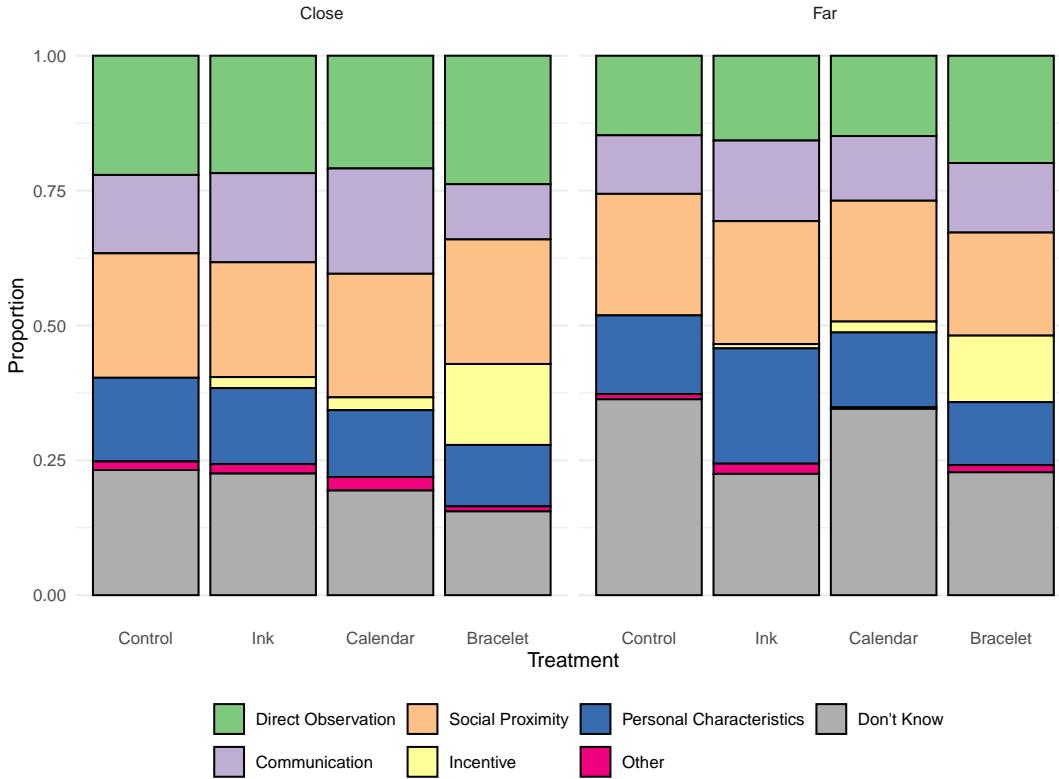


Figure A8: Reasons for the Observability or Non-Observability of Deworming Status

Notes: The figure summarizes why respondents believe other adults know their deworming status. When a respondent stated that a named peer knew whether he or she had dewormed, we asked: “Why do you think they would think that [you came or did not come for deworming]?” Verbatim answers were classified with OpenAI’s GPT-4 into four categories: (i) direct observation (“He saw me at the PoT”, “We went together”, “We never met during the campaign”); (ii) communication (“I told them,” “We talked about it at the shop”, “I told her I will not go”); (iii) social proximity (“She is my neighbor”, “We are close friends”, “We stay far from each other”); and (iv) personal characteristics (“They know I take health seriously”, “I love health things”, “I was pregnant/ sick”). Each stacked bar corresponds to one incentive arm (Control, Ink, Calendar, Bracelet) in either Close (left four bars) or Far (right four bars) communities; colors show the share of reasons in each category, while the gray segment labeled “Don’t Know” shows the fraction of peers whom the respondent believed did not know their deworming status. The analysis covers 1,512 respondents; among the 1,627 adults for whom belief data were collected, 115 gave answers that were missing or could not be classified by the model.

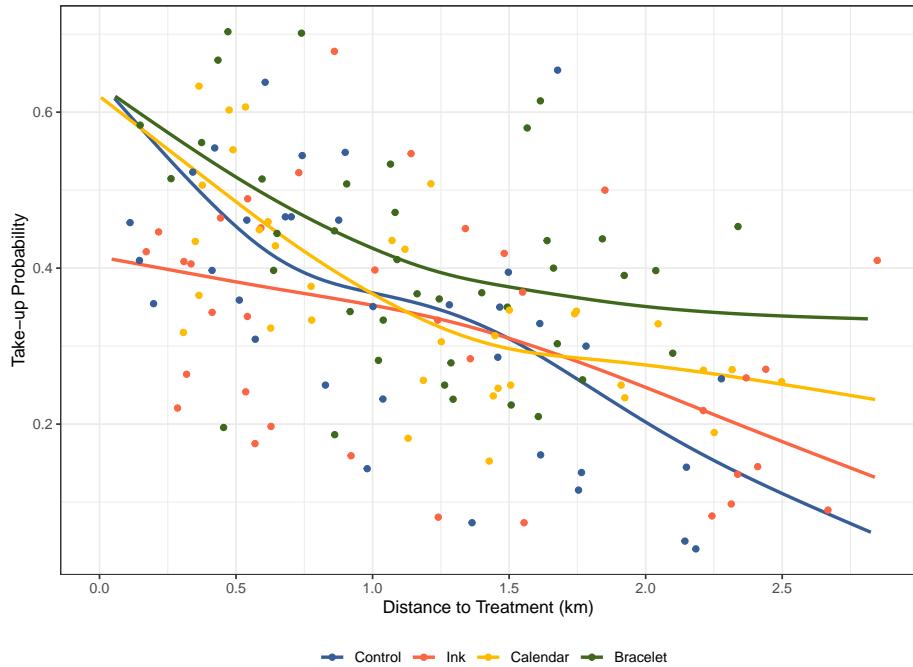


Figure A9: Take-up across Distance by Incentive Arm

Notes: This plot shows non-parametric fits of deworming take-up on distance to the point of treatment, estimated separately by assigned treatment arm (Control, Calendar, Ink, Bracelet). Each dot is a community-level mean among monitored adults; the solid curves are penalized spline estimates of the conditional expectation function. Distance is measured from the community centroid to the assigned point of treatment (km).

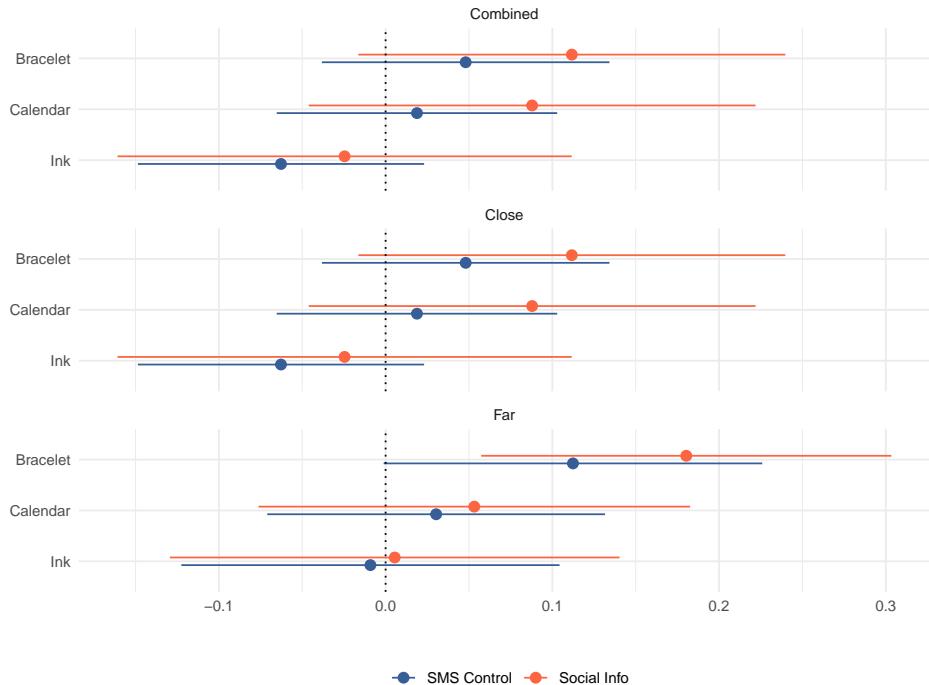


Figure A10: Treatment Effects by SMS Condition

Notes: The figure plots incentive arm treatment effects for phone owners who received Social Info SMS (red) and for phone owners in the SMS-control group (blue). Estimates are shown for the full sample and separately for Close and Far communities. The similarity of the red and blue estimates shows that the bracelet and ink incentives do not operate primarily through a reminder or social learning channel—otherwise their effects would fall in the presence of the SMS. The analysis includes 2,635 SMS recipients and 7,363 SMS-control phone owners.

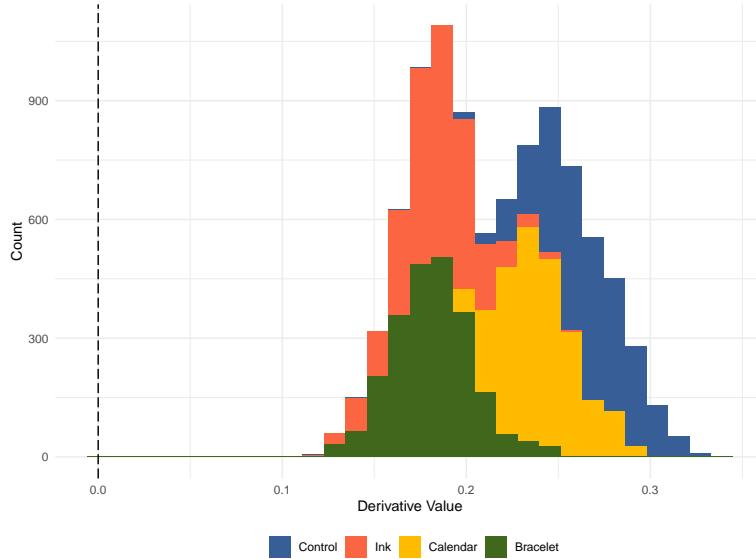


Figure A11: Posterior Distribution of the Key Derivative Ensuring a Unique Equilibrium

Notes: For every posterior draw of the structural parameters and for the range of distances used in the paper, the social-multiplier derivative $(\delta - \mu'(d) \Delta(w^*(d))) / (1 + \mu(d) \Delta(w^*(d)))$ is strictly positive. Hence the model satisfies the single-equilibrium condition $1 + \mu' \Delta(v^*) > 0$ discussed by [Benabou and Tirole \(2025, p. 6\)](#), ruling out multiple equilibria or self-sustaining norms.

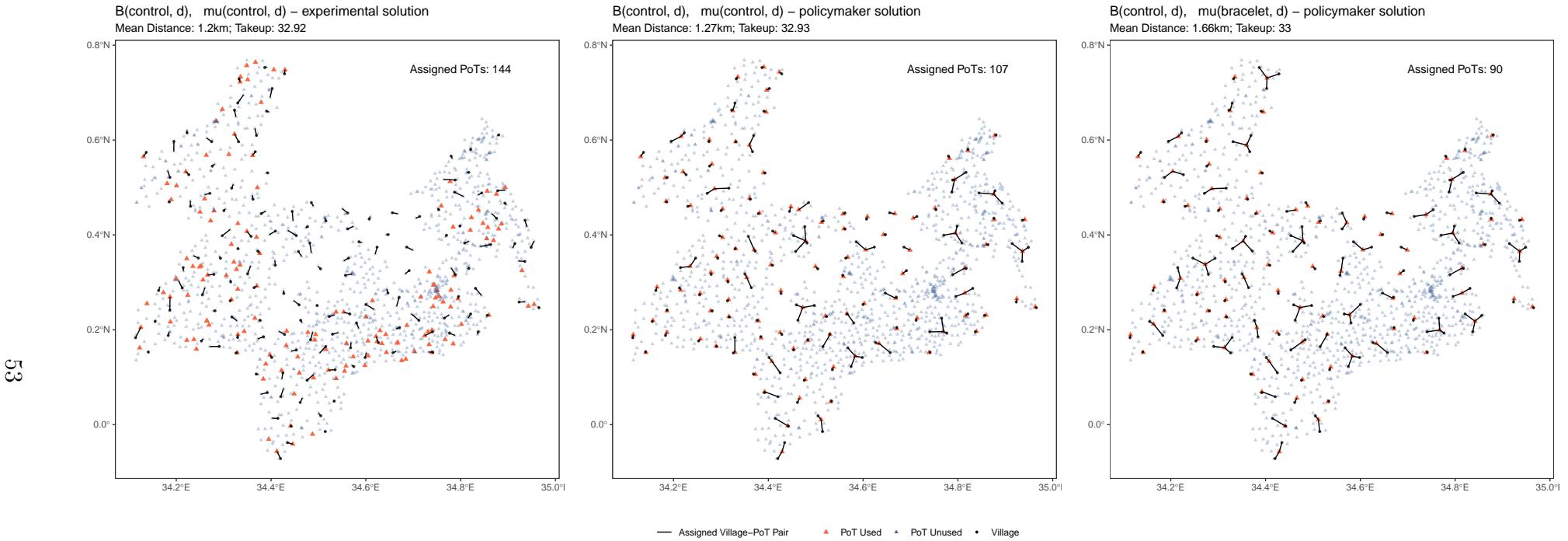


Figure A12: Optimal Placement of Treatment Sites (with Experimental Benchmark)

Notes: Each panel depicts one allocation scenario computed with the posterior-median parameters of the structural model. From left to right: (i) the experimental assignment with every village treated as if it were in the Control arm; (ii) the policymaker's optimal allocation under Control observability; and (iii) the policymaker's optimal allocation under Bracelet observability. Symbols—black dots mark the 144 villages; red triangles are funded points of treatment (PoTs); blue triangles are candidate PoTs left unused. Black line segments connect each village to its assigned PoT. In both optimal scenarios, the integer program chooses among 1,451 candidate sites subject to a 3.5 km distance cap.

B Appendix Tables

Table B1: Baseline Balance, Descriptives, and Implementation Checks

	Sample Mean	Close						Far						Joint Test	
		Con	Ink - Con	Cal - Con	Bra - Con	Bra - Cal	Joint Test	Con	Ink - Con	Cal - Con	Bra - Con	Bra - Cal	Joint Test		
Panel A: Takeup sample, N = 9,805															
Age	39.8	38.126 (0.758)	-0.714 [0.427]	-0.919 [0.286]	0.232 [0.813]	1.151 [0.208]	0.5221 [0.068]	37.995 [0.838]	-0.804 [0.403]	-0.861 [0.324]	-0.798 [0.419]	0.063 [0.945]	0.7592 [0.7592]	0.7776	
Female	0.553	0.543 (0.013)	-0.006 [0.673]	-0.002 [0.907]	-0.036 [0.028]	-0.034 [0.005]	0.0913 [0.014]	0.539 [0.414]	-0.014 [0.015]	0.002 [0.006]	-0.023 [0.182]	-0.025 [0.172]	0.4555 [0.4555]	0.2458	
Phone owner	0.751	0.816 (0.015)	0.008 [0.675]	-0.006 [0.755]	-0.005 [0.797]	0.001 [0.947]	0.9154 [0.015]	0.798 [0.998]	0 [0.307]	0 [0.703]	0 [0.501]	0.006 [0.7542]	0.7542 [0.7542]	0.8231	
Number of individuals per community	667	825.127 [72.089]	-122.914 [0.119]	-124.104 [0.139]	-98.487 [0.171]	25.617 [0.078]	40.000 [0.644]	741.955 [67.644]	13.533 [0.235]	-53.557 [0.235]	-85.044 [0.235]	-31.488 [0.066]	0.5469 [0.5469]	0.5395	
Distance to PoT	1.18	0.699 (0.079)	0.001 [0.988]	0.079 [0.411]	0.154 [0.116]	0.075 [0.432]	0.3271 [0.033]	1.81 [0.345]	0.345 [0.074]	0.074 [0.06]	-0.06 [0.134]	-0.134 [0.567]	0.0366 [0.23]	0.0000	
Panel B: Pretreatment, N = 3,678															
Floor made of tile/cement	0.21	0.259 [0.010]	-0.054 [0.161]	-0.009 [0.844]	-0.023 [0.552]	-0.011 [0.759]	0.5177 [0.061]	0.228 [0.123]	-0.031 [0.084]	-0.030 [0.404]	-0.051 [0.333]	-0.039 [0.412]	0.1790 [0.1790]	0.4125	
Completed primary schooling	0.488	0.484 [0.043]	-0.006 [0.893]	-0.013 [0.35]	-0.041 [0.277]	-0.005 [0.904]	0.4986 [0.037]	0.412 [0.818]	0.01 [0.013]	0.113 [0.114]	0.074 [0.044]	-0.039 [0.028]	0.0419 [0.0419]	0.1556	
Main ethnicity/Luhya	0.63	0.5 [0.067]	0.007 [0.896]	0.02 [0.714]	-0.061 [0.354]	-0.081 [0.201]	0.6361 [0.075]	0.501 [0.101]	-0.097 [0.163]	-0.082 [0.405]	-0.054 [0.635]	0.028 [0.3880]	0.3880 [0.3880]	0.2317	
Christian	0.972	0.973 [0.012]	-0.01 [0.691]	0.000 [0.813]	0.015 [0.266]	0.012 [0.391]	0.5230 [0.008]	0.981 [0.344]	-0.022 [0.773]	0.004 [0.165]	-0.033 [0.141]	-0.037 [0.3808]	0.0649	0.3808	
Panel C: Baseline knowledge, N = 2,056															
Know adults get worms	0.701	0.701 [0.042]	0.005 [0.917]	-0.048 [0.239]	0.058 [0.228]	0.106 [0.019]	0.1278 [0.05]	0.663 [0.186]	0.072 [0.569]	0.038 [0.529]	0.038 [0.000]	0 [0.6129]	0.6129 [0.3626]	0.3626	
Know children get worms	0.936	0.963 [0.014]	-0.039 [0.035]	-0.02 [0.332]	-0.024 [0.245]	-0.003 [0.881]	0.2079 [0.018]	0.926 [0.521]	0.014 [0.136]	0.03 [0.082]	-0.006 [0.146]	-0.035 [0.002]	-0.01 [0.3483]	0.3483 [0.3394]	0.3394
Believe deworming is for the sick	0.018	0.000 [0.008]	0.014 [0.101]	0.025 [0.123]	0.007 [0.392]	-0.018 [0.278]	0.2451 [0.007]	0.007 [0.011]	0 [0.967]	0.012 [0.426]	0.002 [0.88]	-0.01 [0.496]	0.8390 [0.6505]	0.6505	
Know medication treats worms	0.782	0.783 [0.015]	0.118 [0.002]	0.012 [0.063]	0.042 [0.063]	0.01 [0.014]	0.0025 [0.045]	0.501 [0.522]	0.054 [0.522]	-0.057 [0.101]	-0.057 [0.053]	-0.003 [0.052]	0.0819 [0.0819]	0.0024	
Dewormed in the past	0.686	0.75 [0.033]	-0.072 [0.072]	-0.047 [0.347]	-0.05 [0.149]	-0.008 [0.877]	0.3033 [0.036]	0.336 [0.283]	0.042 [0.138]	0.069 [0.16]	0.042 [0.83]	-0.061 [0.056]	0.4599 [0.3361]	0.3361	
Dewormed in the past year	0.376	0.446 [0.04]	-0.064 [0.228]	-0.022 [0.63]	-0.041 [0.338]	-0.018 [0.681]	0.6334 [0.031]	0.336 [0.031]	0.142 [0.022]	0.073 [0.022]	0.129 [0.022]	0.056 [0.022]	0.0015 [0.0124]	0.0124	
Know bi-yearly treatment recommended	0.49	0.533 [0.044]	0.053 [0.343]	-0.045 [0.415]	-0.033 [0.528]	0.012 [0.815]	0.2950 [0.042]	0.476 [0.948]	0.004 [0.023]	0.136 [0.51]	0.029 [0.047]	-0.107 [0.047]	0.1169 [0.1913]	0.1913	
Know worms impose externalities	0.368	0.356 [0.049]	0.058 [0.303]	0.01 [0.852]	-0.024 [0.651]	-0.035 [0.497]	0.4651 [0.051]	0.35 [0.522]	0.04 [0.825]	-0.014 [0.378]	0.056 [0.231]	0.07 [0.6011]	0.6011 [0.7249]	0.7249	
Panel D: Social image concerns, N = 2,056															
Would you judge: Not deworming?	0.698	0.744 [0.041]	-0.042 [0.345]	-0.011 [0.824]	0.011 [0.796]	0.022 [0.602]	0.5227 [0.061]	0.742 [0.083]	0.014 [0.682]	-0.03 [0.525]	0.041 [0.186]	0.071 [0.009]	0.6053 [0.6514]	0.6514	
Would you judge: Not immunizing a child?	0.778	0.805 [0.033]	-0.014 [0.739]	-0.049 [0.311]	-0.018 [0.633]	0.067 [0.183]	0.5077 [0.053]	0.926 [0.666]	0.014 [0.982]	0.023 [0.833]	-0.001 [0.844]	-0.011 [0.086]	0.8861 [0.8693]	0.8693	
Would you praise: Deworming?	0.946	0.945 [0.029]	0.012 [0.511]	-0.015 [0.408]	-0.013 [0.548]	-0.003 [0.941]	0.6038 [0.032]	0.937 [0.991]	0.002 [0.235]	-0.007 [0.809]	0.006 [0.205]	-0.046 [0.6586]	0.6586 [0.8512]	0.8512	
Would you praise: Immunizing a child?	0.959	0.94 [0.025]	0.009 [0.583]	-0.023 [0.241]	-0.011 [0.624]	0.012 [0.618]	0.3986 [0.027]	0.947 [0.202]	-0.022 [0.139]	-0.085 [0.601]	-0.011 [0.211]	0.074 [0.4583]	0.3516 [0.5483]	0.5483	
Panel E: Implementation, N = 3,678															
Did a CHV visit you?	0.829	0.863 [0.03]	-0.013 [0.757]	0.041 [0.886]	0.005 [0.244]	-0.036 [0.043]	0.3886 [0.043]	0.793 [0.51]	0.033 [0.415]	0.041 [0.064]	0.039 [0.212]	0.049 [0.1420]	0.2316 [0.1420]	0.1420	
Announcement about MDA in your community?	0.805	0.797 [0.097]	0.038 [0.711]	0.015 [0.9]	0.029 [0.793]	0.014 [0.096]	0.9851 [0.072]	0.964 [0.077]	-0.184 [0.387]	-0.078 [0.171]	-0.134 [0.59]	-0.057 [0.2840]	0.2840 [0.5860]	0.5860	
CHV share deworming practices?	0.981	0.994 [0.017]	0.024 [0.189]	0.005 [0.853]	0.022 [0.208]	0.018 [0.283]	0.3462 [0.01]	1 [0.86]	-0.001 [0.569]	-0.008 [0.162]	-0.008 [0.241]	-0.009 [0.4716]	0.3165 [0.5198]	0.5198	
CHV share where to get dewormed?	0.967	1 [0.024]	-0.002 [0.878]	-0.151 [0.081]	-0.001 [0.957]	0.15 [0.068]	0.3241 [0.024]	1 [0.236]	-0.088 [0.479]	-0.008 [0.172]	-0.017 [0.451]	-0.009 [0.4716]	0.4716 [0.6145]	0.6145	

Notes: Column “Con” reports the control group mean with its standard error in parentheses. Each subsequent column shows the difference between the treatment arm and the control mean, with the corresponding *p*-value in parentheses below. “Joint Test” reports the F-test *p*-value for equality of means across all four incentive arms within each distance condition (“Close” or “Far”) and across both distance conditions combined. Standard errors are clustered at the community level; county fixed effects are included to reflect stratification. Variable definitions:

- “Number of individuals per community” is calculated from the household census and aggregated to the community level.
- “Distance to PoT” is the straight-line distance (km) from the community centroid to its assigned point of treatment.
- “Know adults get worms” equals 1 if the respondent lists ‘adults’ among those who can be infected.
- “Know children get worms” equals 1 if the respondent lists ‘children’ among those who can be infected.
- “Believe deworming is for the sick” equals 1 if the respondent states that only sick people should deworm.
- “Know medication treats worms” equals 1 if the respondent states that medicine cures worm infections.
- “Know worms impose externalities” equals 1 if the respondent answers ‘Yes’ to ‘Can a person sick with worms spread worms to others?’ or to both of the following: ‘If you have worms, does that affect your neighbors’ or relatives’ health?’ and ‘If your neighbors or relatives have worms, does that affect your health?’.

Table B2: Praise and Stigma by Baseline Externality Knowledge

	Praise	Stigma
<i>Panel A: Above median externality knowledge</i>		
Above median (=1)	-0.026 (0.019)	-0.006 (0.028)
Constant	0.959 (0.008)	0.696 (0.018)
<i>Panel B: Community externality knowledge</i>		
Community knowledge (0–1)	-0.018 (0.031)	0.072 (0.063)
Constant	0.952 (0.013)	0.669 (0.027)
Observations	144	144

Notes: This table reports how likely respondents are to praise someone who participates in the deworming campaign or to look down on someone who abstains at baseline, by externality knowledge. Each column reports coefficients from community-level OLS regressions where the dependent variable is the share of baseline respondents in the community who (i) say they would praise someone who participates in the deworming campaign (left column) or (ii) say they would look down on someone who abstains (right column). Panel A uses a binary measure of externality knowledge: “Above median knowledge” equals 1 if the community’s baseline externality knowledge exceeds the sample median. Panel B uses the community’s baseline average externality knowledge (“Know worms impose externality,” see Table B1). Heteroskedasticity-robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B3: Distance and Travel to Deworming Sites

	Combined	Close	Far
Distance to treatment (km)	1.279	0.825	1.84
Fraction walked to treatment	0.898	0.938	0.826
<i>Travel time (minutes)</i>			
Mean	23	18	32
Median	20	15	30
Standard deviation	19	15	24
25th Percentile	10	10	20
75th Percentile	30	30	35
Observations	2,058	1,309	749

Notes: “Fraction walked to treatment” and “Travel time” are calculated from endline survey data for respondents who dewormed. “Distance to PoT” is computed from GIS data for the full analysis sample of 9,805 adults.

Table B4: Baseline Balance by Incentive Arm: Far – Close Differences

	Con	Con _{far - close}	Ink	Ink _{far - close}	Cal	Cal _{far - close}	Bra	Bra _{far - close}
Panel A: Takeup sample, N = 9,805								
Age	38.126 (0.758)	-0.132 [0.889]	37.412 (0.717)	-0.222 [0.811]	37.207 (0.749)	-0.074 [0.925]	38.358 (0.949)	-1.162 [0.252]
Female	0.543 (0.013)	-0.004 [0.798]	0.537 (0.011)	-0.012 [0.484]	0.541 (0.015)	0 [1]	0.507 (0.014)	0.009 [0.599]
Phone owner	0.816 (0.015)	-0.018 [0.242]	0.824 (0.017)	-0.027 [0.195]	0.81 (0.02)	0.009 [0.705]	0.811 (0.018)	-0.007 [0.693]
Number of individuals per community	825.127 (72.089)	-83.172 [0.311]	702.211 (61.564)	53.277 [0.539]	701.023 (70.246)	-12.625 [0.856]	726.64 [(50.457)]	-69.729 [0.243]
Distance to PoT	0.699 (0.079)	1.111 [0]	0.7 (0.081)	1.454 [0]	0.778 (0.09)	1.106 [0]	0.852 (0.083)	0.897 [0]
Panel B: Pretreatment, N = 3,678								
Floor made of tile/cement	0.259 (0.04)	0.029 [0.605]	0.205 (0.032)	0.001 [0.965]	0.249 [(0.049)]	0.037 [0.459]	0.236 [(0.033)]	0.002 [0.967]
Completed primary schooling	0.484 (0.043)	-0.072 [0.129]	0.478 (0.031)	-0.056 [0.152]	0.441 [(0.037)]	0.084 [0.052]	0.437 [(0.029)]	0.049 [0.247]
Main ethnicity/Luhya	0.5 (0.067)	0 [0.994]	0.508 (0.068)	-0.104 [0.058]	0.52 [(0.067)]	-0.102 [0.052]	0.439 [(0.081)]	0.008 [0.91]
Christian	0.973 (0.012)	0.008 [0.566]	0.963 [(0.022)]	-0.004 [0.905]	0.977 [(0.011)]	0.008 [0.603]	0.989 [(0.008)]	-0.04 [0.087]
Panel C: Baseline knowledge, N = 2,056								
Know adults get worms	0.701 (0.042)	-0.039 [0.473]	0.707 [(0.048)]	0.028 [0.56]	0.653 [(0.036)]	0.048 [0.401]	0.759 [(0.042)]	-0.059 [0.274]
Know children get worms	0.963 (0.014)	-0.037 [0.069]	0.924 [(0.015)]	0.017 [0.435]	0.943 [(0.018)]	0.013 [0.523]	0.939 [(0.019)]	-0.019 [0.479]
Believe deworming is for the sick	0.001 (0.008)	0.006 [0.614]	0.016 [(0.008)]	-0.008 [0.387]	0.027 [(0.015)]	-0.007 [0.701]	0.009 [(0.008)]	0 [0.966]
Know medication treats worms	0.783 (0.035)	0.057 [0.266]	0.901 [(0.027)]	-0.098 [0.039]	0.795 [(0.037)]	0.133 [0.003]	0.805 [(0.031)]	0.032 [0.464]
Dewormed in the past	0.75 (0.033)	-0.06 [0.147]	0.678 [(0.034)]	0.054 [0.148]	0.703 [(0.043)]	0.048 [0.342]	0.696 [(0.031)]	0.048 [0.158]
Dewormed in the past year	0.446 (0.04)	-0.109 [0.005]	0.381 [(0.046)]	0.097 [0.103]	0.423 [(0.039)]	-0.014 [0.778]	0.405 [(0.036)]	0.06 [0.177]
Know bi-yearly treatment recommended	0.533 (0.044)	-0.057 [0.285]	0.585 [(0.045)]	-0.105 [0.104]	0.488 [(0.045)]	0.124 [0.045]	0.5 [(0.038)]	0.005 [0.906]
Know worms impose externality	0.356 (0.049)	-0.006 [0.919]	0.415 [(0.046)]	-0.025 [0.658]	0.367 [(0.043)]	-0.03 [0.579]	0.332 [(0.041)]	0.074 [0.175]
Panel D: Social image concerns, N = 2,056								
Would you judge: Not deworming?	0.744 (0.041)	-0.003 [0.967]	0.702 [(0.035)]	0.053 [0.203]	0.733 [(0.042)]	-0.021 [0.703]	0.755 [(0.033)]	0.027 [0.486]
Would you judge: Not immunizing a child?	0.803 (0.033)	0.012 [0.815]	0.789 [(0.04)]	0.048 [0.298]	0.753 [(0.046)]	0.06 [0.294]	0.821 [(0.033)]	-0.017 [0.662]
Would you praise: Deworming?	0.935 (0.029)	-0.008 [0.716]	0.947 [(0.027)]	-0.02 [0.33]	0.92 [(0.033)]	-0.072 [0.284]	0.922 [(0.034)]	0.011 [0.632]
Would you praise: Immunizing a child?	0.94 (0.025)	0.007 [0.681]	0.949 [(0.023)]	-0.024 [0.137]	0.917 [(0.019)]	-0.055 [0.651]	0.929 [(0.011)]	0.007 [0.182]
Panel E: Implementation, N = 3,678								
Did a CHV visit you?	0.863 (0.03)	-0.07 [0.134]	0.849 [(0.036)]	-0.023 [0.629]	0.903 [(0.026)]	-0.069 [0.047]	0.867 [(0.026)]	0.016 [0.655]
Announcement about MDA in your community?	0.797 (0.097)	0.166 [0.095]	0.836 [(0.073)]	-0.056 [0.607]	0.813 [(0.105)]	0.073 [0.517]	0.826 [(0.08)]	0.003 [0.979]
CHV share deworming practices?	0.994 (0.017)	0.021 [0.213]	1.018 [(0.013)]	-0.005 [0.476]	0.999 [(0.019)]	0.009 [0.651]	1.016 [(0.011)]	-0.086 [0.182]
CHV share where to get dewormed?	1 (0.024)	0.013 [0.343]	1.026 [(0.018)]	-0.073 [0.305]	0.877 [(0.076)]	0.156 [0.063]	1.027 [(0.018)]	-0.003 [0.81]

Notes: This table reports the same variables as Table B1 but reorganized to show within-arm differences by distance. For each incentive arm (Control, Ink, Calendar, Bracelet), the left column reports the arm specific baseline mean with its standard error in parentheses; the adjacent column reports the Far – Close difference within that arm, with the corresponding *p*-value listed below. Standard errors are clustered at the community level; county fixed effects are included to reflect stratification.

Table B5: Endline Incentive Check

	Received incentive	Have incentive currently	Seen incentive	Link incentive to deworming
Bracelet	0.022 (0.015)	0.666*** (0.035)	0.277*** (0.032)	-0.017 (0.012)
	-0.001 (0.018)	0.81*** (0.027)	0.073 (0.044)	-0.065** (0.021)
Ink mean	0.95 (0.013)	0.144 (0.025)	0.674 (0.03)	0.966 (0.008)
Observations	1,613	1,254	2,815	2,239

Notes: Columns report respondent level indicators collected at endline. “Received incentive” is the share of dewormed respondents who say they received the assigned bracelet, calendar, or ink mark at the PoT. “Have incentive currently” is the share of those dewormed respondents who still possess the item at endline. “Seen incentive” is the share of respondents who report seeing the item in their community (“Have you seen ink on people’s fingers?” or, in the Calendar arm, “Have you seen this calendar before?”). “Link incentive to deworming” is the share of respondents who, when asked what the item means, mention any of the words “deworming/medication/treatment/tablet/drug/worms”. Rows for Bracelet and Calendar show estimated treatment effects relative to the mean of the Ink arm (the reference group); standard errors are clustered at the community level and appear in parentheses. For outcomes defined for all respondents, we use the full endline sample ($N = 3,678$) minus those assigned to the Control arm—who were ineligible to receive an incentive—yielding $N = 2,815$; this is the sample used for *Seen incentive*. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B6: Effects of Signals and Distance on the Observability of Deworming - Continuous Distance

Dependent variable: Observability	Reduced Form			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.135*** (0.037)	0.052 (0.049)	0.235*** (0.052)	0.183** (0.067)
	0.04 (0.039)	0.036 (0.054)	0.046 (0.054)	0.01 (0.075)
Calendar	0.088** (0.04)	0.037 (0.054)	0.15** (0.052)	0.113 (0.07)
	0.689 (0.031)	0.745 (0.043)	0.623 (0.044)	-0.122** (0.061)
Control mean	999	999	999	999
H0: Any Signal = No Signal, p -value	<0.001	0.432	<0.001	0.001
H0: Bracelet = Calendar, p -value	0.001	0.671	<0.001	0.001

Notes: This table replicates the specification in Table 1 replacing the binary Close/Far indicator with the continuous distance from each community’s centroid to its assigned treatment location. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B7: Effects of Signals and Distance on the Observability of Deworming - Full Sample

Dependent variable: Observability	Reduced Form			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.129*** (0.031)	0.06 (0.041)	0.217*** (0.047)	0.157** (0.062)
Calendar	0.034 (0.036)	0.03 (0.05)	0.04 (0.054)	0.011 (0.074)
Ink	0.087** (0.035)	0.036 (0.048)	0.153** (0.05)	0.117* (0.069)
Control mean	0.705 (0.026)	0.763 (0.034)	0.63 (0.04)	-0.133** (0.053)
Observations	1,564	1,564	1,564	1,564
H_0 : Any Signal = No Signal, p -value	<0.001	0.288	<0.001	0.004
H_0 : Bracelet = Calendar, p -value	0.001	0.446	<0.001	0.015

Notes: This table replicates the specification in Table 1 using the full sample of endline respondents who completed the beliefs module, including SMS recipients and individuals whose deworming take-up was not observed at treatment sites. The larger sample yields slightly smaller standard errors, making the Far–Close difference for Ink significant at the 10 percent level and the Far–Close difference for Control significant at the 5 percent level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B8: Effects of Signals and Distance on the Perceived Observability of Deworming

Dependent variable: Perceived Observability	Reduced Form			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.06 (0.054)	0.029 (0.069)	0.099 (0.08)	0.07 (0.102)
Calendar	0.013 (0.051)	0.047 (0.069)	-0.027 (0.074)	-0.074 (0.1)
Ink	0.031 (0.054)	-0.01 (0.074)	0.08 (0.074)	0.089 (0.101)
Control mean	0.646 (0.042)	0.687 (0.056)	0.596 (0.059)	-0.091 (0.079)
Observations	999	999	999	999
H_0 : Any Signal = No Signal, p -value	0.28	0.766	0.053	0.093
H_0 : Bracelet = Calendar, p -value	0.256	0.746	0.065	0.119

Notes: This table replicates the specification in Table 1 using the respondent-level measure of *perceived* observability, defined as the fraction of recognized community members who the respondent reports know the respondent's deworming status. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B9: Predicted Deworming Take-up at Endline

Dependent variable: Predicted Take-up	Reduced Form			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.097*** (0.022)	0.06** (0.026)	0.142*** (0.037)	0.082* (0.045)
Calendar	0.041* (0.022)	0.015 (0.03)	0.072** (0.033)	0.057 (0.044)
Ink	0.017 (0.023)	-0.002 (0.027)	0.041 (0.04)	0.043 (0.048)
Control mean	0.573 (0.018)	0.628 (0.021)	0.507 (0.032)	-0.121** (0.039)
Observations	2,312	2,312	2,312	2,312

Notes: This table replicates the specification in Table 2 using respondents' *predictions*—rather than observed data—of deworming take-up. At endline each participant was told: "I will now ask you about deworming take-up in your community and the decision that people made to come for deworming or not take up treatment. If you can correctly guess what people in your community did, you will be able to win airtime credit. I will transfer you the credit tomorrow. Out of 10 people in your community, how many do you think came for deworming?" The dependent variable is the stated number divided by 10.
 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B10: Effect of Signals and Distance on Deworming Take-up - Continuous Distance

Dependent variable: Take-up	Reduced Form			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.085** (0.028)	0.06* (0.032)	0.115** (0.044)	0.056 (0.052)
Calendar	0.038 (0.027)	0.031 (0.03)	0.046 (0.037)	0.014 (0.04)
Ink	-0.013 (0.03)	-0.047 (0.031)	0.029 (0.042)	0.076* (0.041)
Control mean	0.328 (0.022)	0.405 (0.023)	0.232 (0.033)	-0.173*** (0.033)
Observations	9,805	9,805	9,805	9,805
H_0 : Any Signal = No Signal, p -value	0.36	0.651	0.069	0.049
H_0 : Bracelet = Calendar, p -value	0.046	0.344	0.037	0.336

Notes: This table replicates the specification in Table 2 replacing the binary Close/Far indicator with the continuous distance from each community's centroid to its assigned treatment location. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B11: Heterogeneous Treatment Effects by Covariates

Model:	Age > 40 (1)	Female (2)	Phone Owner (3)	Judgemental (4)	Prev Dewormed (5)	Externality Knowledge (6)
Bracelet	0.096*** (0.031)	0.061* (0.034)	0.096*** (0.033)	0.065* (0.036)	0.096** (0.043)	0.106** (0.043)
Calendar	0.042 (0.029)	0.008 (0.029)	0.035 (0.032)	0.019 (0.034)	0.064* (0.038)	0.063* (0.035)
Ink	0.017 (0.030)	-0.021 (0.032)	0.016 (0.037)	-0.032 (0.037)	-0.021 (0.043)	-0.019 (0.038)
Covariate	0.187*** (0.022)	0.067*** (0.019)	0.056* (0.030)	-0.017 (0.049)	-0.012 (0.047)	0.060 (0.047)
Covariate \times treatment = Bracelet	-0.034 (0.030)	0.041 (0.029)	-0.020 (0.035)	0.059 (0.065)	-0.033 (0.059)	-0.061 (0.061)
Covariate \times treatment = Calendar	-0.022 (0.031)	0.039 (0.027)	-0.007 (0.037)	0.042 (0.058)	-0.060 (0.057)	-0.081 (0.060)
Covariate \times treatment = Ink	-0.080*** (0.029)	0.006 (0.025)	-0.046 (0.038)	0.042 (0.068)	0.001 (0.062)	-0.004 (0.061)
Control mean	0.357	0.357	0.357	0.357	0.357	0.357
Observations	9,805	9,805	9,805	9,805	9,805	9,805

Notes: Each column reports heterogeneous treatment effects for a single covariate group. We estimate $Y_{ic} = X_{ic}\beta + \sum_{z=1}^4 (\gamma_z \text{treatment}_{icz} \times X_{ic} + \beta_z \text{treatment}_{icz}) + \alpha_s$, where Y_{ic} is a deworming indicator, X_{ic} is the covariate listed in the column header, and α_s is a county-strata fixed effect. Age > 40 – individual is older than 40. Female – individual is female. Phone Owner – individual owns a phone. Community Judgment – community’s praise/stigma score above the sample mean. Previously Dewormed – community’s prior deworming rate above the sample mean. Externality Knowledge – community’s knowledge of externalities score above the mean. Standard errors are clustered at the community level.
 $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table B12: Effects of Reminder and Social Info SMS on Take-up in Control Group

Dependent variable: Take-up	
Social Info	0.103*** (0.028)
Reminder Only	0.136*** (0.031)
Control, no SMS mean	0.356 (0.479)
Observations	2,564

Notes: This table reports estimates from a linear probability model (OLS) with a binary outcome for deworming take-up. The regressors are indicators for receiving a “Social Info” SMS or a “Reminder Only” SMS; the omitted category is phone owners who received no SMS. Because the Reminder treatment was implemented only in Control communities, the analysis sample is restricted to phone owners in the Control arm. All regressions include strata fixed effects and the LASSO-selected controls—sex, age, and expected distance to the PoT. Bootstrapped standard errors, clustered at the community level, are in parentheses.
 $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Table B13: Externality Knowledge at Endline

Dependent variable: Externality Knowledge	Reduced Form			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.031 (0.039)	0.004 (0.058)	0.063 (0.051)	0.059 (0.077)
Calendar	0.022 (0.039)	0.03 (0.061)	0.012 (0.045)	-0.017 (0.076)
Ink	0.081* (0.043)	0.107* (0.058)	0.051 (0.062)	-0.056 (0.084)
Control mean	0.313 (0.03)	0.318 (0.05)	0.308 (0.033)	-0.01 (0.06)
Observations	2,312	2,312	2,312	2,312
H_0 : Any Signal = No Signal, p -value	0.1	0.283	0.196	0.85
H_0 : Bracelet = Calendar, p -value	0.785	0.586	0.286	0.262

Notes: Each column reports OLS estimates of the respondent-level externality knowledge measure, equal to 1 if the respondent answers ‘Yes’ to ‘Can a person sick with worms spread worms to others?’ or to both of the following: ‘If you have worms, does that affect your neighbors’ or relatives’ health?’ and ‘If your neighbors or relatives have worms, does that affect your health?’ ‘Control’ is the mean in the Control arm. Rows for Bracelet, Ink, and Calendar give treatment effects relative to Control. Column (4) reports the difference in those effects between Far and Close communities. We pool Bracelet and Ink as “Any Signal” and Control and Calendar as “No Signal” to test H_0 : Any Signal = No Signal. The second test, H_0 : Bracelet=Calendar, compares the bracelet and active control incentive. All regressions include strata fixed effects and the LASSO-selected controls—sex, age, and expected distance to the PoT. Bootstrapped standard errors, clustered at the community level, are reported in parentheses.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B14: Difference in Monetary Valuation of Calendar vs. Bracelet

Parameter	Posterior estimates
<i>Panel A: Model parameters</i>	
Valuation difference (KSh), mean	0.472 (0.404, 0.54)
<i>Panel B: Estimated preferences</i>	
Pr(Prefer calendar), offered 50KSh	0.897 (0.876, 0.917)
Pr(Prefer calendar), offered 0KSh	0.731 (0.705, 0.755)
Pr(Prefer calendar), offered -50KSh	0.486 (0.448, 0.519)

Notes: The table reports posterior means and 95 percent credible intervals from the willingness to pay (WTP) submodel, one of the three blocks of the structural estimation (see Section 6). Private utility parameters are restricted to be equal, $\beta_{\text{calendar}} = \beta_{\text{bracelet}}$, so the scalar ψ measures the mean difference (KSh) in respondents’ valuation of a calendar relative to a bracelet (Panel A). Panel B gives posterior predicted probabilities that a respondent prefers the calendar when the switch offer is +50, 0, and -50KSh. The two-stage probit likelihood is detailed in Appendix G.

Table B15: Gift-Choice Preference for a Bracelet vs. Calendar by Distance

	Reduced Form			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Dependent variable: Prefer Bracelet				
Bracelet	-0.066** (0.028)	-0.136*** (0.037)	-0.002 (0.042)	0.134** (0.056)
Calendar	0.159*** (0.035)	0.199*** (0.054)	0.122** (0.044)	-0.077 (0.069)
Ink	0.036 (0.031)	0.022 (0.038)	0.049 (0.048)	0.027 (0.061)
Control mean	0.235 (0.019)	0.249 (0.026)	0.222 (0.028)	-0.026 (0.038)
Observations	1,808	1,808	1,808	1,808

Notes: The table tests whether bracelets carry extra private or prevalence-related value with distance (Far vs. Close). The dependent variable equals 1 if a respondent chooses the bracelet (rather than the calendar) when offered one of the two items for free at endline. The sample is restricted to adults who did not deworm and therefore did not already possess either item; gift choice thus reveals private consumption value, including any premium associated with local item prevalence. Estimates follow the main reduced-form specification: OLS with strata fixed effects, LASSO-selected controls (sex, age, and expected distance to the point of treatment, PoT), and standard errors bootstrapped and clustered at the community level. Rows for Bracelet, Ink, and Calendar report effects relative to Control in each column; Column (4) reports the Far–Close difference in those effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B16: Structural Parameters: Posterior Estimates and CrIs

Parameter	Estimate
Take-up	
λ	0.645 (0.328, 0.936)
$\beta_{Control}$	-0.833 (-1.173, -0.449)
β_{Ink}	-0.145 (-0.242, -0.051)
$\beta_{Calendar}$	0.102 (0.033, 0.178)
$\beta_{Bracelet}$	0.102 (0.033, 0.178)
δ	0.21 (0.176, 0.247)
σ_u	0.317 (0.139, 0.609)
Observability	
$\beta_{Control}^O$	1.668 (1.428, 1.941)
β_{Ink}^O	-0.257 (-0.598, 0.048)
$\beta_{Calendar}^O$	-0.147 (-0.437, 0.146)
$\beta_{Bracelet}^O$	-0.172 (-0.529, 0.191)
$\gamma_{Control}^O$	-0.381 (-0.52, -0.268)
γ_{Ink}^O	0.426 (0.266, 0.598)
$\gamma_{Calendar}^O$	0.148 (-0.008, 0.3)
$\gamma_{Bracelet}^O$	0.531 (0.357, 0.706)
WTP	
ψ	0.472 (0.403, 0.547)
ρ	8e-05 (0, 0.00021)

Notes: This table reports posterior means and 95% credible intervals from the joint Bayesian estimation in Section 6. Take-up parameters include social image weight λ , arm specific private value shifts β_z ($z \in \text{Control, Ink, Calendar, Bracelet}$), the marginal disutility of distance δ (standardized km units), and the idiosyncratic shock s.d. σ_u . Observability parameters include arm specific intercept and distance slope shifts. A negative $\gamma_{Control}^O$ means observability falls with distance in Control; positive $\gamma_{Bracelet}^O$ or γ_{Ink}^O offset this decline. The WTP parameter ψ is the mean KSh valuation difference from the two-stage choice exercise.

Table B17: Structural Model: Signal-by-Distance Effects on Observability

Dependent variable: Observability	Structural			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.093 (0.054, 0.132)	0.035 (-0.005, 0.075)	0.15 (0.094, 0.217)	0.115 (0.058, 0.186)
Calendar	0.015 (-0.019, 0.05)	-0.003 (-0.041, 0.037)	0.033 (-0.007, 0.077)	0.036 (-0.001, 0.08)
Ink	0.061 (0.019, 0.101)	0.011 (-0.029, 0.051)	0.11 (0.054, 0.168)	0.099 (0.048, 0.159)
Control mean	0.752 (0.72, 0.784)	0.797 (0.765, 0.828)	0.707 (0.656, 0.754)	-0.09 (-0.145, -0.045)
Δ (Any Signal–No Signal)	0.062 (0.032, 0.092)	0.026 (-0.006, 0.06)	0.098 (0.058, 0.142)	0.072 (0.032, 0.124)
Δ (Bracelet–Calendar)	0.078 (0.045, 0.11)	0.038 (-0.001, 0.077)	0.118 (0.075, 0.168)	0.08 (0.035, 0.138)

Notes: This table reports the structural counterparts to the reduced form estimates in Table B6. Point estimates are posterior means; parentheses contain 95 percent credible intervals. The structural model is fit in Stan via Hamiltonian Monte-Carlo (four chains, 400 warm-up draws, and 400 saved iterations per chain). Sampler diagnostics show no divergent transitions and split $\hat{R} < 1.1$ for all parameters. The final two rows report differences in expected observability: Any Signal minus No Signal and Bracelet minus Calendar.

Table B18: Structural Model: Signal-by-Distance Effects on Deworming Take-up

Dependent variable: Take-up	Structural			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.073 (0.049, 0.098)	0.056 (0.029, 0.083)	0.095 (0.065, 0.124)	0.039 (0.013, 0.067)
Calendar	0.04 (0.017, 0.064)	0.037 (0.01, 0.065)	0.044 (0.021, 0.067)	0.008 (-0.01, 0.03)
Ink	-0.019 (-0.045, 0.005)	-0.04 (-0.071, -0.011)	0.008 (-0.019, 0.033)	0.047 (0.024, 0.075)
Control mean	0.332 (0.315, 0.35)	0.399 (0.377, 0.42)	0.249 (0.228, 0.27)	-0.15 (-0.173, -0.129)
Δ (Signal–No Signal)	-0.013 (-0.029, 0.005)	-0.029 (-0.044, -0.011)	0.007 (-0.014, 0.029)	0.035 (0.02, 0.055)
Δ (Bracelet–Calendar)	0.033 (0.017, 0.051)	0.019 (0.006, 0.036)	0.05 (0.027, 0.075)	0.031 (0.014, 0.054)

Notes: This table reports the structural counterparts to the reduced form estimates in Table B10. Point estimates are posterior means; parentheses contain 95 percent credible intervals. The structural model is estimated in Stan via Hamiltonian Monte-Carlo (four chains, 400 warm-up draws, and 400 saved iterations per chain). Sampler diagnostics show no divergent transitions and split $\hat{R} < 1.1$ for all parameters. The final rows report differences in expected take-up: Any Signal minus No Signal and Bracelet minus Calendar.

Table B19: Structural Model: Effects of Social Image and Private Utility on Take-up

Dependent variable: Take-up	Structural			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
<i>Panel A: Social Image</i>				
Bracelet	0.029 (0.012, 0.049)	0.012 (-0.002, 0.029)	0.045 (0.021, 0.076)	0.033 (0.012, 0.062)
Calendar	0.004 (-0.008, 0.017)	-0.001 (-0.017, 0.013)	0.009 (-0.003, 0.025)	0.011 (-0.001, 0.027)
Ink	0.018 (0.004, 0.036)	0.004 (-0.011, 0.018)	0.033 (0.012, 0.058)	0.029 (0.011, 0.054)
<i>Panel B: Private</i>				
Bracelet	0.032 (0.007, 0.058)			
Calendar	0.032 (0.007, 0.058)			
Ink	-0.044 (-0.073, -0.015)			

Notes: Panel A shows the treatment effect on deworming take-up from social image utility (private valuation fixed at its Control value), whereas Panel B shows the treatment effect attributable to private utility (social image utility fixed at the Control value). Point estimates are posterior means; parentheses contain 95 percent credible intervals. The structural model is estimated in Stan via Hamiltonian Monte-Carlo (four chains, 400 warm-up draws, and 400 saved iterations per chain). Sampler diagnostics show no divergent transitions and split $\hat{R} < 1.1$ for all parameters.

Table B20: Density of Communities

Mean (persons/km ²)	St. dev. (persons/km ²)	Prop. within 1 s.d.	Prop. within 2 s.d.	<i>N</i> communities
658.22	309.68	0.78	0.95	144

Notes: The table reports summary statistics for population density in each of the 144 study communities. A community's area is computed as the convex hull of all dwellings mapped in the household census; density equals the census population divided by this area (km²).

Table B21: Number of Points of Treatment Supplied

Observability	Assigned PoTs	Mean take-up	Mean distance (km)
<i>Panel A: Experimental allocation</i>			
Control	144	0.33 (0.310, 0.346)	1.2
<i>Panel B: Policymaker allocation</i>			
Control	107 (100, 114)	0.33 (0.329, 0.331)	1.26 (1.11, 1.411)
Bracelet	90 (79, 102)	0.33 (0.329, 0.33)	1.68 (1.364, 1.957)
Control social image returns at 0.5km	100 (92, 110)	0.33 (0.329, 0.331)	1.41 (1.194, 1.614)
Bracelet social image returns at 0.5km	96 (86, 105)	0.33 (0.329, 0.331)	1.52 (1.297, 1.758)
No social image returns	144 (144, 144)	0.174 (0.101, 0.259)	0.58 (0.58, 0.58)

Notes: Point estimates are posterior medians; parentheses contain 95 percent credible intervals. In Panel A (experimental allocation) the numbers of PoTs and the mean community–PoT distance are fixed by design, so credible intervals are omitted. Panel B reports the policymaker's solution: for each of 200 posterior draws from the structural model we solve an integer program that minimizes the number of PoTs subject to the coverage target and a maximum walking distance of 3.5 km; the table shows the posterior medians and credible intervals of the resulting allocations. No community–PoT pair is allowed if travel would exceed 3.5 km.

C Cluster Selection and Distance Randomization

We randomly selected 158 clusters from a sampling frame of 1,451 primary schools across three study counties. Each cluster consists of a single point of treatment (PoT), located at a primary school, and its associated catchment area. Of the 158 drawn clusters, 144 were ultimately implemented.³⁵

Point of Treatment Selection and Distance Assignment

Primary school coordinates serve both to identify feasible PoTs and to locate nearby communities for the information campaign and data collection.³⁶ Selection proceeded via a geographically constrained randomization:

1. **Cluster selection.** We applied an acceptance–rejection algorithm to draw 158 schools. Each selected school was preliminarily assigned to one of two randomization groups—“Control/Calendar” or “Bracelet/Ink”—and served as the center of a cluster with a 2.5 km catchment radius. We imposed a 3 km buffer for Control/Calendar clusters and a 4 km buffer for Bracelet/Ink clusters to limit potential spillovers. The algorithm iteratively selected schools whose buffer zones overlapped existing clusters by no more than a prespecified threshold. If the required number of clusters was not reached, the procedure re-started. Figure A3 maps the final set of clusters.³⁷
2. **Distance treatment.** Clusters were then randomly assigned, stratified by county, to “Close” or “Far”: 0–1.25 km versus 1.25–2.5 km from the cluster’s centroid to the PoT.
3. **Incentive treatment.** Conditional on distance and county, we randomly assigned clusters to one of four incentive arms—Control, Ink, Calendar, or Bracelet.

Target Community Selection

4. **Anchor schools.** For each cluster we picked an anchor school consistent with its distance assignment. Close clusters kept the initially selected school (or a nearby school); Far clusters used an alternative school to ensure the required separation.³⁸
5. **Community listing.** Enumerators listed all communities near the anchor school and we randomly selected one as the target community.

³⁵We initially targeted 150 clusters and drew eight additional clusters as fallbacks. Field constraints limited implementation to 144 clusters.

³⁶Coordinates are from the Kenya Open Data Portal: <http://www.opendata.go.ke/>.

³⁷For each PoT we verified that treatment was feasible and identified a nearby back-up PoT.

³⁸We re-ran the randomization algorithm using [Borusyak and Hull \(2023\)](#) after data collection was complete, so we are unable to survey additional nearby communities. Therefore, all results calculate potential distance between PoTs and previously surveyed communities. We have good geographic coverage of communities and every counterfactually placed PoT has an eligible community to re-assign.

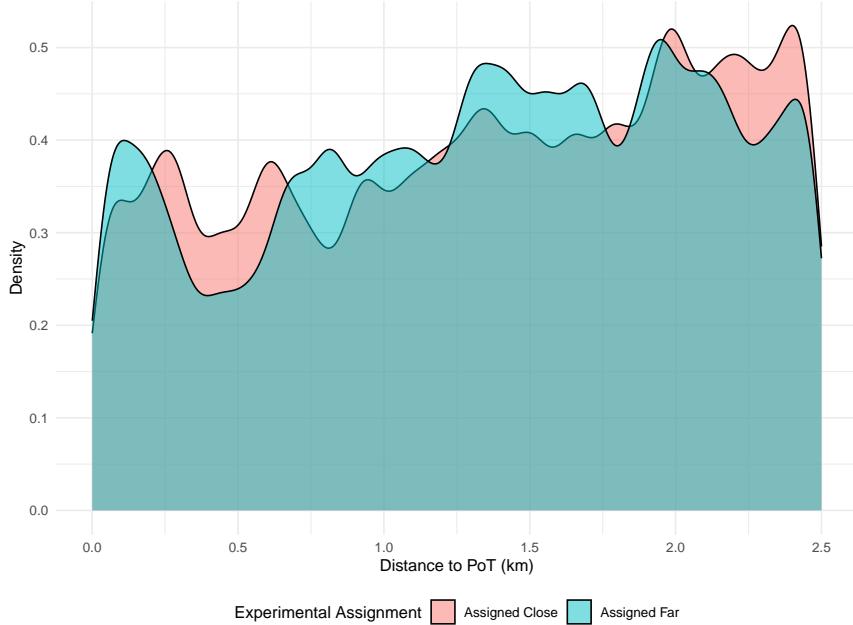


Figure C1: Counterfactual Assignment Distance Density

Notes: The figure plots the distribution of expected distances from each community to a PoT under 100 re-randomizations of the site selection algorithm ([Borusyak and Hull 2023](#)). Red (blue) curves correspond to clusters observed as Close (Far) in our experiment. A Kolmogorov–Smirnov test fails to reject equality of the two distributions ($p = 0.472$), indicating that observed Close (respectively, Far) communities are not systematically predisposed to shorter (longer) potential distances.

D Continuous Distance Randomization Inference

As an additional check that our continuous distance measure is orthogonal to pre-treatment characteristics, we regress each baseline covariate on the distance from the community centroid to the nearest assigned point of treatment (PoT), controlling for county-stratum fixed effects and clustering standard errors at the community level. We then draw 500 within-stratum permutations of the distance assignment and recompute the associated t -statistics. Figure D1 plots the randomization distribution of each statistic under the null of no association; the vertical line marks the realized value, and the exact randomization p -values are reported in the upper right-hand corner of each panel. For all 17 covariates the p -value exceeds 0.10, indicating that continuous distance is not systematically related to observable baseline characteristics.

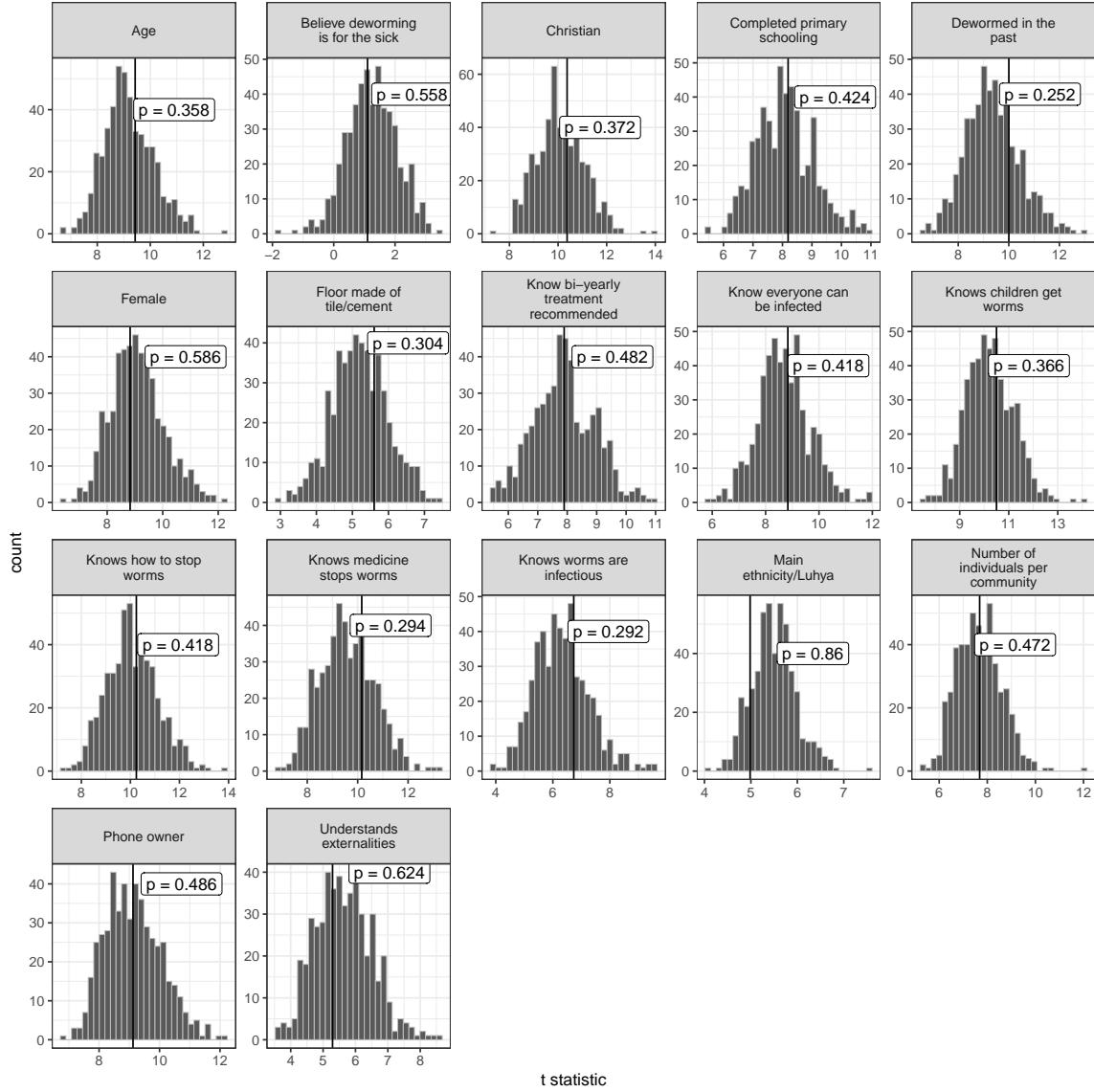


Figure D1: Randomization Distribution – Baseline Covariates and Distance to PoT

Notes: Each histogram shows the distribution of the OLS t -statistic obtained from 500 within-community permutations of community-level distance assignments. Models include county-stratum fixed effects; standard errors are clustered by community. The vertical line is the realized t -statistic from the actual experiment, and the panel reports the corresponding exact randomization p -value.

E Mapping of Preregistered Analyses to Implemented Analyses

This appendix maps the comparisons prespecified in the pre-analysis plan (AEARC-TR-0001643) to the corresponding results in the paper and flags analyses added after registration. All preregistered reduced form comparisons of the incentive and SMS arms—using the preregistered outcomes and the Close/Far distance stratification—are reported, and no additional reduced form comparisons were introduced. After data collection, we added two extensions: (i) a structural interpretation of the social multiplier mechanism and (ii) an allocation exercise that uses the estimated parameters to place treatment sites. Both are clearly labeled as post-PAP extensions and nest the preregistered reduced form tests. References in the table to “Where in PAP” and “Where in paper” are to section or subsection numbers.

Table E1: Mapping of Pre-Analysis Plan (PAP) Items to Reported Analyses

PAP item	Where in PAP	Where in paper	Status
Design: four incentive arms (Control, Calendar, Bracelet, Ink); SMS cross-cut (None, Only, Social Info); distance treatment (Close/Far).	Secs. 4.3–4.6 (design and assignment and assignment)	Sec. 3 (Design and implementation)	Implemented as preregistered.
Primary outcome: adult deworming take-up.	Sec. 6.1 (“deworming choice”)	Sec. 4 (reduced form); main take-up table Table 2	Implemented as preregistered; measured via site monitoring (see Sec. 3).
Observability of peers’ deworming decisions (constructed from beliefs module).	Sec. 5.3.2 (“social knowledge”)	Sec. 4 (reduced form); main observability table Table 1	Constructed from pre-registered questions. Mechanism check, secondary outcome.

Table E1: Mapping of Pre-Analysis Plan (PAP) Items to Reported Analyses (continued)

PAP item	Where in PAP	Where in paper	Status
Heterogeneity by distance (Close vs. Far; continuous distance checks in robustness).	Secs. 4.5–4.6; Sec. 6.4 (heterogeneity)	Sec. 4 (Close/Far comparisons throughout)	Implemented as pre-registered (Close/Far). Continuous distance results reported as robustness.
SMS comparisons (Reminder/Social Info): (i) incentive effects among SMS recipients vs. non-recipients; (ii) Reminder vs. No SMS in Control; (iii) Social Info vs. No SMS within social incentive clusters.	Secs. 6.2.3–6.2.5 (5)–(10)	Sec. 4.2.1; Fig. A9; App. Table B11	Implemented as preregistered; presented as robustness. No new reduced form contrasts beyond PAP.
Active control comparison: Bracelet vs. Calendar.	Sec. 6.2.2	Sec. 4 (reduced form); comparison reported in main tables	Implemented as preregistered.
Calendar/Bracelet preference at endline (gift choice).	Sec. 6.2.6	Sec. 4.2.3; App. Tables B13 and B14	Implemented as preregistered.
<i>Additions beyond the PAP</i>			
Willingness to pay (BDM) for — Calendar vs. Bracelet (Control sample).	—	Sec. 4.2.3; App. Table B13	Benchmarks private valuation; complements preregistered gift-choice module.

Table E1: Mapping of Pre-Analysis Plan (PAP) Items to Reported Analyses (continued)

PAP item	Where in PAP	Where in paper	Status
Theory and structural estimation of the social multiplier; decomposition of distance effects.	—	Secs. 5–6 (model and estimation)	Developed post-analysis; nests preregistered tests and exploits randomized observability and distance.
Optimal allocation / site placement exercise.	—	Sec. 7 (allocation)	Uses structural estimates to solve an implied policy problem.

F Ruling Out Pure Private Utility Substitution

The reduced form evidence produces two key facts: (i) signal incentives (Bracelet, Ink) increase take-up more in Far than in Close communities; (ii) the willingness to pay (WTP) exercise shows no distance-related difference in the private valuation of a bracelet relative to a calendar,

$$\mathbb{E}[v_B - v_C | d] = 0 \quad \text{for all } d, \quad (\text{F.1})$$

where v_B, v_C denote respondents' monetary valuations of the two items. This appendix shows that a binary choice model driven only by private utility and distance cost cannot fit both facts simultaneously.

Latent utility model. Let $B_i, C_i \in \{0, 1\}$ indicate assignment to the Bracelet or Calendar arm (Control: $B_i = C_i = 0$), and let d_i be the community-to-PoT distance. Define y_i^* as the latent net utility of deworming for individual i ; we assume

$$y_i^* = \beta_0 + \beta_B B_i + \beta_C C_i + \delta d_i + \gamma_B (B_i d_i) + \gamma_C (C_i d_i) + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, 1), \quad (6)$$

and set the observed decision $y_i = \mathbf{1}\{y_i^* > 0\}$.

Relative marginal utilities. The ratio of marginal private utilities of a bracelet versus a calendar is

$$R(d) = \frac{\partial y_i^*/\partial B_i}{\partial y_i^*/\partial C_i} = \frac{\beta_B + \gamma_B d}{\beta_C + \gamma_C d}. \quad (7)$$

Suppose (F.1) holds. Then model (6) can replicate the larger Bracelet effect at long distances only if the interaction coefficients satisfy

$$\gamma_B \beta_C - \gamma_C \beta_B \neq 0, \quad (\text{F.2})$$

which contradicts (F.1). Therefore a model with private utility and distance cost alone cannot explain both empirical facts.

Proof. Differentiating (7) with respect to d yields

$$R'(d) = \frac{\gamma_B \beta_C - \gamma_C \beta_B}{(\beta_C + \gamma_C d)^2}.$$

Hence $R(d)$ is constant in d iff $\gamma_B \beta_C - \gamma_C \beta_B = 0$. This equality obtains only in two cases: (i) $\gamma_B = \gamma_C = 0$, implying no distance interaction; (ii) exact proportionality, $\gamma_B/\beta_B = \gamma_C/\beta_C \neq 0$.

Empirically, the signal-distance interaction is positive and significant ($\gamma_B > 0$), whereas the Calendar interaction is indistinguishable from zero ($\gamma_C \approx 0$). Neither case (i) nor case

(ii) holds, so $R(d)$ must vary with d , contradicting the distance-invariant valuation gap (F.1). \square

Empirical check. Table B15 confirms the premise (F.1): the probability that a respondent chooses the bracelet over the calendar when offered a free gift is the same in Close and Far communities ($p = 0.40$ and $p = 0.49$, respectively). Consequently, heterogeneous substitution in private utility cannot by itself account for the stronger signaling incentive effects observed at greater distances; an additional distance modulated component—such as social image utility—is required.

G WTP Likelihood

Let each individual i have a latent valuation

$$g_i^{\text{wtp}} \sim N(\psi, 1),$$

where ψ (to be estimated) is the mean difference in monetary value between calendar and bracelet and the variance is fixed at one for scale identification.

Stage 1 (item choice). Respondent i first chooses between a calendar and a bracelet.

Define

$$c_i = \begin{cases} 1, & \text{calendar chosen,} \\ 0, & \text{bracelet chosen,} \end{cases} \quad c_i = 1 \iff g_i^{\text{wtp}} > 0.$$

Hence $\psi > 0$ means that on average individuals tend to prefer the calendar.

Stage 2 (cash offer and switch decision). After the initial choice, a random cash offer m_i (in KSh) is drawn. The respondent may switch to the other item. Let

$$s_i = \begin{cases} 1, & \text{switch occurs,} \\ 0, & \text{otherwise.} \end{cases}$$

- If $c_i = 0$, switching to the calendar occurs when $-m_i < g_i^{\text{wtp}} < 0$.
- If $c_i = 1$, switching to the bracelet occurs when $0 < g_i^{\text{wtp}} < m_i$.

Likelihood. With $\Phi(\cdot)$ denoting the standard normal c.d.f., the individual contribution to the likelihood is

$$\begin{aligned} L_i(\psi) &= [\Phi(-m_i - \psi)]^{1\{c_i=0,s_i=0\}} \times [1 - \Phi(m_i - \psi)]^{1\{c_i=1,s_i=0\}} \\ &\quad \times [\Phi(-\psi) - \Phi(-m_i - \psi)]^{1\{c_i=0,s_i=1\}} \times [\Phi(m_i - \psi) - \Phi(-\psi)]^{1\{c_i=1,s_i=1\}}. \end{aligned}$$

Posterior summaries of ψ are reported in Table B14.

H Priors

We estimate the structural model (Section 6) in a Bayesian framework and therefore have to specify the distributions from which we draw parameters

$$\begin{aligned}\boldsymbol{\theta}^{\text{deworming}} &= \{ \beta_{\text{control}} \sim N(0, 1), \beta_z \sim N(0, 0.25), \delta \sim N(0, 0.25), \\ &\quad \lambda \sim N^+(0, 0.25), \sigma_u \sim \text{Inv Gamma}(3.3, 1.1), \rho \sim N^+(0, 0.001) \}, \\ \boldsymbol{\theta}^{\text{wtp}} &= \{ \psi \sim N(0, 2) \}, \\ \boldsymbol{\theta}^{\text{observability}} &= \{ \beta_z^O, \gamma_z^O \sim N(0, 0.125) \}.\end{aligned}$$

For the treatment effect parameters (β 's) we adopt weakly informative priors centered at zero, reflecting the view that, absent data, incentives are assumed not to change observability beliefs or confer extra private utility. In other words, we do not rule out social signaling for any of the incentives; instead, we assume ex ante that their effects do not differ.

Two additional parameters are difficult to learn from the data. We constrain ρ to be positive and close to zero with a strongly informative prior, reflecting limited prior knowledge about consumption value differences between bracelets and calendars. Likewise, we place an inverse-Gamma prior on σ_u that keeps this scale parameter away from zero for numerical stability. Both choices are conservative: larger σ_u flattens Δ^* , making signaling harder to detect, while a smaller ρ assigns more of the bracelet effect to private utility rather than signaling.

I Alternative Structural Models

This appendix estimates three departures from the baseline specification in Section 6. Each alternative model preserves the main signaling mechanism—how observability generates social image payoffs—and the baseline distribution of intrinsic types. The only change is in how travel distance enters private utility and/or the social image term.

I.1 Private Distance Costs and Community Social Image Returns

In many settings adults know their own walking distance but form beliefs about others using a public reference point such as the community centroid. Let d_{ic} be household i 's distance and \bar{d}_c the centroid distance of community c . The fixed-point equation therefore uses \bar{d}_c ,

$$w^*(z, \bar{d}_c) = -(z\beta_z - \delta\bar{d}_c) - \mu(z, \bar{d}_c) \Delta(w^*(z, \bar{d}_c)),$$

while individual i deworms if

$$z\beta_z - \delta d_{ic} + \mu(z, \bar{d}_c) \Delta(w^*(z, \bar{d}_c)) + w_i > 0.$$

Substituting $w^*(\cdot)$ shows that the deworming probability is

$$\Pr(\text{Deworm}_{ic} | z, d_{ic}, \bar{d}_c) = F_w(-w^*(z, \bar{d}_c) + \delta(\bar{d}_c - d_{ic})).$$

That is, an individual's decision to deworm equals the community cut-off type plus an offset depending on how much farther (or closer) the household is from the PoT than the centroid. Table I1 reports posterior average treatment effects: the Bracelet and Calendar estimates are very similar to those in the main specification, whereas Ink becomes negative even in Far communities and its Far-Close becomes smaller and insignificant.

I.2 Full Information

Alternatively, suppose every household knows the exact distance for every other household. Then both the private cost and the social image return depend on the household specific distance d_{ic} . Posterior average treatment effects (Table I2) are again stable for Bracelet and Calendar, whereas the Ink effect remains negative and is now flat across distance.

I.3 Excluding Spatially Dispersed Clusters

The centroid may mismeasure cost in communities that are geographically spread out. We drop clusters whose mean-squared household distance from the centroid exceeds 0.5 km^2 and re-estimate the main specification. Results (Table I3) are nearly identical to the main specification.

Table I1: Structural Average Treatment Effects: Private Distance Costs with Community-Level Social Image Returns

Dependent variable: Take-up	Structural			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.065 (0.041, 0.09)	0.041 (0.016, 0.066)	0.098 (0.072, 0.128)	0.057 (0.035, 0.084)
Calendar	0.027 (0.004, 0.051)	0.025 (-0.005, 0.052)	0.035 (0.014, 0.058)	0.01 (-0.007, 0.034)
Ink	-0.031 (-0.055, -0.006)	-0.041 (-0.072, -0.011)	-0.025 (-0.048, -0.002)	0.016 (-0.008, 0.044)
Control mean	0.34 (0.322, 0.36)	0.407 (0.382, 0.43)	0.258 (0.239, 0.277)	-0.149 (-0.171, -0.127)
Δ (Signal–No Signal)	-0.01 (-0.026, 0.005)	-0.025 (-0.041, -0.008)	0.001 (-0.018, 0.022)	0.026 (0.009, 0.044)
Δ (Bracelet–Calendar)	0.038 (0.023, 0.053)	0.016 (0.002, 0.032)	0.063 (0.044, 0.087)	0.046 (0.031, 0.067)

Notes: Estimates are from the same structural specification as Table B18, except that the private cost term now uses each household's own distance to the point of treatment while social image returns are computed at the community centroid. Point estimates are posterior means; 95 percent credible intervals appear in parentheses.

Table I2: Structural Average Treatment Effects: Private Distance Costs with Household-Level Social Image Returns

Dependent variable: Take-up	Structural			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.064 (0.039, 0.088)	0.045 (0.019, 0.07)	0.09 (0.063, 0.117)	0.045 (0.028, 0.064)
Calendar	0.025 (0.002, 0.05)	0.021 (-0.007, 0.05)	0.034 (0.009, 0.056)	0.013 (-0.003, 0.032)
Ink	-0.031 (-0.054, -0.007)	-0.035 (-0.061, -0.004)	-0.035 (-0.062, -0.01)	0 (-0.02, 0.023)
Control mean	0.342 (0.325, 0.36)	0.405 (0.384, 0.427)	0.264 (0.245, 0.282)	-0.141 (-0.16, -0.122)
Δ (Signal–No Signal)	-0.009 (-0.028, 0.008)	-0.016 (-0.036, 0.005)	-0.007 (-0.027, 0.013)	0.009 (-0.002, 0.023)
Δ (Bracelet–Calendar)	0.039 (0.021, 0.055)	0.024 (0.006, 0.042)	0.056 (0.035, 0.077)	0.032 (0.019, 0.046)

Notes: Estimates are from the same structural specification as Table B18, except that both the private cost term and the fixed-point calculation for social image returns are based on each household's own distance to the treatment site. Point estimates are posterior means; 95 percent credible intervals appear in parentheses.

Table I3: Structural Average Treatment Effects: Excluding Spatially Dispersed Clusters

Dependent variable: Take-up	Structural			
	Combined (1)	Close (2)	Far (3)	Far - Close (4)
Bracelet	0.07 (0.046, 0.094)	0.056 (0.029, 0.081)	0.089 (0.061, 0.118)	0.033 (0.012, 0.058)
Calendar	0.041 (0.018, 0.067)	0.041 (0.013, 0.07)	0.042 (0.02, 0.066)	0.001 (-0.016, 0.022)
Ink	-0.025 (-0.052, 0)	-0.042 (-0.071, -0.012)	-0.004 (-0.028, 0.02)	0.038 (0.018, 0.063)
Control mean	0.334 (0.315, 0.354)	0.398 (0.374, 0.422)	0.252 (0.23, 0.276)	-0.147 (-0.172, -0.12)
Δ (Signal–No Signal)	-0.019 (-0.037, 0)	-0.034 (-0.051, -0.016)	0.001 (-0.023, 0.023)	0.035 (0.018, 0.053)
Δ (Bracelet–Calendar)	0.029 (0.013, 0.048)	0.015 (0.001, 0.031)	0.047 (0.022, 0.071)	0.032 (0.013, 0.053)

Notes: Estimates are based on the same structural specification as Table B18, except that clusters whose mean-squared household-to-centroid distance exceeds 0.5km^2 are excluded from the sample. Point estimates are posterior means; 95 percent credible intervals appear in parentheses.

J Sensitivity to the Distribution of w

The main specification assumes that the composite type $w = v + u$ is Gaussian. This section shows that our qualitative conclusions are robust to departures from normality.

Conditions for a Social Multiplier. As in [Benabou and Tirole \(2025\)](#), a social multiplier different from one requires (i) non-zero observability, $\mu(z, d) \neq 0$, and (ii) a non-flat social image schedule, $\Delta'(w) \neq 0$. The latter fails, for example, under a uniform type distribution, in which gains in honor are exactly offset by losses in stigma.

A potential concern. Suppose instead that w is bimodal—one mass of adults highly motivated to deworm and another strongly opposed—and that the local mix of types happens to covary with distance. Could such heterogeneous private values alone generate the distance-by-signal interactions we observe?

Simulation design. We simulate two data-generating processes: Unimodal: $w \sim \mathcal{N}(0, 1)$ (our baseline) and Bimodal: an equally weighted mixture of $\mathcal{N}(-\eta, 1)$ and $\mathcal{N}(+\eta, 1)$ with η large enough that the two modes are distinct. For each process we vary the travel cost b and compute (i) the equilibrium participation rate $\bar{y}(b)$, (ii) the social image return $\Delta(w^*(b))$, and (iii) the implied social multiplier $M(b)$.

Key findings.

- **Flat regions occur only under a bimodal type density.** Figure [J1a](#) contrasts the unimodal (Gaussian) and bimodal mixtures we simulate. When the cut-off type moves through the density gap between the two modes, the participation curve $\bar{y}(b)$ in Figure [J1b](#) becomes locally flat. Over exactly the same distance band the social image return $\Delta(w^*(b))$ is constant (Figure [J1c](#)), so the social multiplier collapses to $M(b) = 1$ (Figure [J1d](#)). No such flat segment appears in the empirical distance range, ruling out bimodality driven substitution as the sole explanation for our findings.
- **Any local curvature in the type density is sufficient to generate $M \neq 1$.** With the Gaussian benchmark, the slope of the density is never zero in the experimental window, so changes in distance always shift $\Delta(w^*(b))$ and hence $M(b)$. The conclusion is distribution robust: any density that is not locally flat yields amplification or mitigation effects similar to those we estimate.
- **Large unobserved noise would also flatten $M(b)$ —but is rejected by the data.** A large idiosyncratic shock σ_u would effectively smooth the density even if v were Gaussian, pushing $M(b) \approx 1$. Our posterior places little mass on such high variance values; the estimated Δ retains marked curvature, consistent with the signal-by-distance interactions in the reduced form.

Implication. The simulations show that a social multiplier equals one only when the type density is locally flat—either because the cut-off moves through the gap between two modes, or because idiosyncratic noise σ_u is so large that the effective density is nearly uniform. In the data we see neither (i) any distance band with flat take-up nor (ii) estimates of σ_u in the flattening range. Hence heterogeneous private utility substitution alone cannot explain the stronger distance interaction for highly observable incentives.

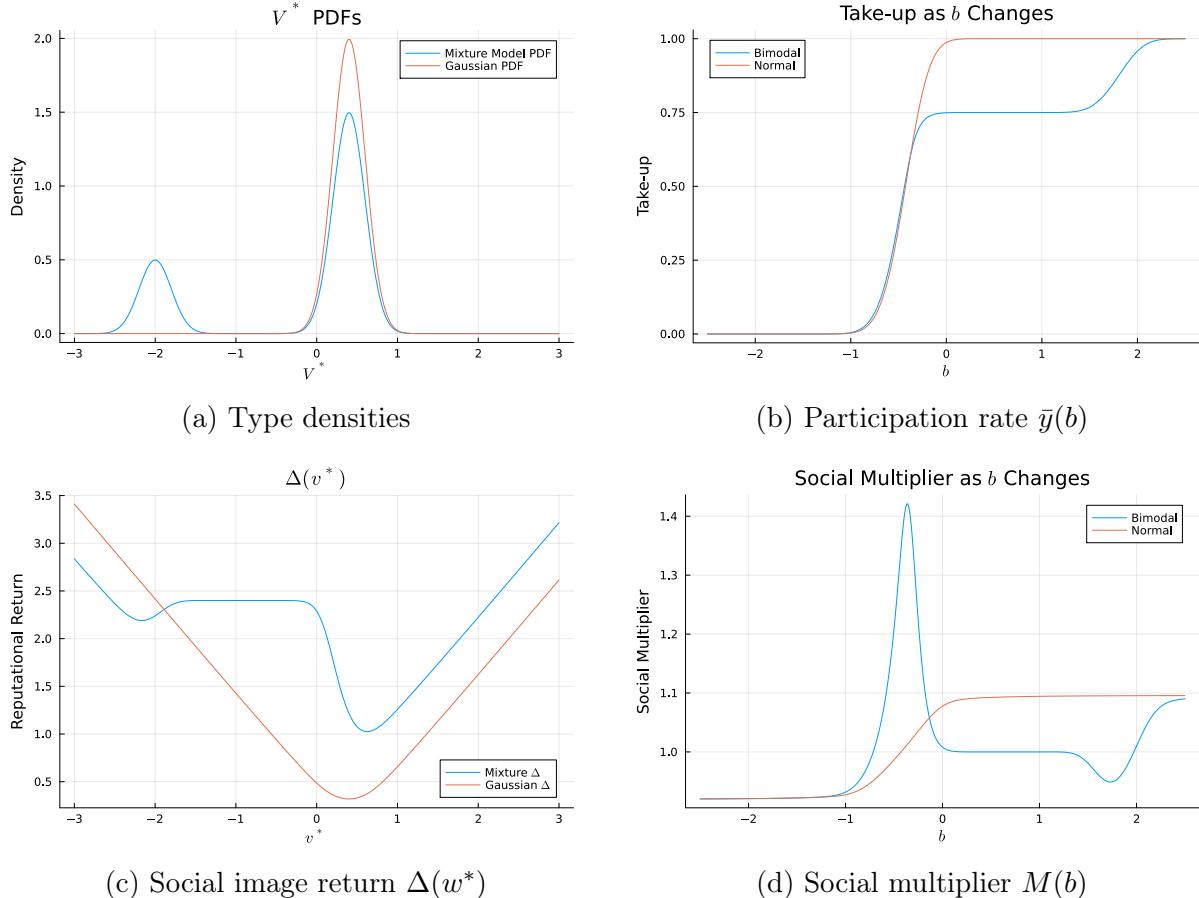


Figure J1: Simulation under Gaussian (red) versus Bimodal (blue) Type Distributions

Notes: Panel (a) plots the assumed densities for w . Panel (b) shows the participation rate as the private cost b falls (distance decreases). Panel (c) traces the implied social image return. Panel (d) gives the resulting social multiplier. Flat segments—visible only in the bimodal case—occur where the cut-off type moves through the density gap between the two modes; in those intervals $\Delta(w^*)$ and $M(b)$ are constant at one.

K Policymaker PoT Optimization

Table K1: Policymaker PoT Optimization

Observability	Number of PoTs	Mean take-up	Mean distance (km)
Panel A: Experimental allocation			
Control	144	0.33 (0.31, 0.346)	1.2
Panel B: Policymaker allocation, 3.5km			
Control	107 (100, 114)	0.33 (0.329, 0.331)	1.26 (1.11, 1.411)
Bracelet	90 (79, 102)	0.33 (0.329, 0.33)	1.68 (1.364, 1.957)
Bracelet social image returns at 0.5km	96 (86, 105)	0.33 (0.329, 0.331)	1.52 (1.297, 1.758)
No social image returns	144 (144, 144)	0.174 (0.101, 0.259)	0.58 (0.58, 0.58)
Panel C: Policymaker allocation, 4.5km			
Control	107 (100, 114)	0.33 (0.329, 0.332)	1.27 (1.123, 1.416)
Bracelet	90 (79, 102)	0.33 (0.329, 0.331)	1.68 (1.374, 1.957)
Bracelet social image returns at 0.5km	96 (86, 105)	0.33 (0.329, 0.331)	1.51 (1.299, 1.758)
No social image returns	144 (144, 144)	0.174 (0.101, 0.259)	0.58 (0.58, 0.58)
Panel D: Policymaker allocation, 5.5km			
Control	106 (99, 113)	0.33 (0.329, 0.331)	1.33 (1.145, 1.507)
Bracelet	90 (79, 102)	0.33 (0.329, 0.331)	1.67 (1.364, 1.953)
Bracelet social image returns at 0.5km	96 (86, 105)	0.33 (0.329, 0.332)	1.52 (1.294, 1.758)
No social image returns	144 (144, 144)	0.174 (0.101, 0.259)	0.58 (0.58, 0.58)
Panel E: Policymaker allocation, 10km			
Control	106 (99, 113)	0.33 (0.329, 0.331)	1.51 (1.257, 1.755)
Bracelet	90 (79, 102)	0.33 (0.329, 0.332)	1.69 (1.374, 2.007)
Bracelet social image returns at 0.5km	96 (86, 105)	0.33 (0.329, 0.332)	1.57 (1.309, 1.876)
No social image returns	144 (144, 144)	0.174 (0.101, 0.259)	0.58 (0.58, 0.58)

Notes: This table reports the number of Points of Treatment (PoTs), the mean community–PoT travel distance, and mean take-up by observability regime and distance cap. Point estimates are posterior medians; parentheses contain 95 percent credible intervals. In Panel A (experimental allocation), the number of PoTs and the mean community–PoT distance are fixed by design, so credible intervals are omitted. The remaining panels report the policymaker’s solution: for each of 200 posterior draws from the structural model, we solve an integer program that minimizes the number of PoTs subject to the coverage target and a maximum travel distance (with caps from 3.5 to 10 km); the table reports the posterior medians and credible intervals of the resulting allocations.

L $\Delta(w(b, c))$ Closed Form Derivation

In this section, we show how we generate the closed form solution for $\Delta(w)$ using the Gaussian distribution - both the unbounded case used in the paper, and the bounded case for completeness since [Benabou and Tirole \(2025\)](#) place bounds on v .

L.1 $\Delta(w(b, c))$ Derivation

We want to find:

$$\Delta(b, c) = \frac{-1}{F_w(w(b, c))[1 - F_w(w(b, c))]} \int_{-\infty}^{\infty} v F_u(w(b, c) - v) f_v(v) dv$$

We suppress $w(b, c)$ notation and just write w . $V \sim N(0, 1)$, $U \sim N(0, \sigma)$ and the two are independent.

Ignoring the first fraction:

$$\begin{aligned} & \int v F_u(w - v) f_v(v) dv \\ &= \int \underbrace{v\phi(v)}_{d(-\phi(v))} \Phi\left(\frac{w-v}{\sigma}\right) dv \\ &= \int -\Phi\left(\frac{w-v}{\sigma}\right) d\phi(v) \\ &= \left[-\phi(v)\Phi\left(\frac{w-v}{\sigma}\right) \right]_{-\infty}^{\infty} - \int \frac{-1}{\sigma} \phi\left(\frac{w-v}{\sigma}\right) (-)\phi(v) dv \quad (\text{Integration by parts}) \\ &= \left[-\phi(v)\Phi\left(\frac{w-v}{\sigma}\right) \right]_{-\infty}^{\infty} - \frac{1}{\sigma} \int \phi\left(\frac{w-v}{\sigma}\right) \phi(v) dv \end{aligned}$$

Now focus solely on the remaining integral:

$$\begin{aligned}
\int \phi\left(\frac{w-v}{\sigma}\right) \phi(v) dv &= \frac{1}{2\pi} \int \exp \left[-\frac{1}{2} \left(\left(\frac{w-v}{\sigma} \right)^2 + v^2 \right) \right] dv \\
&= \frac{1}{2\pi} \int \exp \left[-\frac{1}{2} \left(\left(\frac{w}{\sigma} \right)^2 + v^2 \frac{\sigma^2+1}{\sigma^2} - \left(\frac{2wv}{\sigma} \right)^2 \right) \right] dv \\
&= \frac{1}{2\pi} \int \exp \left[-\frac{1}{2} \left(\left(\frac{w}{\sigma} \right)^2 + \frac{\sigma^2+1}{\sigma^2} \left[\left(v - \frac{w}{\sigma^2+1} \right)^2 - \frac{w^2}{(\sigma^2+1)^2} \right] \right) \right] dv \\
&\quad (\text{Completing The Square}) \\
&= \frac{1}{2\pi} \int \exp \left[-\frac{1}{2} \left(\frac{w^2}{\sigma^2+1} + \frac{\sigma^2+1}{\sigma^2} \left(v - \frac{w}{\sigma^2+1} \right)^2 \right) \right] dv \\
&= \frac{1}{2\pi} \exp \left(-\frac{1}{2} \frac{w^2}{\sigma^2+1} \right) \int \exp \left[-\frac{1}{2} \frac{\sigma^2+1}{\sigma^2} \left(v - \frac{w}{\sigma^2+1} \right)^2 \right] dv
\end{aligned}$$

The first exponential term is just a function of w and σ whilst the the second term is very nearly a normal pdf:

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}} \int \exp \left[-\frac{1}{2} \frac{\sigma^2+1}{\sigma^2} \left(v - \frac{w}{\sigma^2+1} \right)^2 \right] dv \\
&= \sqrt{\frac{\sigma^2}{\sigma^2+1}} \times \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\frac{\sigma^2}{\sigma^2+1}}} \int \exp \left[-\frac{1}{2} \left(\frac{v - \frac{w}{\sigma^2+1}}{\sqrt{\frac{\sigma^2}{\sigma^2+1}}} \right)^2 \right] dv \\
&= \sqrt{\frac{\sigma^2}{\sigma^2+1}}
\end{aligned}$$

Putting this back together:

$$\begin{aligned}
&\left[-\phi(v)\Phi\left(\frac{w-v}{\sigma}\right) \right]_{-\infty}^{\infty} - \frac{1}{\sigma} \int \phi\left(\frac{w-v}{\sigma}\right) \phi(v) dv \\
&= \left[-\phi(v)\Phi\left(\frac{w-v}{\sigma}\right) \right]_{-\infty}^{\infty} - \frac{1}{\sigma} \frac{\exp\left[-\frac{1}{2} \frac{w^2}{\sigma^2+1}\right]}{\sqrt{2\pi}} \times \sqrt{\frac{\sigma^2}{\sigma^2+1}} \\
&= \frac{-1}{\sigma} \frac{\exp\left[-\frac{1}{2} \frac{w^2}{\sigma^2+1}\right]}{\sqrt{2\pi}} \times \sqrt{\frac{\sigma^2}{\sigma^2+1}}
\end{aligned}$$

Reintroducing the first fraction:

$$\begin{aligned}\Delta(w(b, c)) &= \frac{-1}{F_w(w(b, c))[1 - F_w(w(b, c))]} \int_{-\infty}^{\infty} v F_u(w(b, c) - v) f_v(v) dv \\ &= \frac{1}{\Phi\left(\frac{w}{\sqrt{\sigma^2 + 1}}\right) \left[1 - \Phi\left(\frac{w}{\sqrt{\sigma^2 + 1}}\right)\right]} \times \frac{1}{\sigma} \frac{\exp\left[-\frac{1}{2} \frac{w^2}{\sigma^2 + 1}\right]}{\sqrt{2\pi}} \times \sqrt{\frac{\sigma^2}{\sigma^2 + 1}}\end{aligned}$$

L.2 $\Delta'(w(b, c))$ Derivation

$$\begin{aligned}\Delta'[w] &= \frac{\int_{-\infty}^{\infty} v f_u(w - v) f_v(v) dv + f_w(w) [1 - 2F_w(w)] \Delta[w]}{F_w(w)(1 - F_w(w))} \\ &= \frac{\int_{-\infty}^{\infty} v \frac{1}{\sigma} \phi\left(\frac{w-v}{\sigma}\right) \phi(v) dv + \frac{1}{\sqrt{1+\sigma^2}} \phi\left(\frac{w}{\sqrt{1+\sigma^2}}\right) \left[1 - 2\Phi\left(\frac{w}{\sqrt{1+\sigma^2}}\right)\right] \Delta[w]}{\Phi\left(\frac{w}{\sqrt{1+\sigma^2}}\right) \left(1 - \Phi\left(\frac{w}{\sqrt{1+\sigma^2}}\right)\right)}\end{aligned}$$

Focusing on the integral and recognising we have derived this above:

$$\begin{aligned}\int_{-\infty}^{\infty} v \frac{1}{\sigma} \phi\left(\frac{w-v}{\sigma}\right) \phi(v) dv &= \frac{1}{\sigma} \frac{\exp\left(-\frac{1}{2} \frac{w^2}{\sigma^2 + 1}\right)}{\sqrt{2\pi}} \Sigma \times \frac{1}{\sqrt{2\pi}} \frac{1}{\Sigma} \int_{-\infty}^{\infty} v \exp\left(-\frac{1}{2} \left(\frac{v-\mu}{\Sigma}\right)^2\right) dv\end{aligned}$$

Where $\mu = \frac{w}{\sigma^2 + 1}$, $\Sigma = \sqrt{\frac{\sigma^2}{\sigma^2 + 1}}$. Define:

$$H = \frac{1}{\sigma} \frac{\exp\left(-\frac{1}{2} \frac{w^2}{\sigma^2 + 1}\right) \Sigma}{\sqrt{2\pi}}$$

Now perform change of variables:

$$\begin{aligned}\frac{v - \mu}{\Sigma} &= y \\ dy &= dv \frac{1}{\Sigma}\end{aligned}$$

Giving:

$$\begin{aligned}H \frac{1}{\sqrt{2\pi}} \frac{\Sigma}{\Sigma} \int_{-\infty}^{\infty} (\Sigma y + \mu) \exp\left(-\frac{1}{2} y^2\right) dy &= H\Sigma [-\phi(y)]_{-\infty}^{\infty} + H\mu \\ &= H\mu\end{aligned}$$

Plugging this back into the formula:

$$\Delta'[w] = \frac{H\mu + \frac{1}{\sqrt{1+\sigma^2}}\phi\left(\frac{w}{\sqrt{1+\sigma^2}}\right)\left[1 - 2\Phi\left(\frac{w}{\sqrt{1+\sigma^2}}\right)\right]\Delta[w]}{\Phi\left(\frac{w}{\sqrt{1+\sigma^2}}\right)\left(1 - \Phi\left(\frac{1}{\sqrt{1+\sigma^2}}\right)\right)}$$

L.3 Bounded $\Delta(w(b, c))$ Derivation

Now we go back and bound types between \underline{v}, \bar{v} . There are three things we need to do here: keep any terms that dropped out due to integration limits being infinite before, replace the normal pdf with the truncated normal pdf for V , and calculate the convolution of $W \sim V + U$ when V is truncated normal.

$$\begin{aligned} \Delta(w(b, c)) &= \frac{-1}{F_w(w)(1 - F_w(w))} \frac{1}{\Phi(\bar{v}) - \Phi(\underline{v})} \\ &\times \left(\left[-\phi(v)\Phi\left(\frac{w-v}{\sigma}\right) \right]_{\underline{v}}^{\bar{v}} - \Gamma \left[\Phi\left(\frac{v - \frac{w}{\sigma^2+1}}{\frac{\sigma}{\sqrt{\sigma^2+1}}}\right) \right]_{\underline{v}}^{\bar{v}} \right) \end{aligned}$$

Where:

$$\Gamma = \frac{1}{\sigma} \exp\left(-\frac{1}{2} \frac{w^2}{1 + \sigma^2}\right) \times \sqrt{\frac{\sigma^2}{\sigma^2 + 1}}$$

and:

$$\begin{aligned} F_w(w) &= \int_{\underline{v}}^{\bar{v}} \Phi\left(\frac{w-t}{\sigma}\right) \frac{\phi(t)}{\Phi(\bar{v}) - \Phi(\underline{v})} dt \\ (\Phi(\bar{v}) - \Phi(\underline{v}))F_w(w) &= \int_{\underline{v}}^{\bar{v}} \Phi\left(\frac{w-t}{\sigma}\right) \phi(t) dt \end{aligned}$$

That is, the convolution of two independent r.v.s. It can be shown that the RHS:

$$\begin{aligned}
RHS &= \frac{1}{2} (\Phi(z_h) - \Phi(z_l)) \\
&\quad - \frac{1}{2} \left[\frac{\mu}{z_h} < 0 \right] \\
&\quad + \frac{1}{2} \left[\frac{\mu}{z_l} < 0 \right] \\
&\quad - T \left(z_h, \frac{h}{z_h} \right) \\
&\quad + T \left(z_l, \frac{l}{z_l} \right) \\
&\quad - T \left(\frac{\mu}{\rho}, \frac{\mu\sigma + z_h\rho^2}{\mu} \right) \\
&\quad + T \left(\frac{\mu}{\rho}, \frac{\mu\sigma + z_l\rho^2}{\mu} \right)
\end{aligned}$$

where: $U \sim N(0, \gamma)$, $\mu = \gamma^{-1}z$, $\sigma = -\gamma^{-1}$, $\rho = \sqrt{1 + \sigma^2}$, $z_l = \frac{l-\mu}{\sigma}$, $z_h = \frac{h-\mu}{\sigma}$, $h = \frac{z-\bar{v}}{\gamma}$

Using a change of variables on the above gives:

$$\int_l^h \Phi(x) \frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) dx$$

where T stands for Owen's T.