Policy Research Working Paper 10967

Can Facebook Ads Prevent Malaria?

Two Field Experiments in India

Dante Donati
Nandan Rao
Victor Orozco-Olvera
Ana Maria Muñoz-Boudet

WORLD BANK GROUP
Development Economics
Development Impact Group
November 2024

Policy Research Working Paper 10967

Abstract

This study uses a cluster randomized controlled trial to evaluate the impact of a nationwide malaria prevention advertising campaign delivered through social media in India. Ads were randomly assigned at the district level, and the study relies on data from two independently recruited samples (8,257 individuals) and administrative records. Among users residing in solid (concrete) dwellings, where malaria risk is lower, the campaign led to an 11 percent increase in mosquito net usage and a 13 percent increase in timely treatment seeking. Self-reported malaria incidence decreased by 44 percent. Consistently, recorded health facility data indicate a reduction in urban monthly incidence of 6.2 cases per million people, corresponding to 30 percent of the overall monthly incidence rate of malaria. Conversely, the study finds no impact on households living in non-solid dwellings, which face higher malaria risk, nor among rural settlements where such dwellings are more prevalent. To disentangle if this lack of impact stems from ineffective content or insufficient reach, an individual-level trial was conducted (1,542 individuals), ensuring campaign exposure for both household types. The findings indicate an increase in bed net usage and timely treatment seeking for both groups, underscoring the need for improved targeting in social media campaigns to fulfill public health goals..

This paper is a product of the Development Impact Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at dd3137@gsb.columbia.edu, amunozboudet@worldbank.org, vorozco@worldbank.org or nandan@vlab.digital. A verified reproducibility package for this paper is available at http://reproducibility.worldbank.org, click **here** for direct access.

Can Facebook Ads Prevent Malaria? Two Field Experiments in India*

Dante Donati†
Nandan Rao‡
Victor Orozco-Olvera§
Ana Maria Muñoz Boudet¶

*This research is part of the entertainment-education program of the World Bank's Development Impact Department (DIME). Computational replicability was verified by DIME Analytics. We thank our study partners at Facebook's Campaigns for a Healthier World initiative and Upswell, in particular Nisha Deolalikar, Sarah Francis, and Drew Bernard. We also thank Malaria No More for their collaboration. We are grateful to attendees at CEGA-DIME Measuring Development Conference, Centro de Investigacion y Docencia Economicas seminar, Columbia Business School seminar, Marketing Science Annual Conference, Marketing Science DEI Conference, MIT's Annual Conference on Digital Experimentation, Quantitative Marketing and Economics Conference, SSRC Workshop on the Economics of Social Media, Universitat Autonoma de Barcelona seminar, and Yale School of Management seminar. The research received ethical clearance from Columbia University IRB (AAAU6317). All errors are our own.

†Columbia Business School and CESifo; dd3137@gsb.columbia.edu;
‡Virtual Lab and UAB
§DIME, The World Bank
¶Gender Group, The World Bank

1 Introduction

With over half of the world's population engaged in social media (We Are Social, 2023), governments and organizations increasingly leverage these platforms in their social marketing campaigns to influence individuals' beliefs, attitudes and real-world behaviors.[1]

Typically, the effectiveness of these campaigns is assessed using online engagement and direct response metrics, such as likes, clicks, or website activity, along with "Brand Lift" surveys that measure short-term effects (Athey et al. 2023; Shawky et al. 2019). However, while these methods may be adequate to evaluate ads' impact on purchase intent and sales (e.g., Bart et al. 2014; Johnson et al. 2017), they fall short when assessing the longer-term objectives of social marketing campaigns, such as those promoting public health. These campaigns require outcome data that go beyond what can be measured on ad platforms and may need offline data, as their goal is to influence offline behaviors. Moreover, even when ad content is proven effective, campaigns may fail to reach the intended audience due to the ad delivery optimization algorithms used by social media platforms (Lambrecht and Tucker 2019; Sapiezynski et al. 2022). Additionally, recent changes in privacy policies by regulators and tech companies, which restrict access to third-party data that facilitates targeting, can further hinder both the success and assessment of these advertising efforts (Goldfarb and Tucker 2011*b*; Gordon et al. 2021; Johnson et al. 2020).

This paper addresses these measurement and targeting challenges by employing a newly-developed experimental methodology and toolkit (Rao et al. 2020) to evaluate the impact of a nationwide malaria prevention ad campaign and its heterogeneous treatment effects across sub-populations in India. By linking survey responses and administrative records with (potential) ad exposure data, our design enables us to assess the causal impact of online public health messages on offline behaviors and real-world outcomes, such as consistent bed net usage, timely treatment-seeking, and malaria incidence. Additionally, we investigate whether the ad content reached its intended audience by explicitly targeting and oversampling respondents at higher risk of contracting the disease. Our findings show that the ad campaign was effective at improving health outcomes and reducing malaria incidence only among individuals who were ex-ante less likely to contract the disease, underscoring the need for improved measurement and targeting in public health campaigns.

Malaria is transmitted by the bite of infected mosquitoes and disproportionately affects poor and rural people in India (Dev et al., 2004). The ad campaign was designed by the public health NGO Malaria No More (MNM) and used behaviorally-informed ads to provide information and encourage individuals to take preventive measures (e.g., sleeping under mosquito nets) and response actions (e.g., timely seeking treatment).[2] It was run on Meta platforms and aimed to achieve its public health objectives by maximizing engagement with the ads and MNM's pages. While it succeeded in achieving these engagement metrics (e.g., the ads reached 130 million people and garnered 46

[1]Social marketing entails applying commercial marketing techniques to influence the behavior of target audiences to enhance personal welfare and societal well-being (Andreasen 1994; Kotler and Zaltman 1971).

[2]For example, 75% of households in our baseline survey have mosquito nets, but only 70% of them use them regularly, implying that almost 50% of households remain unprotected to some degree.

million reactions on Facebook and Instagram across 22 Indian states),[3] our study focuses on its offline behavioral impacts: "Did the campaign work?" and "Did it work for those most at risk?"

We adopt a multi-stage methodology. We start by recruiting survey respondents stratifying across 80 districts in the three study states: Uttar Pradesh, Jharkhand, and Chhattisgarh. Specifically, we use ads on Facebook and Instagram, where we offer mobile credit as an incentive to fill out a baseline survey using a chatbot integrated with Facebook Messenger. During the initial phase of recruitment, we conduct formative research to identify households most at risk of contracting malaria and ensure their inclusion in the study sample. After considering various demographic and socio-economic characteristics, we find that dwelling type is the most robust household-level predictor of malaria incidence.

Respondents living in solid houses (i.e., dwellings made of stones, bricks, cement, and concrete) are 43% less likely to report recent malaria cases compared to those living in semi-solid or non-solid dwellings (i.e., made of mud, straw, and tin).[4] Beyond being a strong predictor of malaria risk, dwelling type also provides a convenient mapping into administrative data on malaria incidence, which are disaggregated across urban and rural health facilities. In India, non-solid and semi-solid dwellings make up approximately 15% of housing in urban areas and around 50% in rural areas (India IIPS, 2022).

During the initial recruitment phase, respondents living in solid dwellings were more likely to join our survey. To correct for this imbalance, we stratify our recruitment by dwelling type. Since Facebook does not offer the option to micro-target ads based on housing characteristics, we use the Facebook Lookalike Audience tool to target individuals living in non-solid dwellings. We show that this technique works to increase the number of respondents living in non-solid houses.

We apply this technique to recruit survey respondents into a longitudinal sample (N=5,205) and a cross-sectional sample (N=3,052), totaling 8,257 individuals in the three study states, balancing across 80 districts and the two dwelling types. To estimate the population-level effects of the malaria-prevention campaign, we work with the MNM advertising team and conduct a cluster randomized controlled trial, where we expose half of the 80 districts to the same ad content and delivery strategy used in the nationwide campaign. It is important to note that MNM employed conventional targeting techniques (i.e., based on socio-demographic characteristics available on the ad platform), optimizing ad delivery for engagement. Survey respondents in treatment districts were not necessarily exposed to the campaign, hence our results estimate intent-to-treat (ITT) effects.

We rely on the longitudinal sample to assess the impact of the campaign on a range of self-reported attitudes, behaviors, and malaria-related outcomes *during* the campaign. This process was accomplished via an "experience" sampling method that gathered data across several survey rounds, for a maximum of 10 waves. We find an overall moderate impact of the intervention on bed net use among mosquito net owners. Overall, the likelihood of sleeping under a bed net both at the individual and household level increased by 3 percentage points (p.p.), a 4.5% increase compared

[3]The campaign cost 202,000 USD and reached users an average of 3.6 times over a 6-month period.

[4]In the context of India, houses constructed with durable materials are referred to as *pucca* dwellings, while those made from temporary or non-permanent materials are known as *kutcha* dwellings.

to the control group. At the same time, we find evidence of significant heterogeneous treatment effects across dwelling categories. For those living in solid houses, bed net use increase by 9-to-11% compared to control, yet we find no significant effect on those living in non-solid dwellings. These effects are amplified in the last 3 months of the campaign, and for individuals with moderate levels of concerns about getting malaria. While incidence of high fever and malaria during the campaign was not affected, the campaign increased the likelihood of solid dwellers seeking medical assistance promptly (within 24 hours) in the event of a fever by 13%.

We use the cross-sectional sample to study impacts 1-to-3 months *after* the campaign ended. The results indicate that prolonged exposure to the ad campaign, which led to cumulative behavior change, resulted in decreased malaria incidence. Self-reported household-level malaria cases decreased among solid dwellers by 44 % at the extensive margin, and by 53% at the intensive margin, compared to control, but not among those living in non-solid houses. The reduction in malaria incidence is larger among bed net owners and more so among households with at least 1 mosquito net per every 2 members, emphasizing the importance of technology adoption to prevent infection. Our survey-based findings are bolstered by the evidence of baseline balance in demographic and socio-economic characteristics across experimental arms and the consistency of impact estimates across different model specifications and estimators.

We complement this survey evidence with administrative health-facility data on recorded malaria cases from the Indian Health Management Information System (HMIS). Over the 9-month period following the launch of the campaign, monthly malaria incidence in urban areas decreased by 6.2 cases per million people, which is 30% of the overall monthly malaria incidence rate in the pre-treatment period. By contrast, and consistent with our survey data, incidence did not significantly change in rural settlements, where non-solid dwellings are substantially more common. Moreover, while the campaign had no impact on the number of tests conducted to diagnose malaria, the urban positivity rates (positive cases/tests) decreased by 2 p.p. and the effect is null in rural areas. These results are robust to different model specifications and estimators.

Overall, our evidence indicates that the MNM campaign benefited households residing in solid dwellings and in urban settlements, but less so those that are particularly susceptible to malaria. The potential cause behind this distinct effect could stem from two factors: (1) content effectiveness (where, upon receiving the ad, it only influenced households residing in solid dwellings) or (2) targeting (the campaign did not sufficiently reached households living in non-solid dwellings). To discern which mechanism might underlie this effect, first we consider ad recall outcomes at the end of the campaign. If the ads reached both solid and non-solid dwellers at similar frequencies, then one would expect both audiences to equally recall them. Yet, we find that ad recall is larger only among treated individuals in solid houses, suggesting that non-solid dwellers were not sufficiently exposed to the ads.

To corroborate this mechanism, we nest an individual-level *feed* experiment within our cross-sectional sample. We use Facebook's Custom Audience tool to directly retarget ads to a randomly selected half of our survey respondents. Over a 2-week campaign, we reached individuals living in both dwelling types on their Facebook feeds, with an average frequency of 4.9 ads per person.[5] Two weeks after the end of this remarketing campaign, we sent them a follow-up survey (N=1,542) asking about bed net use and treatment seeking in case of symptoms. We find that the campaign was equally effective for users residing in both types of dwellings, reporting an 8-pp-increase in the probability of sleeping under a bed net the previous night, and a 2.5-pp-increase in timely treatment seeking in the event of a fever.

We conclude that the ad content itself was effective at informing people and persuading them to take preventative and response actions to fight malaria. The campaign was also effective at large scale for those it reached, but it did not sufficiently reach those most at risk, whose engagement on social media is lower and more expensive. This emphasizes the importance of accounting for algorithmic targeting in social marketing. Cost effectiveness analysis based on administrative data indicates that the costs per avoided (officially reported) malaria case range between $3.4 in urban areas and $6.5 overall. Given that the estimated number of cases is roughly twenty-two times greater than the amount of reported cases (WHO, 2021), it is plausible that the cost per averted estimated case could fall below $0.20. Altogether, our results suggest that social media campaigns with public health objectives are cost effective tools, yet they should employ more advanced targeting strategies to ensure ads reach not only those who are the cheapest to advertise to or the most likely to engage with the ad, but also those who are most in need of the campaign.

The rest of this paper is organized as follows: Section 2 describes the contribution to the literature and Section 3 discusses the intervention and its theoretical framework. Section 4 outlines the study design. Section 5 presents the study findings, while Section 6 disentangles the mechanisms of content effectiveness and targeting. Section 7 presents the cost-effectiveness analysis and Section 8 concludes.

2 Contribution to the Literature

Over the last decade, a growing body of the marketing literature has studied the effectiveness of display advertising, finding positive impacts on both brand and performance marketing metrics (Bart et al. 2014; Goldfarb and Tucker 2011*b*; Johnson et al. 2017; Lewis and Reiley 2014), and highlighting its inefficiencies and measurement challenges (Gordon et al. 2021; Johnson 2023). Our work builds and expands upon this literature providing new evidence on the impact of ads on societal outcomes beyond traditional marketing outcomes, and describing a new methodology to measure ad effectiveness on off-platform behaviors.

In this respect, the paper is close to recent studies assessing the effects of advertising campaigns with public health objectives, in both developed and developing countries. Digital and social media campaigns have been shown to be effective in reducing alcohol and tobacco consumption (Thompson et al. 2013; Wang et al. 2021), changing attitudes and behavior towards COVID-19 vaccination and routine immunization (Athey et al. 2023; Evans et al. 2023), promoting better physical and eating

[5] The ads show up on their Facebook timeline from the MNM page, like any ad, and respondents do not have any way of knowing that any particular ad came from the research team.

habits (Carins and Rundle-Thiele 2014) and reducing social stigma in different areas (for a complete review, see Shawky et al. 2019).[6] However, the effectiveness of such campaigns has been, to our knowledge, entirely restricted to assessing on-platform engagement metrics or short-term "Brand Lift" surveys,[7] which may not be informative about medium-term behavioral change and public health outcomes.

We contribute to the existing literature by measuring the effects of a social media campaign on malaria-related behaviors and incidence, using both self-reported surveys and objective health facility data for up to five months post-advertisement exposure. Our study provides new and substantial evidence that social media can be an important channel in fighting malaria and similar diseases through the promotion of preventative behaviors and response actions. In this respect, our research complements evaluations of traditional interventions conducted in the field to combat malaria, including initiatives to subsidize and distribute antimalarials, rapid tests, and bed nets. (Cohen et al. 2015; Tarozzi et al. 2014). We show that digital ads that remind individuals to use bed nets and timely seek treatment in case of symptoms can be a cost-effective way of reducing malaria incidence. Moreover, in line with previous studies (Lambrecht and Tucker 2019; Sapiezynski et al. 2022), our analysis of the mechanisms sheds light on the importance of accounting for algorithmic targeting when delivering ads with public health objectives, as populations that are more at risk of the disease are also less likely to engage with the ads.

In this respect, our study demonstrates the importance of data in targeting users and tracking them to measure ad effectiveness, contributing to the ongoing debate on privacy and the use of third-party data in digital marketing (Acquisti et al. 2016; Tucker et al. 2018). The constraints imposed by privacy regulations, such as the GDPR, CCPA, and recent changes by Apple and Google, pose significant challenges (Goldfarb and Tucker 2011*b*; Johnson et al. 2020; Wernerfelt et al. 2022).[8] However, they also underscore the need for innovative approaches that respect consumer privacy while leveraging data for effective targeting and measurement (Godinho de Matos and Adjerid 2022; Goldfarb and Tucker 2011*a*; Rafieian and Yoganarasimhan 2021). By integrating privacy considerations into our experimental methodology, our approach highlights how first-party data, directly collected from users on social media, can inform targeting and help reach the most valuable segments of the population—in our context, those most at risk of malaria.

More broadly, our study relates to the work emphasizing the importance of marketing techniques to address social issues (Chandy et al. 2021; Kotler and Levy 1969; Kotler and Zaltman 1971), with

[6] Traditional media (television/radio/print) campaigns have also been shown effective for public health outcomes such as the prevention of childhood obesity and COVID-19 (Croker et al. 2012; Ghosh Dastidar et al. 2023. For a complete review of the effects of mass media on health outcomes, see Orozco-Olvera and Malhotra (2023).

[7] Brand Lift Studies are commonly used to measure ad effects on outcomes such as ad recall, brand sentiment, and purchase intent. Companies such as Meta and Google provide Brand Lift experiments to advertisers, typically executed via a brief three-question survey administered to respondents one to two days following exposure to the advertisement.

[8] Regulatory measures such as the General Data Protection Regulation (GDPR) in Europe and the California Consumer Privacy Act (CCPA) in the United States have been enacted to protect consumer privacy and limit the extent to which personal data can be used without explicit consent. Concurrently, tech giants like Apple have restricted user tracking through device IDs on mobile devices, and Google has announced plans to deprecate third-party cookies in its Chrome browser.

a focus on emerging markets (Narasimhan et al., 2015). Examples include strategies to promote consumer literacy (Viswanathan et al. 2021) and environmental consciousness (Brecko and Hartmann 2023; Sun et al. 2021; Zhang et al. 2021), to reshape attitudes towards gender norms and violence against women (Donati et al. 2022), the impact of marketing skills on business performance (Anderson et al. 2018; Bulte et al. 2017), learning about new technologies through social networks (Miller and Mobarak, 2015) and using social norms to promote behavior change (Burchell et al. 2013). Our study has also practical implications for practitioners and public policy (Davis et al. 2021), as it highlights the potential of expanding off-the-shelf social media tools to micro-target social media campaigns and measure impacts beyond short-term engagement metrics and across different sub-populations. We do so by using a newly developed survey chatbot (Rao et al., 2020) that can be integrated into social media platforms to deliver and evaluate online interventions.

3 Background and Intervention

3.1 Malaria

Malaria is a mosquito-borne disease caused by a parasite and transmitted by the bite of infected mosquitoes. In line with global trends of the past two decades, malaria has seen a significant decline in India, where its incidence (per 1,000 population at risk) decreased from 19.9 in 2000 to 3.2 in 2020 (World Bank, 2022). Yet, malaria continues to stand as a significant public health challenge. It is estimated that 93% of the population in India is at risk of the disease (WHO, 2018). The country accounts for 1.2% of global malaria-related deaths and 82% of all malaria deaths in Southeast Asia. In 2020 alone, there were approximately 7,300 deaths attributed to malaria, along with an estimated 4.1 million malaria cases. Of them, approximately 186,000 were officially recorded, equating to one every 22 estimated cases (WHO, 2021). This difference is driven by factors such as asymptomatic cases, under-reporting (especially where health care access is limited), and diagnostic and data challenges. The disease disproportionately affects poor and rural people in the country (Dev et al., 2004). COVID-19 brought additional health and data challenges in 2020. It is estimated that each episode of malaria leads to approximately 11 days of work or school lost (Singh et al., 2019).

3.2 The Intervention

The campaign Bite Ko Mat Lo Lite ("Don't take the bite lightly") is a social marketing campaign designed and implemented by Malaria No More (henceforth MNM), an international nonprofit organization established in 2006 by business leaders seeking to apply private sector expertise to combat malaria. Funding for this national social media ad campaign came from Facebook's Campaigns for a Healthier World initiative, with technical support provided by Upswell, a social media marketing firm specializing in promoting social good campaigns.

The campaign was scheduled to coincide with the monsoon season and the months that follow[9]

[9]Substantial rainfall runs from June to September in diverse regions, with subsequent precipitation patterns differing according to each specific area. The highest incidence of malaria transmission occurs as a result of the

and went live on Facebook and Instagram from July 2020 to early January 2021. With a total ad expenditure of 202,000 USD, the campaign reached approximately 130 million users on these platforms across 22 states known for historically high malaria incidence rates (WHO, 2019).[10] Annex Figure C1 shows the targeted states.

The campaign consisted of a series of ads[11] centered on promoting preventive and response actions, such as consistently sleeping under bed nets, using mosquito repellent/sprays, and promptly seeking testing and treatment for malaria symptoms within 24 hours of the onset of a fever.[12] Given the overlap with the COVID-19 pandemic, the messages emphasized the importance of testing due to the shared symptom of fever between the two diseases.

The creation of the content was user-centered and inspired by audience personas. This involves constructing archetypes or typologies that embody distinct characteristics of actual users, including factors like motivations, interests, and communication preferences (Miaskiewicz and Kozar 2011; Nielsen 2013). For the intervention, six personas were constructed using combinations of gender, age, and location: Young Urban Men, Young Urban Women, Young Rural Men, Young Rural Women, Urban Seniors, and Rural Seniors. Through formative research utilizing data insights from their online behaviors, barriers to preventative and responsive behaviors were identified for each persona. Subsequently, content was tailored per persona based on these findings.[13] Annex A provides a detailed description of the content development process, campaign management and examples of ads targeting each persona.

The objective of the campaign was to maximize engagement with the ads and with the page of MNM, which was managed daily for the duration of the campaign to promote positive interactions. Across all ads, the campaign reached individuals at an average frequency of 3.6 impressions per user (and a maximum of 5), and triggered more than 45 million actions (e.g., clicks) and reactions (e.g., likes and comments). While MNM initially deployed over 50 different ads, not all performed equally well in terms of engagement metrics. As outlined in Annex A, the MNM team opted to deactivate ads with lower engagement and reallocated the budget to the top-performing content, shown in Annex Figure A7. Among the top performing ads, four mentioned the importance of using bed nets for protection, and three stressed the urgency of timely seeking for help/treatment in case of fever.

accumulation of rainwater, which creates favorable conditions for the breeding of mosquitoes.

[10]These states have a total population of 1.2 billion people, representing 86% of the country's population. The campaign reached about 45% of Indian Facebook users in 2020.

[11]The central campaign content was also employed for a broader outreach initiative extending beyond Facebook, encompassing various digital platforms, broadcasts, print media, radio, and social events.

[12]Initial symptoms might appear mild, resembling various fever-related illnesses, which can make recognizing malaria challenging. If left untreated, malaria can rapidly develop into severe illness and death. Individuals at heightened risk include infants, children under 5 years, pregnant women, travelers, and those with HIV/AIDS.

[13]For example, for women, ads focused on maternity and protecting children and other family members. For men, content focused on information and services (government, healthcare), symptoms and treatment. Younger audiences were targeted using meme-styled content and humor; and older individuals received content encouraging them to spread the word and advocate for prevention and response, for example, by encouraging grandparents to have their grandchildren tested if they experienced malaria symptoms. For rural areas, ads in Hindi and Hinglish were designed, as well as ads leveraging relevant health advisory figures such as community health workers.

We conducted an independent impact evaluation using the same content and targeting strategy of this national campaign. In coordination with MNM, we limited our study to three states — Uttar Pradesh, Chhattisgarh, and Jharkhand—, where MNM agreed to postpone the launch of the nationwide campaign until late August 2020. These three states collectively comprise 126 districts and a combined population of over 300 million people. Traditionally, they have contributed to a significant proportion of the reported malaria cases in the country (WHO, 2019). To enhance internal validity, albeit at the cost of external validity, we excluded the state capitals from the study. These capitals were outliers concerning population size, malaria case numbers, and rates of social media penetration. MNM and the evaluation team agreed to utilize randomization for assigning districts in these states to either the treatment or control group. The evaluation team had no additional role in the implementation of the campaign beyond overseeing the randomization process.

3.3 Theoretical Framework

The campaign design was based on an assessment of what would motivate individuals to take preventative and response actions to malaria. As such, the ads aimed not only to create the intention to act, but the actions themselves, and depending on the behavioral phenotype of the different individuals — represented in the personas — ensure that they had sufficient and adequate information, motivation, and opportunity to act. In this section, we outline the behavioral models and theories from which the ads indirectly drew. Subsequently, we elaborate on the precise implementation of these models in relation to the advertisements.

Behavior Change Frameworks Several models help understand how individuals decide to adopt protective actions against risks (see Baldwin et al. 2022 and Zimmerman et al. 2016 for a review). The Health Belief Model (HBM) emphasizes cost-benefit analysis and perceived vulnerability and severity of impact (Rosenstock et al., 1988). The Theory of Planned Behavior (TPB) links beliefs to behaviors through attitude, subjective norms, and perceived behavioral control (Ajzen, 1991).[14] Expanding on TPB, the Protective Action Decision Model (PADM) includes physical and social environments influencing the decision-making process, from pre-decision to behavior implementation (Lindell and Perry, 2012). Protective Motivation Theory (PMT) focuses on threat appraisal (balancing harm probability and non-protective behavior rewards) and coping appraisal (balancing self-efficacy and solution effectiveness) (Babcicky and Seebauer 2019; Botzen et al. 2019).

These models share common concepts: motivation drivers, resources/capabilities (internal and external), and the opportunity to act. The Fogg Behavior Model (FBM) adds that behavior results from motivation, ability, and prompts (Fogg, 2009). Belief in one's ability to implement action and its effectiveness is crucial. Regarding malaria, additional behavioral factors like optimism bias

[14] Perceived behavior refers to the individual's assessment of self-efficacy or control to execute a behavior; subjective norms refer to the social perception of the behavior as a positive one, and the assessment of how many others are engaging in it. Attitude is the individual's perception of the behavior and its outcome. The interaction of these three components generates intentions to engage in protective action (Najafi et al., 2017).

(underestimating personal risk) and the gambler's fallacy (believing a frequently occurring event is less likely to recur) can lead to procrastination or a false sense of safety (Sharot 2011; Tversky and Kahneman 1974).

Theory-based Ads These frameworks were instrumental for the MNM creative team in guiding the design and content of the ads. First, theories helped understand and leverage the motivation to take action, whether that entailed caring for one's family or oneself. Secondly, these models guided the identification of practical preventive and responsive actions that could be widely adopted by significant segments of the general population in India. These actions included measures like sleeping under bed nets or seeking medical attention in the presence of fever. Third, the ads were strategically engineered to present clear and compelling prompts for action and the consequences (cost) of inaction. By emphasizing specific risk-reducing steps, the messages aimed to enhance individuals' self-efficacy, an essential element in fostering a sense of empowerment to undertake protective measures. Lastly, the ads intentionally displayed a limited amount of information. This approach was chosen to alleviate cognitive overload and to address the natural tendency of individuals to focus primarily on a limited set of factors that they perceive as significant (Schwenk, 1984).

Our study evaluates the effectiveness of the entire MNM campaign, which encompassed multiple ads. Our evaluation did not probe individual mechanisms underlying each ad.

4 Study Design

This section introduces our clustered design as well as our research methodology using Virtual Lab, an open-source platform (Rao et al., 2020), to recruit respondents through Facebook advertising with mobile credit incentives. Additionally, we discuss how formative research helped identify dwelling type as a key predictor of malaria risk, guiding recruitment and stratification strategies to study potential treatment effects on policy-relevant populations.

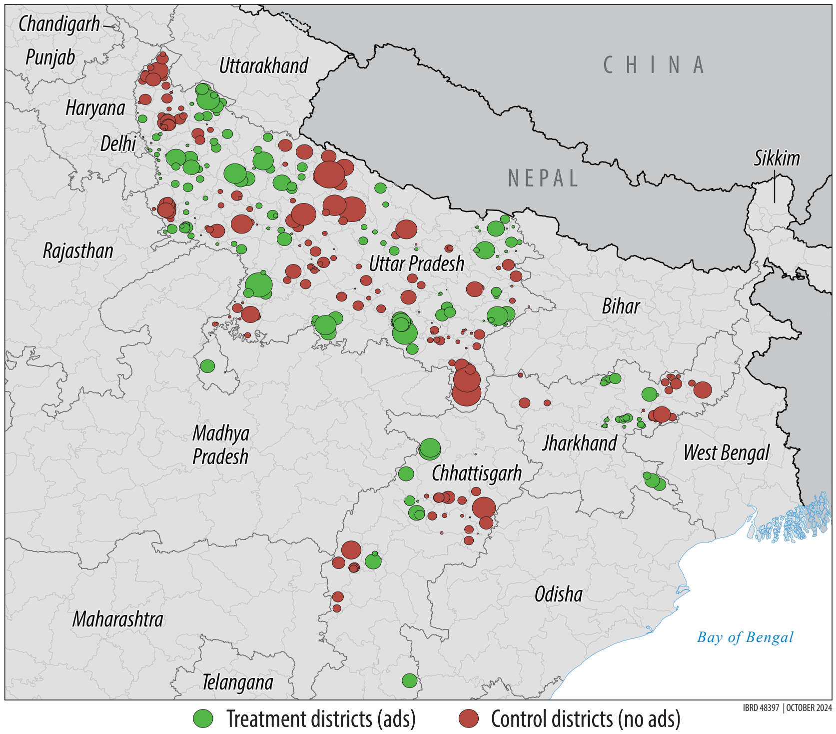
4.1 Cluster-randomized Controlled Trial

We use a cluster randomized controlled trial, where guided by power calculations, we randomly assigned 40 of 80 districts to the treatment condition (ads) and the remaining 40 to the control condition (no ads), balancing on several district-level characteristics. These 80 study districts were selected based on convenience, favoring areas where recruiting study participants on Facebook was the most cost-effective.[15]

To work around Facebook's restriction on targeting ads by district boundaries, we identified the principal cities within each district and employed GIS software to delineate circles around them. These circles expanded until they reached the district's borders. Figure 1 shows the distribution

[15] The cost of recruiting on Facebook is influenced by several factors but primarily higher costs are linked to a lack of an engaged population. Thus, a limitation of our study's participant pool is that it doesn't encompass the entirety of the state; rather, it excludes some of the less connected and less densely populated areas as well as the state capitals.

Figure 1: Cluster-randomization of districts (and circles therein)



WORLD BANK GROUP This map was produced by the Cartography Unit of the World Bank Group. The boundaries, colors, denominations and any other information shown on this map do not imply, on the part of the World Bank Group, any judgment on the legal status of any territory, or any endorsement or acceptance of such boundaries.

of treatment and control districts (and circles therein) over space. While one district can contain multiple circles, all circles in a district are in the same experimental condition. Circles are identified by the coordinates of their centroids and a radius. We passed this information to the MNM advertising team, which used it to exclude the control (red) circles from their nationwide campaign, while running it on the treatment (green) circles as well as on all the remaining parts of the districts and the other targeted states. Due to an implementation issue, we dropped one district from the analysis, leaving 79 study districts in total.[16]

We use a cluster design for the following two reasons (Hayes and Moulton, 2017). First, we acknowledge the presence of different types of spillovers.[17] A cluster design allows us to account for the correlation across observations within a cluster and to minimize the risk of contamination between treated and control areas. Second, related to the spillovers of the infectious disease itself, we are interested not just in the effects of the ad campaign on the direct reduction of one individual's

[16]The randomization list given to the MNM advertising team included two districts named Balrampur, each in a different state. This duplication made it impossible to determine whether one of these districts belonged to the treatment or control group. Consequently, we decided to exclude it from the analysis.

[17]These may (i) occur within families, where treated users may help their family members to adopt precautions against malaria exposure; (ii) ripple across the social graph as Facebook ads and posts are shared by treated users, thereby amplifying exposure to their Facebook friends; and (iii) include health-related spillovers connected to the infectious disease itself, wherein the decrease in malaria incidence among treated individuals and their families could contribute to a reduction in malaria cases among their neighbors.

risk, but on the population-level effect of mass behavior change on the risk of an average member from the population. A cluster design allows us to estimate such population-level effects.

4.2 Stratified Recruitment

We used Facebook advertising to recruit survey respondents from the previously mentioned circles in both treatment and control districts, incentivizing participation with mobile credit ranging from 2.6 to 3.6 USD. Respondents were asked to complete a series of surveys conducted via chatbot on Facebook Messenger and delivered by a page that we created called "Global Insights".[18] The main benefit of employing chatbot surveys on Messenger lies in their seamless integration into the ad platform itself. Moreover, the chatbot streamlines follow-up with respondents, as inviting them to additional surveys can be as simple as sending a new chat message.

Before conducting the study, we carried out a formative research phase to guide our sampling design. Specifically, in July 2020, we started recruiting respondents asking them several demographic and socio-economic questions, as well as whether they or anyone in their household had had malaria in the past 5 years and in the previous two weeks. These surveys indicated that one of the original study states, Odisha, exhibited low malaria rates among our target population. Consequently, we opted to exclude it from our trial due to statistical power constraints.

Importantly, the analysis of our formative research data in the other three study states revealed that dwelling type was the most robust household-level predictor of past and current malaria incidence.[19] Its predictive capacity, both in terms of Pearson pairwise correlation and OLS regression coefficients, surpassed that of other factors such as proximity to health facilities, membership to a backward caste, family size, religion, and the educational background of the respondent (Annex Tables D1 and D2). Specifically, past and current malaria incidence rates in households living in solid dwellings (i.e., those with walls and roof made of bricks, stones and cement) were 21% and 43% lower compared to those living in non-solid and semi-solid houses (i.e., made of mud, tin and straw), respectively.[20] For our analysis, we combined the latter two dwelling categories.

Understanding the most important risk factors enables us to estimate conditional treatment effects for the most vulnerable populations, addressing the research question of *whether social media campaigns reach and impact these high-risk groups*. In our initial recruitment, however, people in non-solid dwellings were significantly less likely to be reached and to respond to our survey than those in other dwelling types. In fact, the vast majority of respondents lived in solid dwellings. For this reason, we decided to stratify our recruitment by district and dwelling type (160 strata in total, 80 for the districts and 2 for the dwellings). Besides being a strong predictor of malaria risk, dwelling type captures the dimensions of poverty and rurality (this variable is correlated with education,

[18] Annex Figure C2 shows examples of ads used to recruit survey participants.

[19] Following the India DHS survey, respondents were asked which kind of house they lived in among three options: non-solid houses (made of mud, tin, straw); solid houses (built with durable materials like stones, bricks, cement, or concrete, encompassing the entirety of the structure, including the floor, roof, and outer walls); semi-solid houses, which occupy an intermediate position between the previous two housing types.

[20] This is in line with previous work in India (Yadav et al., 2014) and elsewhere (Wanzirah et al., 2015).

Table 1: Baseline cluster characteristics by experimental condition

	Treatment *Mean* (1)	Control *Mean* (2)	Normalized Difference (3)	*T-test* *(p-value)* (4)
Cluster covariates				
N. of clusters	39	40		
Population (Million)	2.756	2.629	0.077	0.629
Share living in solid houses	0.502	0.494	0.078	0.623
Share with university degree	0.598	0.593	0.063	0.693
Share unemployed	0.439	0.446	-0.103	0.519
Share with malaria in last 2 weeks	0.017	0.016	0.072	0.652
Share with malaria in last 5 years	0.190	0.184	0.051	0.750
Average click-through rate (%)	1.830	1.847	-0.048	0.764
Average cost per mille (USD)	0.764	0.758	0.050	0.756
Average ad cost per survey started (USD)	0.252	0.241	0.072	0.650
Average ad cost per survey completed (USD)	0.820	0.858	-0.052	0.744

Notes: The table reports clusters' characteristics measured at baseline, by experimental condition. Population numbers come from census data. The normalized difference is the difference between the two means divided by the square root of the sum of the variances of the two variables.

distance from the nearest medical center and caste (Column 3 of Annex Table D2) and provides a convenient mapping into administrative data on malaria incidence, which are disaggregated across urban and rural locations. In India, non-solid and semi-solid dwellings make up approximately 15% in urban areas and around 50% in rural areas (India DHS).[21]

To overcome the challenge that Facebook does not offer the option to target ads by dwelling type, we created Lookalike Audiences in Facebook and dynamically increased budget to them in districts where non-solid dwellers were underrepresented.[22] The Lookalike Audience tool allows advertisers to connect with fresh users who mirror their current customers or desired target audience. Drawing from our formative research surveys, we provided Facebook with the "seed audience" of individuals that lived in non-solid dwellings. Leveraging its rich data and algorithms, Facebook then assembles a Lookalike Audience by extrapolating patterns identified within the seed audience.

We used this procedure to ensure that in each district we recruited a sufficient sample size for measuring the effects of the campaign on those most at risk of the disease. The benefits of this stratification can be seen in Annex Figure C3, which shows an improvement in the share of non-solid dwellers per district and a reduction in its variance across districts as a result of the optimization. Table 1 shows that, eventually, the stratified recruitment and cluster-randomization resulted in a sample with two important features: (1) baseline cluster characteristics are balanced across treatment and control districts, a condition for internal validity; (2) the average share of cluster population living in solid houses is around 50%, which allows the estimation of conditional treatment effects for both types of dwellings. The table also provides insights into the main metrics of our recruitment campaign, pointing out an average ad cost of 85 cents per final survey respondent.

[21] According to the India DHS, the distribution of housing types in urban areas is as follows: 85.8% solid, 13.3% semi-solid, and 0.9% non-solid houses. In rural areas, these percentages change to 48.7% solid, 44.8% semi-solid, and 6.5% non-solid houses, respectively.

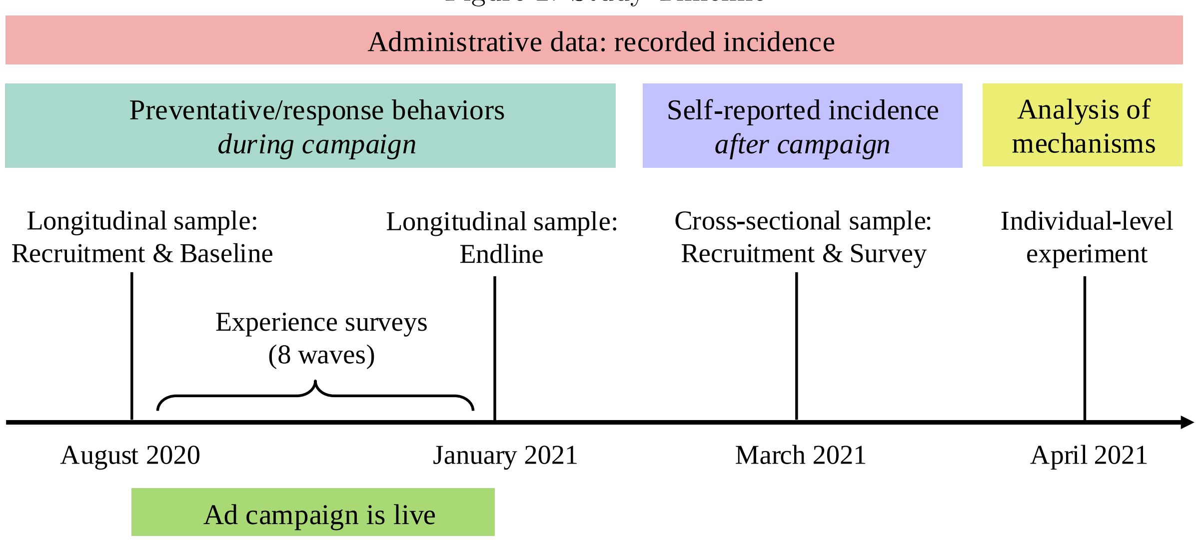
[22] By leveraging real-time response data, Virtual Lab optimized the budget allocation across ad sets to get all the strata filled as quickly as possible.

4.3 Timeline and Data

4.3.1 Study Timeline

Figure 2 shows the study's timeline and data collection phases. In the three study states, the MNM campaign was live for almost 5 months, from August 19th 2020 to January 6th 2021. To assess its effectiveness, we recruited two separate survey samples over distinct periods: (i) a longitudinal (experience) survey from early September 2020 to end of January 2021 to gauge effects *during* the campaign; and (ii) a cross-sectional sample from January to late March 2021 to study accumulated and medium-term effects *after* the campaign had ended. As outlined in Section 6, we also carried out an additional survey wave in April 2021 as part of an individual-level *feed* experiment, aimed at distinguishing between limited effectiveness and limited reach of the ad campaign for vulnerable populations. Finally, we augmented these survey samples with administrative health facility records covering the period from four months prior to the campaign's initiation (April 2020) to five months after its completion (May 2021).

Figure 2: Study Timeline



4.3.2 Survey Samples and Administrative Data

During the campaign, the longitudinal (experience) survey assessed malaria-related preventive behaviors (such as bed net usage) and response actions (such as seeking medical care when experiencing fever symptoms). This survey involved a baseline questionnaire (before or at the start of the campaign), an endline survey (towards the end of the campaign and right after it), and a series of 8 short surveys in between, conducted every 17 days. All baseline respondents regardless of how many intermediate waves they completed, were invited to complete the endline.[23] To analyze potential impacts during the campaign, we filtered the data to include responses collected from September

[23]Hence, early enrollees had the potential to provide responses for all 8 experience surveys and the endline survey

1st to January 31st and aggregated them at the individual level, across different waves. The final experience sample contains information from 5,205 individuals, equally split between treatment and control districts.

The cross-sectional sample measured the accumulated and medium-terms impact of the campaign on malaria incidence, risk perceptions, and bed net-related actions and intentions. Some questions from the longitudinal survey were reiterated, with variations in the time periods for respondents to report on. For example, while the experience surveys inquired whether participants or their family members had experienced malaria in the past two weeks, the cross-sectional survey posed a similar question but with reference to the last 6-8 months, i.e. since August 2020, when the study campaign went live. We also asked about malaria risk perceptions, and intentions to buy or treat bed nets with insecticide. The cross-sectional sample comprises 3,052 observations. As discussed in Section 6, in the cross-sectional study we also nested an individual-level study to explore the mechanisms behind the lack of program impacts on households living in non-solid dwellings.

For both samples, recruitment was stratified across districts and dwelling type, as described previously. The enrollment questionnaire started asking participants about their preferred language (English, Hindi, or Odia), followed by providing study details to obtain individuals' informed consent for their participation in the research. Only individuals aged 18 or above and whose consent was obtained are part of our survey samples. Prior to analysis, we standardized and refined the samples to enhance data integrity. Specifically, we removed respondents who hastily completed the survey[24] and those reporting to be above 90 years old and indicating a household size exceeding 30 members, retaining only observations where all outcome and control variables were present.

We complemented survey data with administrative health facility data collected at the sub-district-month level from the Indian Health Management Information System (HMIS).[25] This dataset provides monthly information from health facilities at the sub-district level, detailing the number of malaria tests performed and the positive cases identified. It separates the data for urban and rural health facilities, which is used for analyzing treatment effects specific to each setting. The dataset spans from April 2020 to May 2021.[26] We designated April to August 2020 as the pre-treatment period and September 2020 onwards as the treatment period, aligning with the dissemination of the campaign in the study states. The use of administrative records complements and expands upon the analysis of survey data in two important aspects. (1) The use of objective malaria indicators alleviates the concern that our results could be driven by social desirability bias and experimenter demand effect, as these might be present in survey responses but not in health facility records. (2)

[24]By employing time data collected for each survey question by Virtual Lab, we omitted observations where the 90th percentile of response time was less than 3 times the fastest response time. This procedure accommodates variations in respondents' capabilities and response velocities.

[25]HMIS is a monitoring system that has been put in place by the Indian Ministry of Health & Family Welfare (MoHFW). It collects facility-wise monthly information on a variety of health outcomes, and processes and reports the information at Sub-district, District, State and National levels. The records in this dataset are considered to meet the required standards for completeness for health research in India (Sharma et al., 2016).

[26]Although data from previous periods is available, the total number of sub-districts per state and their boundaries were different and matching across years was imprecise. Hence, we excluded previous observations and focused on an homogeneous set.

Administrative data offer a representative measure of disease incidence in each district, enabling the estimation of the broader impacts of the campaign, including indirect or spillover effects on individuals not directly exposed to it. As a result, this district level analysis also facilitates more precise evaluations of the campaign's cost-effectiveness.

4.3.3 Outcome Measures

We evaluate the impact of the social media campaign on three outcome categories: (1) behaviors associated with malaria prevention and response as reported in surveys, (2) malaria incidence through self-reported surveys, and (3) objective malaria incidence from the health facility HMIS data. Below we list the primary behavioral and health outcomes and explain their construction. The questions underlying the outcomes are reported in Appendix B.

Self-reported outcomes *during* the campaign. Respondents in the experience survey were asked every 17 days a series of questions that captured the adoption of preventative behaviors and response actions related to malaria. From these questions, we constructed the following outcomes: (i) the share of times a respondent reported to use a bed net or other protection measures, (ii) the average share of household members who used a bed net, (iii) any fever/malaria case in the household, and (iv) medical seeking and testing behaviors. We aggregated the responses for every individual across survey waves by considering the *mean* for outcomes (i-ii) and the *maximum* for outcomes (iii-iv). In the analysis, we weight observations by the number of survey waves answered by the respondent.

Self-reported outcomes *after* the campaign. In the cross-sectional survey, we inquired about the occurrence and number of individuals who tested positive for malaria since August 2020 (the campaign launch date), households' concerns about contracting COVID-19 or malaria in the upcoming year, intentions to purchase mosquito nets, and whether mosquito nets had been treated with insecticide.

Health facility recorded outcomes. These data were used to construct objective measures of malaria incidence. Specifically, for every subdistrict, we considered the monthly number of recorded malaria cases and tests conducted, as well as the total population counts.[27] From these, we constructed malaria incidence rates (number of recorded cases per million people), and test positivity rates (number of positive cases per test conducted).

[27]We collected population data from WorldPop (www.worldpop.org), which produces various gridded population count datasets. We used the 2020 1km-resolution data, which were adjusted to match United Nations national population estimates. We used GIS software to compute the total population in each subdistrict.

4.4 Empirical Specifications

4.4.1 Analysis of Survey Data

In the survey analysis, each observation is an individual. Respondents in districts where ads were shown are considered "treated" and those in districts where ads were not shown are considered the "control" group. Because survey respondents in treatment districts were not necessarily exposed to the MNM campaign, our results estimate intent-to-treat (ITT) effects. Our workhorse model, estimated via Ordinary Least Squares, takes the following form:

$$y_{idz} = \alpha + \beta\, T_{dz} + \gamma\, \mathbf{x}_{idz} + \delta\, \mathbf{w}_{dz} + \theta_z + \epsilon_{idz} \tag{1}$$

where i is an individual residing in district d of state z. T is the treatment indicator, which is randomized across districts. β is our coefficient of interest, which captures the effect of being assigned to the treatment on the outcome Y. Note that when the outcome is binary, the equation describes a Linear Probability Model. $\mathbf{x}$ and $\mathbf{w}$ represent, respectively, vectors of individual-level and district-level control covariates measured at baseline and potentially correlated with our outcomes. Since the treatment was randomly assigned, incorporating these control variables should not influence our estimates of β. However, their inclusion enhances the efficiency of the estimator. Lastly, θ denotes state fixed-effects, which account for unobserved state-level factors that could impact our outcomes of interest, given the potential influence of such factors in the Indian context.[28] We cluster the standard errors at the district level to account for intra-cluster correlations in the error term ϵ. We also investigate the robustness of our results to different model specifications (e.g., adding controls and interaction terms) and estimators. For the latter, we estimate a logistic regression for all binary outcomes.

4.4.2 Analysis of Administrative Data

Administrative data were gathered at the subdistrict level, covering a total of 395 subdistricts across the 79 districts included in the analysis. Focusing on smaller areas rather than the entire district reduces measurement error and increases the precision of our estimates. Moreover, the subdistrict-level analysis also alleviates concerns about health facility records in control districts potentially including outcomes for individuals exposed to the campaign. The smaller scale of subdistricts allows for a more precise matching with our targeted areas.[29] In practice, we use GIS tools to identify

[28] State governments in India play a significant role in the administration of health services, including hospitals, public health programs, and healthcare infrastructure within their territories.

[29] Recall that the implementation of our cluster-level design did not rely on the boundaries of the districts but on a series of circles around their major cities. These circles were used in the control group to identify areas within a district to be excluded from the campaign (see Section 4.1). This implies that some parts of the districts in the control group received ads. While this is not a problem for the survey-based analysis -in fact, respondents were precisely sampled within the circles, it creates potential concerns when analyzing administrative records. In the control group, district-level data would reflect records from individuals who were potentially exposed to the campaign, which would confound and downward-bias our estimates. Focusing on subdistricts alleviates this concern, as their smaller dimension allows us to consider those that better coincide with our circles.

the intersecting areas between each subdistrict and our circles. We then compute the share of the subdistrict population that lives in these areas and use it to define different subsamples. The higher the share, the lower the potential contamination in the control group.

In the benchmark estimation, we focus on a subsample of 119 subdistricts with at least 50% of their population living within our pre-defined circles. We will also show the sensitivity of our estimates to different choices of the population threshold (e.g., between 10% and 50%).[30] Annex Table D7 shows the mean outcomes for the entire number of subdistricts and for the subgroup used in the benchmark analysis. While absolute incidence (number of malaria cases) is similar across samples, its rate per million people is significantly lower in our subsample compared to the entire sample (about 21 vs. 51 cases per million). The difference reflects the larger average population in the subsample, which is due to the fact that our circles included the major towns and cities of the subdistrict. The table also reports the means disaggregated across urban and rural health facilities.[31] The numbers highlight the severity of the disease in rural areas. For example, in our subsample, absolute and relative incidences recorded by rural health facilities (9 and 19 cases, respectively) are twice as large as those in urban areas (4.8 and 9.5, respectively). To improve data integrity, we excluded from the analysis subdistricts-months where the number of positive malaria cases exceeded the number of malaria tests conducted.

To assess the impact of the ad campaign, we exploit the within-subdistrict variation between April 2020 and May 2021 and estimate the following Difference-in-Differences model:

$$y_{sdt} = \alpha + \beta\, T_d \times \text{Post}_t + \gamma\, \mathbf{x}_{sd} \times \text{Post}_t + \delta\, \mathbf{w}_d \times \text{Post}_t + \mu_s + \psi_t + \epsilon_{sdt} \tag{2}$$

where y is the outcome of subdistrict s in district d in month t. T takes value 1 for the subdistricts that belong to a treated district. *Post* is equal to 1 after August 2020, i.e., after the start of the campaign. μ_s represents subdistrict fixed-effects, which account for time-invariant subdistrict-level factors that may be correlated with the outcomes, while ψ_s includes month and year fixed-effects to control for seasonality in the outcomes. We include state fixed-effects interacted with *Post* in all the regressions. Depending on the specification, we also include time-invariant pre-treatment district- and subdistrict-level controls ($\mathbf{w}$ and $\mathbf{x}$, respectively) interacted with *Post* to account for potentially confounding factors that may differentially affect the outcomes over time. We cluster the standard errors at the district level to account for both serial and cross-sectional correlations in the errors. β is the coefficient of interest, which captures the causal impact of the campaign on malaria incidence under the identifying assumption of parallel trends. This requires that in the absence of the treatment, the *changes* in malaria incidence in treatment and control subdistricts would have been the same. The random assignment of the ad campaign should ensure this assumption is valid, but we will present evidence to support it.

[30] A sufficiently large sample is needed to guarantee power for statistical inference. Considering thresholds above 50% would lead to an insufficient number of observations.

[31] Note that not all the subdistricts contain both types of health facilities. Among the 119 subdistricts in our subsample, 116 of them have rural health facilities, while 65 have urban ones.

5 Results

5.1 Samples Characteristics and Randomization Checks

Tables D3 and D4 in the Appendix describe the characteristics of the individuals and households in the experience and cross-sectional samples, respectively. The tables report the variables' means in the treatment and control conditions, their normalized difference[32] and the *p-value* associated to the t-test for their comparison. For the experience sample, we report both covariates (Panel A) and main outcomes (Panel B) measured at baseline (i.e., during the first survey wave). For the cross-sectional sample, conducted after campaign completion, we only report time-invariant covariates.

Both tables exhibit a general balance in baseline individuals' and households' characteristics and pre-treatment outcomes across experimental groups, validating the randomization process. For the longitudinal (experience) sample, none of the differences between treatment and control groups is significant. In the cross-sectional sample, the treatment group happens to include a slightly higher proportion of unemployed individuals and households with pregnant women compared to the control group. However, these disparities are minor, and we address them in the analysis by incorporating these and other factors as controls in the regressions.

In both samples, respondents are roughly evenly split among treatment and control districts, and approximately 55% of them reside in non/semi-solid dwellings, reflecting a successful implementation of the stratified recruitment across districts and dwelling type. Attrition in the experience sample was equally balanced across experimental groups, with respondents completing 3 survey waves, on average. In terms of socio-demographics, males make up about 90% of the samples, with an average age of 27 years, and 55% to 60% report university completion, while about a half indicate being unemployed. Around 60% of the respondents identify with socially and economically disadvantaged castes (OBC, SC/Dalit and ST), which prior studies suggest exhibit a lower awareness regarding malaria prevention (Singh et al., 2020). Household size ranges between 6 and 7 members. One of every eight households is home to a pregnant woman. Three-quarters of participants reside within 60 minutes (by walk) of a medical center, and a similar proportion has a mosquito net at home.

In terms of outcomes, 2% of respondents report a household member having experienced malaria in the past two weeks (and 20% within the last five years). About 70% of those with a bed net at home report having used it the previous night, with about 70% of household members sleeping under it. Combining these numbers with the average bed net ownership in the sample (75%) suggests that just over 50% of individuals are regularly protected from mosquito bites at night. Relatedly, approximately one-third of the participants express some level of concern about getting malaria in the following month.

As expected, our online sample exhibits a higher socioeconomic status compared to the general population. This is highlighted by population-level statistics from the 2019-2021 India Demographic

[32]The normalized difference is the difference between the two means divided by the square root of the sum of the variances of the two variables.

and Health Survey (India IIPS, 2022), indicating that only around a third of households in India possess at least one mosquito net, with an estimated 8% having an insecticide-treated mosquito net (ITN). In the three study states, population-level estimates suggest that only one in three women and six in ten men have accessed the internet at some point. These statistics emphasize that our study and evaluated intervention reflect the results from a specific subgroup of the population, and do not generalize to individuals and households without access to the internet.

As anticipated, both covariates and pre-treatment outcomes vary among different types of dwellings. Annex Tables D5 and D6 demonstrate that households residing in solid dwellings exhibit greater economic advantages, such as higher levels of education, lower unemployment rates, and belonging to higher castes. Additionally, they tend to reside in closer proximity to urban centers with medical facilities and are less prone to experiencing malaria in the last 5 years and 2 weeks compared to those living in non/semi-solid dwellings. Although bed net ownership rates are comparable between the two groups, residents in solid dwellings are less inclined to use bed nets consistently. This could be attributed to their perception of lower malaria contraction risks. These differences emphasize the importance of having both types of dwellers equally represented in our sample, which allows us to explore potential effect heterogeneity.

5.2 Survey-based Results

5.2.1 Preventative behaviors and response actions during the campaign

Table 2 reports the results on sleeping under mosquito nets during the campaign, one of the main messages of the ads that received more budget. Columns (1-4) pertain to the respondent's behavior, while columns (5-8) capture the behavior of the entire household. When considering the complete sample (of respondents that had at least one bed net), we note modest yet favorable impacts of the campaign on increased bed net utilization in columns (1-2) and (5-6). Individually, there is an approximate increase of 3 percentage points (p.p.) in usage (5% relative to control), and for the entire households, there is an increase of 2.8 p.p. (4% relative to control). In columns (3) and (7) we report heterogeneous treatment effects across dwelling type, which reveal that the previous results are solely driven by solid dwellers. Specifically, in solid houses (columns 4 and 8), our estimates indicate that the campaign improved the frequency of bed net use at the individual level by 5.5 p.p. (or 9% relative to control), and increased the share of family members using bed nets by almost 7 p.p. (or 11%). Note that these effects are unlikely to be driven by differential bed net ownership rates across dwelling type, as only households with at least one bed net are considered in the analysis. Moreover, Annex Tables D5 and D6 show a balance across dwelling type in terms of bed net ownership rates and number of bed nets per family member (i.e., approximately one for every two individuals).[33]

Figure 3 reports the heterogeneous treatment effects of the campaign on bed net use throughout the campaign period (left panel) and across baseline risk perceptions levels (right panel). The

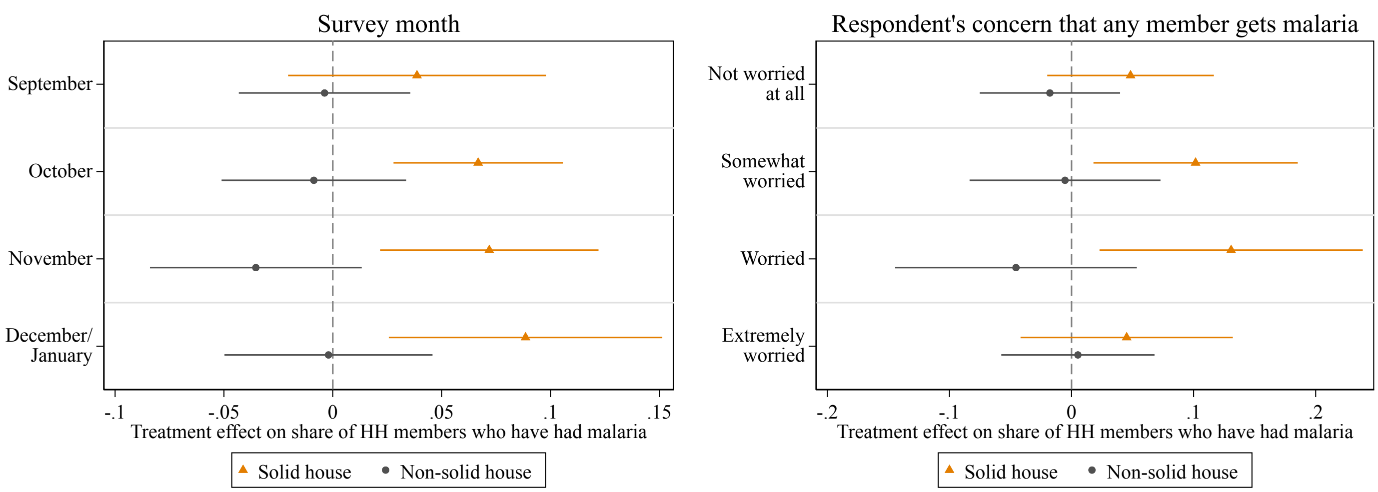
[33]The price of a basic bed mosquito net in India ranges from \$1.5 to \$5, making them affordable for a significant portion of the population.

Table 2: Sleeping under bed net (during the campaign)

	Share of times respondent used bednet				Share of HH members who used bednet			
	(1) Full	(2) Full	(3) Full	(4) Solid house	(5) Full	(6) Full	(7) Full	(8) Solid house
T	0.030	0.031*	0.007	0.055**	0.025	0.028**	-0.010	0.068***
	(0.019)	(0.017)	(0.024)	(0.025)	(0.017)	(0.014)	(0.018)	(0.019)
T × Solid house			0.051				0.081***	
			(0.037)				(0.025)	
Solid house		-0.017	-0.041			-0.039***	-0.078***	
		(0.020)	(0.028)			(0.013)	(0.019)	
Individual controls		✓	✓	✓		✓	✓	✓
Cluster controls		✓	✓	✓		✓	✓	✓
Observations	3900	3900	3900	1789	3900	3900	3900	1789
Adj. R-squared	0.006	0.016	0.017	0.031	0.005	0.033	0.036	0.039
Mean in control	0.633	0.633	0.633	0.603	0.661	0.661	0.661	0.610
SD in control	0.422	0.422	0.422	0.430	0.344	0.344	0.344	0.361

Notes: Each observation is an individual. Responses collected from September 1, 2020 to January 31, 2021 are included and are weighted by the number of survey waves the individual completed in this period. Only households with at least one bednet are included. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of an air conditioner and 5-year malaria incidence at baseline. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 3: Heterogeneous treatment effects by survey month and malaria concern levels



Notes: The figure reports coefficient estimates from the regression described in columns (5-6) of Table 2, estimated on different sub-samples, based on the month in which the survey was taken (left panel) and the levels of malaria concern measured at baseline (right panel), across dwelling type. Non-solid houses include semi-solid dwellings. Only households with at least one bed net are included. Bars represent the 95% confidence intervals. Standard errors are clustered at the district level.

former panel shows that, only among solid dwellers, bed net adoption intensified over the period the campaign was live, from an insignificant 4 p.p. effect in September to nearly a 10 p.p. impact towards the end of the campaign. This temporal pattern indicates that as the campaign reach and frequency grew, its effectiveness increased. This suggests that there is likely a minimum threshold of reach and frequency necessary for ads to generate any discernible behavioral changes. The right panel shows that the campaign successfully increased the use of mosquito nets among households that initially showed moderate levels of concern. Yet, this is visible only among households living in solid dwellings. These results suggest that the ads did not convince individuals who were unconcerned about malaria risks. Instead, they served as a reminder for those with moderate levels of awareness and concerns, prompting them to take action. At the same time, the campaign did not further change the behaviors of those who were already extremely worried about the disease. Both panels confirm the null effects for households living in non/semi-solid dwellings.

We also investigated the effects of the MNM campaign on other protective behaviors, such as wearing long-sleeve clothing and using body and wall sprays to repel mosquitoes and prevent bites. Annex Table D8 shows absence of impact on these outcomes, both overall and across dwelling type. A plausible explanation is the limited coverage that these protective behaviors received in the campaign. In fact, although the initial iteration of content discussed these actions, the ads received lower engagement compared to those on bed nets and treatment seeking. Thus, the MNM team allocated limited budget to them when the campaign was scaled-up in September-December 2021 (see Section 3.2).

Table 3: Incidence of high fever and timely help-seeking (during the campaign)

	Any member has had a fever				Any member sought for help within 24h			
	(1) Full	(2) Full	(3) Full	(4) Solid house	(5) Full	(6) Full	(7) Full	(8) Solid house
T	0.007	0.008	0.016	0.008	0.006	0.019	-0.023	0.107***
	(0.017)	(0.014)	(0.018)	(0.020)	(0.030)	(0.026)	(0.040)	(0.035)
T × Solid house			-0.016				0.107*	
			(0.028)				(0.058)	
Solid house		-0.017	-0.009			0.051	-0.004	
		(0.016)	(0.020)			(0.036)	(0.046)	
Individual controls		✓	✓	✓		✓	✓	✓
Cluster controls		✓	✓	✓		✓	✓	✓
Observations	5158	5158	5158	2381	724	724	724	301
Adj. R-squared	0.002	0.033	0.033	0.037	-0.000	0.034	0.038	0.021
Mean in control	0.156	0.156	0.156	0.145	0.797	0.797	0.797	0.808
SD in control	0.363	0.363	0.363	0.352	0.403	0.403	0.403	0.395

Each observation is an individual. Responses collected from September 1, 2020 to January 31, 2021 are included and are weighted by the number of survey waves the individual completed in this period. In columns (1-4), the dependent variable takes value 1 if the respondent/family member has had a fever (100.4°F / 38°C or above) at least once in the period under consideration. In columns (5-8), the dependent variable takes value 1 if the respondent/family member with fever has sought for help within 24 hours at least once in the period under consideration. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and air conditioner, and 5-year malaria incidence at baseline. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

As for response actions, participants in the experience sample were surveyed every two weeks regarding any occurrences of a fever (at or above 100.4°F / 38°C) within their family. For those who reported a fever, we recorded whether they sought medical assistance within 24 hours. Table 3 reports impacts on these events and behaviors. We do not find evidence that the treatment affected incidence of fever during the campaign (columns 1-4). Yet, we find that the probability to seek for treatment in a timely manner (i.e., within 24 hours) in case of a fever increased by more than

13% among those living in solid houses (column 8), compared to the control. Again, we do not find impacts on respondents living in non/semi-solid dwellings (column 7). We also explicitly asked about malaria cases and testing behaviors. However, Annex Table D9 shows no significant impact of the ads on these outcomes. The general lack of effect on fever and malaria incidence throughout the campaign period can be attributed to the necessity of maintaining protective actions (such as consistently using bed nets) over an extended duration for them to yield health benefits. This justifies our decision to enlist a fresh cohort of respondents and gather data on malaria incidence after the campaign concluded.

5.2.2 Self-reported malaria incidence after the campaign

Table 4 shows the impact on malaria incidence in the household since the launch of the campaign in August 2020, self-reported in the cross-sectional survey conducted 1-3 months after the end of the campaign. The table reports incidence both at the extensive (if any household member had malaria, columns 1-4) and intensive margins (share of household members who had malaria, columns 5-8). In both cases, the estimates suggest null impact of the ads on overall malaria incidence (columns 1-2 and 5-6). Yet, columns (3-4) and (7-8) unmask substantial heterogeneity in ad effectiveness, showing a significant effect on households living in solid dwellings. For them, the ad campaign reduced the probability that any member reported malaria by 2.4 p.p., a 44% reduction compared to control and, consequently, lowered the share of infected members by 0.8 p.p., a 53% decrease.

Figure 4 provides evidence about the role of bed net ownership on moderating the effectiveness of advertisements. The left panel shows that the reduction in malaria cases is more substantial and statistically significant for households living in solid dwellings equipped with at least one bed net. The right panel shows that among households with bed nets, the decrease is greater for those that have at least one bed net for every two members. This suggests that a minimum number of bed nets per member is necessary for the ad campaign to affect outcomes significantly. Finally, the figure also confirms the absence of any effects among non-solid dwellers, for which the coefficient estimates are not significant and point towards the opposite direction.

Table D10 displays results concerning bed net ownership and the reapplication of insecticide to existing ones. There is no indication that the campaign influenced the likelihood of individuals owning a bed net (columns 1-3), or their inclination to purchase one among those who did not already own one (columns 4-6).[34] On the other hand, the ad campaign increased the likelihood by 6.7 p.p. (or 17%) that bed net owners in solid dwellings re-treated their mosquito nets with insecticide (columns 7-9 of Annex Table D10). While the ads did not explicitly suggest this action, they probably reminded people about the importance of insecticide treatment to repel mosquitoes and enhance the effectiveness of bed nets.

Finally, Annex Tables D11 presents impacts on individuals' concerns about getting malaria in the next year, respectively. Among respondents living in solid dwellings and with no mosquito nets

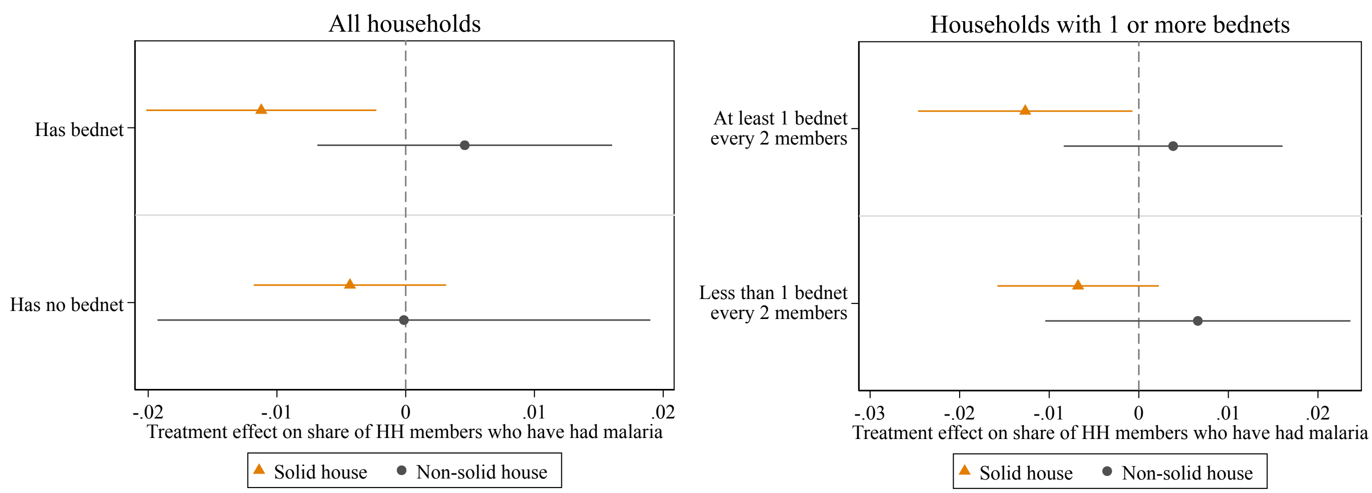
[34] This finding is also in line with the content of the campaign, which was not explicitly trying to persuade individuals to purchase bed nets.

Table 4: Incidence of malaria (since campaign launch)

	1 if any HH member had malaria				Share of HH members who had malaria			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full	Full	Full	Solid house	Full	Full	Full	Solid house
T	0.003	0.001	0.017	-0.024**	-0.002	-0.002	0.003	-0.008***
	(0.011)	(0.009)	(0.013)	(0.011)	(0.003)	(0.003)	(0.004)	(0.003)
T × Solid house			-0.038**				-0.011**	
			(0.018)				(0.005)	
Solid house		-0.019*	-0.001			-0.007**	-0.001	
		(0.010)	(0.014)			(0.003)	(0.005)	
Individual controls		✓	✓	✓		✓	✓	✓
Cluster controls		✓	✓	✓		✓	✓	✓
Observations	3052	3052	3052	1316	3052	3052	3052	1316
Adj. R-squared	-0.000	0.028	0.029	0.020	0.001	0.026	0.027	0.033
Mean in control	0.058	0.058	0.058	0.055	0.016	0.016	0.016	0.015
SD in control	0.235	0.235	0.235	0.229	0.081	0.081	0.081	0.076

Each observation is an individual. Responses collected from January 7, 2021 to March 31, 2021 are included. In columns (1-4), the dependent variable takes value 1 if the respondent/family member has had malaria since August 2020. In columns (5-8), the dependent variable is the share of household members who have had malaria since August 2020. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and of an air conditioner. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 4: Heterogeneous treatment effects by bed net ownership



Notes: The figure reports coefficient estimates from the regression described in columns (5-6) of Table 4, estimated on different sub-samples, based on bed net ownership and across dwelling type. Non-solid houses include semi-solid dwellings. Bars represent the 95% confidence intervals. Standard errors are clustered at the district level.

at home, the campaign increased by 6.7 p.p. or (31%) the share of individuals reporting to be extremely worried about getting malaria. However, as discussed above, we do not find evidence that this increased level of concern translated into higher bed net ownership and lower malaria incidence. Consistently with the above evidence, we find no significant effects among respondents living in non-solid dwellings.

5.3 Evidence from Administrative Data

Table 5 shows the campaign's impact on malaria incidence, expressed as the number of diagnosed monthly cases per million people recorded by health facilities for each sub-district-month. We report the interaction term (Treatment×Post) as the analysis of the administrative data is based on a difference-in-difference specification (see equation 2). We show a parsimonious specification with no controls, as well as with the results with the inclusion of district-level variables, and additionally controlling for subdistrict-level characteristics.[35]

Table 5: Malaria monthly incidence rates (N. of monthly cases/1M people)

	Overall incidence			Urban incidence			Rural incidence		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated × Post	-3.12	-5.66	-4.68	-6.10*	-6.96**	-6.21**	0.99	-1.47	-2.24
	(6.37)	(6.25)	(5.37)	(3.41)	(2.92)	(2.86)	(6.02)	(6.37)	(5.79)
Subdistrict FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Month and year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
District controls		✓	✓		✓	✓		✓	✓
Subdistrict controls			✓			✓			✓
Observations	1498	1498	1498	791	791	791	1338	1338	1338
Adj. R-squared	0.189	0.188	0.187	0.451	0.455	0.461	0.151	0.149	0.148
Baseline mean incidence	20.89	20.89	20.89	20.89	20.89	20.89	20.89	20.89	20.89

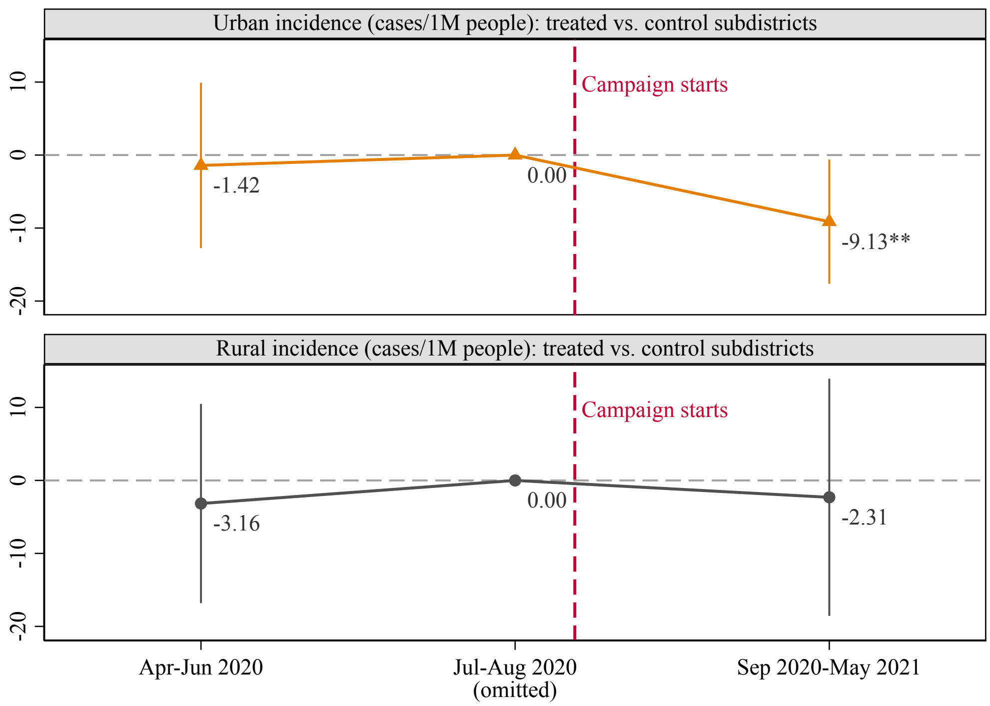
Notes: Each observation is a subdistrict-month-year. The panel spans from April 2020 to May 2021. T takes value 1 if the subdistrict belongs to a treated district. Post takes value 1 after August 2020, i.e. when the Facebook campaign started. Subdistrict controls include: subdistrict area, share of urban population, elevation, terrain ruggedness and distance to the state capital, all interacted with Post. District controls include total population from the Census as well as the shares of baseline survey respondents living in kutcha dwellings, with university degree, sleeping under mosquito nets, and reporting any malaria cases in the last 5 years and 2 weeks, all interacted with Post. All regressions include state fixed-effects interacted with Post. Standard errors are clustered at the district level. Data come from the Indian Health Management Information System. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Columns (1-3) examine overall incidence across both urban and rural health facilities. Although the coefficients are negative, they are not statistically significant. We observe that the campaign was effective in urban settings but not in rural ones. Column (4) shows that in treated subdistricts the number of monthly urban malaria cases per million people decreased by more than 6 cases in the months during and after the campaign, compared to the previous period. The coefficient is significant at the 90% confidence level and its magnitude and significance are robust to controlling for district-level controls (column 5) and subdistrict-level variables (column 6). Specifically, in the fully-saturated specification, the coefficient is significant at the 95% confidence level and suggests a decrease of 6.2 cases in the number of monthly urban cases per million people, which corresponds to about 30% of the overall monthly malaria incidence rate in the pre-campaign period. Finally, we do not find significant effects of the campaign on rural incidence. Columns (8-9) suggest a potential reduction, but the magnitudes are only a quarter of those observed in the urban sample, and the coefficients are not statistically significant.

[35] District-level controls include total population from the census as well as the shares of baseline survey respondents living in non-solid or kutcha dwellings, with university degree, sleeping under mosquito nets, and reporting any malaria cases in the last 5 years and 2 weeks, all interacted with Post. Subdistrict controls include subdistrict area, share of urban population, elevation, terrain ruggedness and distance to the state capital, all interacted with Post.

The differential effect of the ad campaign on recorded malaria incidence across urban and rural settlements aligns with the survey-based evidence. According to the 2019-2021 India Demographic and Health Survey (India IIPS, 2022), solid dwellings represent almost 85% of the total housing units in urban areas, while they are less than 48% in rural areas. Evidence presented above showed significant increases in bed net adoption and decreases of malaria incidence among respondents and households residing in solid houses, with no impacts on households residing in non-solid dwellings. Hence, one may expect that recorded malaria cases are more likely to decrease in localities with a larger proportion of solid dwellers.

Figure 5: Trends in malaria incidence rates (N. of monthly cases/1M people)



Notes: Each point reports the coefficient on the interaction term Treatment×Period, where Period is a dummy variable equal to 1 at each period on the graph (i.e., Apr-Jun 2020, Jul-Aug 2020 and Sep 2020-May 2021). The omitted period is Jul-Aug 2020. Two separate regressions are estimated for rural and urban incidence rates. Regressions include subdistrict and period fixed-effects, as well as district-level controls interacted with period dummies. Standard errors are clustered at the district level. Bars represent the 95% confidence intervals.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Annex Table D19 complements our evidence on malaria incidence with an analysis of testing behaviors and positivity rates. It shows non-significant effects on the number of malaria tests conducted for neither settlement category. However, consistent with the reduction in incidence, the table shows a statistically significant reduction in test positivity rates: the share of positive tests out of the total tests conducted decreased by 1.7 p.p. overall, and by 2.2 p.p. in urban localities, while the effect is virtually null in rural areas.

The identifying assumption behind the Diff-in-Diff estimates presented in Table 5 is the parallel trends assumption: in the absence of the campaign, the *change* in incidence rates in the treatment and control subdistricts should have been the same. This assumption is plausible due to the cluster-level randomization, and we provide evidence to support it. By grouping observations in different time periods, we can depict the differences in malaria incidence between treated and control localities over time. The result, shown in Figure 5, allow us to detect potentially diverging trends before the campaign (i.e., between April and August 2020). Importantly, the figure supports the parallel trend assumption, as trends in incidence rates across treated and control subdistricts are comparable before the campaign, providing reassurance on the identification strategy underlying the analysis of administrative data. The figure also confirms that the campaign significantly impacted urban malaria incidence while having negligible effects on rural cases.

5.4 Placebo Regressions and Robustness Checks

In this section, we present a series of placebo regressions and robustness checks for our results. As demonstrated, our findings remain consistent and reliable across these various checks.

For our survey-based findings, we conducted the following specification checks and placebo exercises. First, we assessed the sensitivity of our results to the inclusion of individual- and district-level covariates, and reported them in the tables above (and in the Appendix). The coefficients showed minimal changes, and the effects remained stable. Second, we estimated logistic regressions for all the binary outcomes presented above. The results are shown in Annex Tables D12-D14. The estimated marginal effects for the treatment indicator reported at the bottom of the tables are identical to the OLS estimates presented so far, validating the use of a Linear Probability Model in this context. Third, since we measured bed net usage at baseline (i.e., before the campaign started), we can conduct placebo regressions to further validate the randomization. Annex Table D15 shows the results of this exercise, where the dependent variable measured in July and August 2020 is regressed on the treatment indicator. The table provides strong evidence in favor of a correct randomization: all coefficients are small and statistically insignificant. Moreover, there is no evidence of heterogeneous treatment effects across dwelling type. Fourth, we measured individuals' concerns about getting COVID-19 in the next year and used this as a placebo outcome, since the campaign should not directly affect it. Consistently, Annex Table D16 shows no impact on risk perceptions about getting COVID-19. This alleviates the potential concerns that our results are driven by reporting bias.

For our health facility data, we test the robustness of our results to different models, samples, and outcome definitions. First, as shown in Table 5, the coefficients of our impact estimators showed minimal changes with or without district- and subdistrict-level controls. Second, Annex Table D17 shows the sensitivity of our estimates to five choices of the population thresholds used to identify subdistricts to include in our sample. The impact of the ad campaign on urban incidence is stable and mostly significant across all subsamples, with estimated reductions in cases ranging between 20% and 30% of the baseline average incidence rate. Conversely, the sign of the coefficient on rural incidence is inconsistent and remains statistically insignificant. Third, in Annex Table D18 we consider the absolute number of malaria cases as dependent variable, and include the subdistrict population among the controls. Columns (1-3) report the estimates from a linear model using the *log* number of cases, while columns (4-6) report the results from a Poisson regression. In both cases, the results are closely aligned with the evidence presented so far, indicating that the campaign had beneficial effects in urban localities only. For example, the *log* specification estimates a 30% reduction in urban incidence, and the effect is even larger in the Poisson model (both estimates are significant at the 95% confidence level).

6 Mechanisms: Limited Effectiveness or Limited Reach?

The analysis of survey and health facility data indicates that the MNM campaign benefited households residing in solid dwellings and in urban settlements. However, it did not benefit households that are most susceptible to malaria. What mechanism underlies these differential effects? There are two potential explanations for our results. First, the ad content itself may not have been effective for the sub-population residing in non-solid dwellings and rural areas. This could indicate shortcomings in the identification and ad design for these type of personas. Second, the ad campaign might not have sufficiently reached this key sub-population. This underscores the importance of accounting for algorithmic targeting (Lambrecht and Tucker 2019; Sapiezynski et al. 2022) when aiming to engage with underrepresented, geographically dispersed (and therefore more expensive to reach) populations on social media platforms. To disentangle these potential mechanisms, we employed two additional analyses. First, in a post-campaign survey we included ad-recall questions to assess how well participants remembered the campaign messages. Second, we implemented a *feed* experiment where we ensured exposure to the campaign for both household types. The results of these analyses are detailed below.

6.1 Post-Campaign Recall

Our post-campaign survey measured ad recall for the MNM campaign. If the ads reached both solid and non-solid dwellers at similar frequencies, then one would expect both audiences to equally recall the Facebook advertising or the advertiser itself. To measure this, in the last wave of the longitudinal sample (completed between mid-December 2020 and the end of January 2021), we explicitly asked respondents if they remembered having seen an ad from Malaria No More on Facebook since August 2020. Table 6 shows the results. Columns (1-2) show a 3 p.p. increase in ad recall across all treated respondents, corresponding to a 27% lift with respect to the control group. However, columns (3-6) show that the effect is driven by individuals residing in solid dwellings, where ad recall increased by 5 p.p., a 50% lift with respect to the control. For non-solid dwellers, the coefficient is close to zero and statistically insignificant. Estimates from a logistic model reported in Annex Table D20 confirm these results, with marginal effects being identical to the OLS coefficients.

The analysis of treatment heterogeneity for the recall question provides initial evidence that the ad campaign did not reach vulnerable populations. Annex Table D21 explores how ad recall varies among younger respondents and those with higher concern about malaria. These groups were

Table 6: Ad recall at endline

	All respondents		Solid house		Non-solid house	
	(1)	(2)	(3)	(4)	(5)	(6)
T	0.030**	0.033**	0.049**	0.049**	0.011	0.009
	(0.014)	(0.013)	(0.024)	(0.023)	(0.023)	(0.023)
Individual controls		✓		✓		✓
Cluster controls		✓		✓		✓
Observations	1492	1492	728	728	764	764
Adj. R-squared	0.000	0.002	0.003	0.004	-0.001	0.001
Mean in control	0.114	0.114	0.101	0.101	0.126	0.126

Each observation is an individual. Responses collected in the endline survey between December 15, 2020 and January 31, 201 are included. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and air conditioner, and 5-year malaria incidence at baseline. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

theoretically more likely to be exposed to and remember the campaign.[36] While the table confirms both hypotheses, the effect is only evident for respondents living in solid dwellings. Among them, ad recall significantly increased only for individuals under 26 (by 16 p.p.) and for those expressing at least some concern about malaria (by 9 p.p.). By contrast, among residents of non-solid dwellings, the coefficients for both younger respondents and those more concerned about malaria were close to zero and statistically insignificant. This lack of recall effect, including those theoretically more receptive to the campaign, suggests the ad content may not have effectively reached this population. To substantiate this hypothesis, we conducted the following *feed* experiment.

6.2 Individual-level *Feed* Experiment

To differentiate between limited ad effectiveness and limited reach of the ad campaign, in April 2021 we nested an individual-level study at the end of the cross-sectional survey (January-March 2021, see timeline 4.3). This study ensured that the campaign ads were distributed equally to both solid and non-solid dwellers (balancing also across gender, caste, education, as well as the treatment assignment of their district in the MNM ad campaign). In other words, by maintaining exposure to the campaign ads constant, this experiment offers a direct assessment of ad effectiveness on both types of households.

Specifically, after completing the cross-sectional survey of the broader trial, respondents were asked if they would be open to being contacted again within 1 to 3 months. To enhance statistical power, we also recruited a new group of participants using the same Facebook Ads and methodology of the longitudinal and cross-sectional samples. Those who agreed to take part in this study were randomly placed into either a treatment or control group, ensuring that each dwelling type was equally represented in each group. We used Facebook remarketing tool to show ads in the Facebook

[36] We expect ad recall to be larger among younger respondents, since they spend more time on social media and thus have more exposure to ads. We also expect larger recall rates among respondents with higher levels of concern about getting malaria, since they would pay more attention to malaria-related ads if these appeared on their feeds.

feed of our study participants. Since the respondents took our survey on Messenger, we could use anonymous identifiers to create a "Custom Audience" for the treatment group.[37] Subsequently, we handed over this audience to MNM, which then remarketed ads from their own Facebook page for two weeks, mirroring the nationwide campaign.[38] The remarketing campaign reached 2,119 individuals (about 90% of the total audience) at least once in two weeks, for an average frequency of 4.9 ads per person. These ads were displayed in the users' feeds, alongside with other competing ads and organic content from various sources. Crucially, participants were unaware that the ads were targeted at them based on their survey participation. Consequently, the remarketing campaign closely resembled the nationwide campaign, lending ecological validity to the individual-level study.

Two weeks after the end of the remarketing campaign, we invited respondents to fill out a follow-up survey. Those who completed it are included in the individual study sample (N=1,542). Annex Table D22 shows the distribution of individuals' and households' characteristics in the treatment and control conditions. The sample is strongly balanced across most covariates such as age, gender, education and bed net ownership, with a few small imbalances on household size and unemployment rate. We include a battery of control variables in the regressions to show that our results are robust to these small imbalances.

We focus on two primary behavioral outcomes where we anticipated effects within our two-week window: bed net usage and seeking medical attention in case of a high fever. To facilitate the comparison, the questions were similar to those used in the longitudinal panel survey (see Annex B). To mitigate priming and survey effects, we asked these questions only once, at follow-up. Our workhorse model is a Linear Probability Model, which is estimated using Ordinary Least Squares. It is represented in the following form:

$$y_{id} = \alpha + \beta_1 \, Individual \; T_{id} + \beta_2 \, T_d + \gamma \, \mathbf{x}_{id} + \epsilon_{id} \tag{3}$$

where i is an individual residing in district d, *Individual* T is the individual-level random treatment indicator for the remarketing campaign, which varies across individuals. T is the district-level random treatment indicator, which varies across districts and captures the potential long-lasting effects of the nationwide campaign. $\mathbf{x}$ represents a vector of individual-level control covariates, which are described in Annex Table D22. We adjust the standard errors to account for heteroskedasticity.

6.3 Targeting Matters

Table 7 reports effects on bed net usage. We find a positive and significant effect of the individual level treatment on the probability that a respondent has slept under a mosquito net the previous

[37] Remarketing is a tool offered by the major ad platforms by which an ad campaign can be run to a group of individuals who have previously interacted with the advertiser in some way. This implies that the individual has either come directly from an ad to a platform (website, app, chatbot, etc.) controlled by the advertiser, has directly connected their account (i.e., via "Login with Facebook"), or has shared personally identifying data and given the advertiser permission to use that data to market to or communicate with them.

[38] This campaign featured the highest-performing ads from the earlier nationwide campaign, which had received the bulk of the budget. See Annex Figure for a selection of these ads A7).

night: the estimated marginal effect is around 7.5 and 8 p.p., and implies an increase in the probability of sleeping under a mosquito net from 66% in the control group to 73% in the treatment group, a 11% increase. This effect maintains its effect size and statistical significance across all specifications and is robust to the inclusion of controls (columns 1-3). Note that the impact of the remarketing campaign is larger compared to the estimated treatment effect of the nationwide MNM campaign described in Section 5.2. A plausible explanation is that the coefficients in the cluster design represent intent-to-treat effects, meaning only a subset of respondents (whose exact number is unknown to us) in our longitudinal sample were exposed to the treatment. In contrast, the individual-level design ensures that the majority of the individuals in the treated group (approximately 90%) saw the ads at least once. Therefore, the effects observed at the individual level are larger and more aligned with treatment-on-the-treated estimates.

Table 7: Sleeping under bed net (individual-level study)

Y=1 if respondent has slept under bed net the previous night

	Any dwelling type			Solid house	Non-Solid house
	(1)	(2)	(3)	(4)	(5)
Individual T	0.075***	0.079***	0.079**	0.076*	0.081**
	(0.029)	(0.029)	(0.038)	(0.043)	(0.039)
Individual T × Solid house			0.000		
			(0.057)		
Solid house	-0.115***	-0.080**	-0.080*		
	(0.029)	(0.031)	(0.044)		
Treatment	0.019	0.021	0.021	0.058	-0.015
	(0.029)	(0.029)	(0.029)	(0.043)	(0.039)
Individual controls		✓	✓	✓	✓
Observations	1043	1043	1043	513	530
Adj. R-squared	0.018	0.033	0.032	0.026	0.030
Mean in control	0.654	0.654	0.654	0.595	0.706
SD in control	0.476	0.476	0.476	0.492	0.456

Each observation is an individual. Invidivual T is the individual treatment indicator for the remarketing campaign. T is the indicator of the district-level treatment assignment. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house and ownership of an air conditioner. Heteroskedasticity-robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Columns (3-5) of Table 7 provide a direct test for our mechanisms and support the hypothesis that the MNM campaign did not sufficiently reach those living in non-solid dwellings. In fact, unlike in the cluster-level treatment, column (3) exhibits no heterogeneous treatment effects of the remarketing campaign across dwelling types, suggesting that -*conditional on receiving the ads- the campaign is similarly effective for both types of dwellers.* This is further supported by columns (4-5), which show a positive and similar impact of the individual treatment on bed net use for both solid and non-solid dwellers (around 8 p.p.). Given the baseline differences in the dependent variable across dwelling types, the estimated effect is of 13% in solid-houses and of 11.5% in non-solid dwellings. Estimates from a logistic model yield identical marginal effects (Annex Table D23).

Note that in all the regressions we also include the cluster-level treatment indicator. Although statistically insignificant, the magnitude and sign of the estimated coefficients are different across dwelling types. In particular, among respondents living in solid houses, the coefficient (5.8 p.p.) is very similar to the treatment effect of the MNM campaign estimated in the cluster design (5.5 p.p.). This point out persistent effects of advertising, although we are unable to precisely estimate them due to limited power. Since half of the respondents in the individual study sample lived in treatment districts, their potential exposure to the nationwide campaign might reduce the impact of the individual treatment. In Annex Table D25 we replicate the above regressions on the subsamples of individuals assigned to treatment and control districts during the MNM campaign. The results show that the individual treatment produces higher effects among respondents who were unlikely to be previously exposed to the ad content (i.e., those residing in control districts). For them, the coefficient estimates are around 10-11 p.p., indicating a lift of 18% in solid houses and 16% in non-solid dwellings.

Table 8: Timely treatment seeking in case of high fever (individual-level study)

Y=1 if respondent would seek treatment within 24 hours of the onset of a high fever

	Any dwelling type			Solid house	Non-Solid house
	(1)	(2)	(3)	(4)	(5)
Individual T	0.026**	0.025**	0.023	0.026	0.023
	(0.013)	(0.013)	(0.016)	(0.019)	(0.016)
Individual T × Solid house			0.004		
			(0.025)		
Solid house	-0.030**	-0.030**	-0.031		
	(0.013)	(0.014)	(0.020)		
T	0.006	0.007	0.007	0.001	0.011
	(0.013)	(0.013)	(0.013)	(0.019)	(0.016)
Individual controls		✓	✓	✓	✓
Observations	1542	1542	1542	807	735
Adj. R-squared	0.004	0.003	0.003	-0.004	0.003
Mean in control	0.921	0.921	0.921	0.905	0.937
SD in control	0.270	0.270	0.270	0.294	0.243

Each observation is an individual. Indivitual T is the individual treatment indicator for the remarketing campaign. T is the indicator of the district-level treatment assignment. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of a bednet and of an air conditioner. Heteroskedasticity-robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8 reports effects on individuals' intentions to seek treatment in the event of a high fever (100.4°F / 38°C or above). Our dependent variable is equal to 1 if the respondent would seek treatment within 24 hours of the onset of a fever. The table shows that the ads were effective at encouraging individuals at seeking treatment promptly. Across all dwelling types (columns 1-2), the estimated coefficients suggest a 2.5 p.p. increase in the share of respondents who would timely seek treatment, a 2.7% lift compared to the control group. Columns (3-5) all seem to suggest the absence of heterogeneous effects across dwelling type. Although not significant because of limited power, the individual treatment effect in each subsample is similar to the overall effect, suggesting that the ad content was effective for both dwellers. Identical results are obtained when we use logistic regression (Annex Table D24).

6.4 Implications for Practitioners

The policy implication of our mechanisms analysis is clear. Organizations running social media ad campaigns with social objectives should adopt more nuanced targeting approaches to effectively reach their target populations. Specifically, these organizations need to implement *explicit* targeting strategies to overcome optimization algorithms of social media platforms that prioritize users that are cheaper to reach to maximize overall engagement numbers. In our study setting, for example, these are wealthier users that may frequently check their social media but are at lower risk of contracting malaria.

In our evaluation of the real-world campaign, Malaria No More employed geographical targeting based on states with high incidence rates, and also segmented audiences using various demographic groups. These groups were defined by age, gender, and levels of urbanization, reflecting the personas of interest. Yet, these audiences were likely too broad, and did not adequately consider the higher engagement costs of users living in non-solid dwellings. Our trial shows that more sophisticated targeting tools, like Facebook Lookalike Audience, can enhance the targeting precision of ad campaigns. Our recruitment method via Facebook ads indicates that reaching individuals in non-solid dwellings could be almost 10 times more costly than those in solid dwellings. Therefore, unless we specifically target these more vulnerable groups, the optimization algorithms are likely to concentrate on lower-risk individuals in order to meet engagement metric goals within a set budget.

We propose a three-step approach for practitioners aiming to ensure their social media ads effectively reach and impact populations of their interest. First, during the formative research stage, organizations should identify these target populations where the potential impact of intervention is greatest. Second, practitioners should move beyond the basic targeting options provided by online platforms and employ more sophisticated tools such as Lookalike models to ensure inclusion of these key groups in the campaign's audience. Third, practitioners should be mindful of the trade-off between maximizing on-platform engagement and concentrating on vulnerable populations that are most likely to benefit from the campaign. Although the cost to reach and engage these at-risk users may be higher, the potential for significant developmental impact per individual reached is also greater.

7 Cost-effectiveness

In this section, we evaluate the cost-effectiveness of the investment in advertising. To determine the range of overall effectiveness in terms of malaria cases avoided per million per month, as *recorded* in health facility records, we use the most conservative impact estimates from Table 5. These include monthly reductions in overall national incidence (3.12 cases) and urban incidence (6.10 cases). Considering that these impacts represent the average during the treatment and after-treatment period from September 2020 to May 2021—a span of 9 months—we multiply the monthly reduction by 9. This calculation suggests a total estimated reduction in reported malaria incidence of approximately 28.08 and 54.90 avoided cases per million people in the three study states.

To extrapolate our results to the 22 states targeted by the MNM national campaign, we assume the effectiveness observed in the three study states mirrors the remaining 19 states. With a total population of about 1.2 billion across the 22 states, and roughly 100 million in control districts within the three study states, the campaign potentially reached 1.1 billion individuals. Consequently, the estimated range of avoided malaria cases attributed to the campaign is between 30,888 and 60,390 cases.

The total ad costs for the national campaign across the 22 states was $202,000 USD, resulting in an estimated cost per avoided case of malaria ranging from $3.35 (for urban cases) to $6.54 (for overall incidence). These figures represent the upper bound estimates of the campaign's cost-effectiveness. Given that the estimated number of malaria cases is about twenty-two times higher than reported cases(WHO, 2021), incorporating these unreported cases could reduce the cost per averted case to below 30 cents in this developing world setting.

While the initial focus here is on direct advertising costs,[39] it is also valuable to compare these figures with the outcomes from similar initiatives in health communication. For instance, the costs per reported malaria case averted are comparable to those seen in social media campaigns promoting COVID-19 vaccinations in developed countries, which range from $3.41 per influenced individual to $5.68 per vaccinated person (Athey et al., 2023). Although under-reporting is likely substantially lower for COVID-19 vaccines and in developed countries, both studies highlight the cost-effectiveness of using social media campaigns for development objectives.

Within malaria prevention interventions, our estimated cost per reported case averted is comparable—and substantially lower for overall averted cases—to that of distributing insecticide-treated nets, which typically range between $5 and $6 per averted case (Conteh et al., 2021). In comparison to traditional offline programs aimed at eradicating malaria, online campaigns not only provide high cost-effectiveness but also offer several other advantages. These include the ease of testing and adapting content and delivery strategies for specific sub-populations and the ability to scale up quickly. Such benefits facilitate the application and rapid expansion of this innovative approach for developmental policies.

However, online campaigns may not be effective in areas with low internet penetration, as there is little opportunity for information to spread to those without internet access. Furthermore, in communities with very few social media users, the costs of reaching these individuals online may be prohibitively high, making offline information campaigns more economical in such cases. Generally, advertising campaigns with public health goals should be seen as supplements to on-the-ground efforts, primarily because of their low costs and their high cost-effectiveness.

8 Conclusions

Social media campaigns are increasingly used for public health objectives, yet measuring their effectiveness is difficult and requires the availability of longer-term attitudinal and behavioral outcomes

[39]The costs detailed above are limited to advertising expenses and do not include staff and management costs associated with designing and executing the national ad campaign. These will be included in a future draft.

that are not directly measurable on the ad platform itself. Malaria prevention campaigns offer an example. Malaria is a disease which disproportionately affects rural people in India. Behaviorally informed ads could provide information and persuade individuals to take precautionary measures. However, while digital advertising provides the ability to micro-target messages to specific subgroups of the population, certain segments remain harder/more expensive to reach than others. At the same time, measuring the impact of such ads on offline behaviors like using bed nets is challenging, and requires availability of individual survey responses that could be connected with ad exposure data.

In this paper, we first introduce a new methodology and a novel toolkit for conducting surveys and experiments on social media. We then use this tool to evaluate a Facebook malaria-prevention ad campaign designed and run by Malaria No More, a global public health non-profit with a strong presence in India. The campaign's goals were the promotion of behaviors to prevent malaria infection and to seek timely treatment in case of symptoms. It consisted of more than 50 creatives targeted to different personas. With an average frequency of 3.6 impressions per user, the campaign had a total reach of about 130 million people on Facebook and Instagram over 6 months. In the evaluation, we consider the question "Did the campaign work?" along with the question "Did it work for those most at risk?".

We recruited two groups of respondents (a longitudinal and a cross-sectional sample) on Facebook and Instagram offering mobile credit as an incentive to fill out a series of chat surveys conducted on Messenger. We stratified the recruitment across 80 districts in three high-burden states, and used real-time optimization to guarantee representativeness of individuals living in non-solid dwellings, where malaria risk is higher. We then worked with the non-profit advertising team to exclude 40 randomly-picked districts from the delivery of the campaign. Relying on a set of self-reported attitudes and behavioral outcomes, we found evidence that the ad campaign increased the use of bed nets among individuals living in solid houses (by 11% compared to control), while it was ineffective for households living in non-solid dwellings. Moreover, self-reported household malaria incidence decreased by 44-53% among those living in solid houses compared to control, but not among non-solid dwellers. The reduction in malaria incidence was predominantly driven by bed net owners. Consistently, the analysis of administrative health facility data showed that monthly malaria incidence in urban areas decreased by 6.2 cases per million people. This is about 30% of the overall monthly malaria incidence rate in the pre-treatment period and implies an upper bound cost per averted *reported* urban case of USD 3.4. By contrast, rural incidence was not significantly affected.

Was the ad content ineffective for households at greater malaria risk or did it fail to sufficiently reach them? To answer this question, in a second trial we experimentally varied exposure to the same ads at the individual level, using the re-marketing tools of the ad platform. We found that bed net use and timely treatment seeking increased for both types of households, suggesting that social media campaigns need to invest in improved targeting strategies to reach their development objectives.

Our findings demonstrated the value of data for targeting users and tracking them to measure ad effectiveness, contributing to the ongoing debate on privacy and the use of third-party data in digital marketing. We proposed a series of micro-targeting techniques to maximize impacts in sub-populations of interest. By integrating privacy considerations into our experimental methodology, our approach highlighted how first-party data, directly collected from users on social media, can inform targeting and help reach the most valuable segments of the population—in our context, those most at risk of malaria. The methodology used in this paper is sector-agnostic and can be applied to both social and commercial campaigns.

Funding and Competing Interests

This research has received funds from the World Bank's Development Impact Department (DIME) and the Facebook's Campaigns for a Healthier World initiative. Three authors do not have any competing interests to this work. One author has ownership interests in Virtual Lab LLC, a company that uses the open-source methodology described in the paper.

References

Acquisti, A., Taylor, C. and Wagman, L. (2016), 'The economics of privacy', *Journal of economic Literature* **54**(2), 442–492.

Ajzen, I. (1991), 'The theory of planned behavior', *Organizational behavior and human decision processes* **50**(2), 179–211.

Anderson, S. J., Chandy, R. and Zia, B. (2018), 'Pathways to profits: The impact of marketing vs. finance skills on business performance', *Management Science* **64**(12), 5559–5583.

Andreasen, A. R. (1994), 'Social marketing: Its definition and domain', *Journal of Public Policy and Marketing* **13**(1), 108–114.
URL: *http://www.jstor.org/stable/30000176*

Athey, S., Grabarz, K., Luca, M. and Wernerfelt, N. (2023), 'Digital public health interventions at scale: The impact of social media advertising on beliefs and outcomes related to covid vaccines', *Proceedings of the National Academy of Sciences* **120**(5), e2208110120.
URL: *https://www.pnas.org/doi/abs/10.1073/pnas.2208110120*

Babcicky, P. and Seebauer, S. (2019), 'Unpacking protection motivation theory: evidence for a separate protective and non-protective route in private flood mitigation behavior', *Journal of Risk Research* **22**(12), 1503–1521.

Baldwin, A. S., Rochefort, C. and Geary, B. (2022), 'Understanding health behaviour change: Guiding theoretical models'.

Bart, Y., Stephen, A. T. and Sarvary, M. (2014), 'Which products are best suited to mobile advertising? a field study of mobile display advertising effects on consumer attitudes and intentions', *Journal of Marketing Research* **51**(3), 270–285.

Botzen, W. W., Kunreuther, H., Czajkowski, J. and de Moel, H. (2019), 'Adoption of individual flood damage mitigation measures in new york city: An extension of protection motivation theory', *Risk analysis* **39**(10), 2143–2159.

Brecko, K. and Hartmann, W. R. (2023), 'Pro-social change for the most challenging: Marketing and testing harm reduction for conservation'.

Bulte, E., Lensink, R. and Vu, N. (2017), 'Do gender and business trainings affect business outcomes? experimental evidence from vietnam', *Management Science* **63**(9), 2885–2902.

Burchell, K., Rettie, R. and Patel, K. (2013), 'Marketing social norms: social marketing and the 'social norm approach'', *Journal of Consumer Behaviour* **12**(1), 1–9.

Carins, J. E. and Rundle-Thiele, S. R. (2014), 'Eating for the better: A social marketing review (2000–2012)', *Public health nutrition* **17**(7), 1628–1639.

Chandy, R. K., Johar, G. V., Moorman, C. and Roberts, J. H. (2021), 'Better marketing for a better world', *Journal of Marketing* **85**(3), 1–9.

Cohen, J., Dupas, P. and Schaner, S. (2015), 'Price subsidies, diagnostic tests, and targeting of malaria treatment: evidence from a randomized controlled trial', *American Economic Review* **105**(2), 609–645.

Conteh, L., Shuford, K., Agboraw, E., Kont, M., Kolaczinski, J. and Patouillard, E. (2021), 'Costs and cost-effectiveness of malaria control interventions: a systematic literature review', *Value in Health* **24**(8), 1213–1222.

Croker, H., Lucas, R. and Wardle, J. (2012), 'Cluster-randomised trial to evaluate the 'Change for Life' mass media/ social marketing campaign in the UK', *BMC Public Health* **12**(1), 1.
URL: *BMC Public Health*

Davis, B., Grewal, D. and Hamilton, S. (2021), 'The future of marketing analytics and public policy', *Journal of Public Policy & Marketing* **40**(4), 447–452.

Dev, V., Phookan, S., Sharma, V. P. and Anand, S. P. (2004), 'Physiographic and entomologic risk factors of malaria in assam, india', *The American journal of tropical medicine and hygiene* **71**(4), 451–456.

Donati, D., Orozco-Olvera, V. and Rao, N. (2022), 'Using social media to change gender norms: An experiment within facebook messenger in india', (10199).
URL: *http://hdl.handle.net/10986/38113*

Evans, W. D., Bingenheimer, J. B., Long, M., Ndiaye, K., Donati, D., Rao, N. M., Akaba, S., Nsofor, I. and Agha, S. (2023), 'Outcomes of a social media campaign to promote covid-19 vaccination in nigeria', *Plos one* **18**(9), e0290757.

Fogg, B. J. (2009), A behavior model for persuasive design, *in* 'Proceedings of the 4th international Conference on Persuasive Technology', pp. 1–7.

Ghosh Dastidar, A., Sunder, S. and Shah, D. (2023), 'Societal spillovers of tv advertising: Social distancing during a public health crisis', *Journal of Marketing* **87**(3), 337–358.

Godinho de Matos, M. and Adjerid, I. (2022), 'Consumer consent and firm targeting after gdpr: The case of a large telecom provider', *Management Science* **68**(5), 3330–3378.

Goldfarb, A. and Tucker, C. (2011*a*), 'Online display advertising: Targeting and obtrusiveness', *Marketing Science* **30**(3), 389–404.

Goldfarb, A. and Tucker, C. E. (2011*b*), 'Privacy regulation and online advertising', *Management science* **57**(1), 57–71.

Gordon, B. R., Jerath, K., Katona, Z., Narayanan, S., Shin, J. and Wilbur, K. C. (2021), 'Inefficiencies in digital advertising markets', *Journal of Marketing* **85**(1), 7–25.

Hayes, R. J. and Moulton, L. H. (2017), 'Cluster randomised trials, second edition', *Cluster Randomised Trials, Second Edition* pp. 1–398.

India IIPS (2022), *India national family health survey NFHS-5 2019-21.*, International Institute for Population Sciences - IIPS/India and ICF, Mumbai, India.
URL: *https://www.dhsprogram.com/pubs/pdf/FR375/FR375.pdf*

Johnson, G. A. (2023), 'Inferno: A guide to field experiments in online display advertising', *Journal of economics & management strategy* **32**(3), 469–490.

Johnson, G. A., Lewis, R. A. and Reiley, D. H. (2017), 'When less is more: Data and power in advertising experiments', *Marketing Science* **36**(1), 43–53.

Johnson, G. A., Shriver, S. K. and Du, S. (2020), 'Consumer privacy choice in online advertising: Who opts out and at what cost to industry?', *Marketing Science* **39**(1), 33–51.

Kotler, P. and Levy, S. J. (1969), 'Broadening the concept of marketing', *Journal of marketing* **33**(1), 10–15.

Kotler, P. and Zaltman, G. (1971), 'Social marketing: an approach to planned social change', *Journal of marketing* **35**(3), 3–12.

Lambrecht, A. and Tucker, C. (2019), 'Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads', *Management science* **65**(7), 2966–2981.

Lewis, R. A. and Reiley, D. H. (2014), 'Online ads and offline sales: measuring the effect of retail advertising via a controlled experiment on yahoo!', *Quantitative Marketing and Economics* **12**, 235–266.

Lindell, M. K. and Perry, R. W. (2012), 'The protective action decision model: Theoretical modifications and additional evidence', *Risk Analysis: An International Journal* **32**(4), 616–632.

Miaskiewicz, T. and Kozar, K. A. (2011), 'Personas and user-centered design: How can personas benefit product design processes?', *Design studies* **32**(5), 417–430.

Miller, G. and Mobarak, A. M. (2015), 'Learning about new technologies through social networks: experimental evidence on nontraditional stoves in bangladesh', *Marketing Science* **34**(4), 480–499.

Najafi, M., Ardalan, A., Akbarisari, A., Noorbala, A. A. and Elmi, H. (2017), 'The theory of planned behavior and disaster preparedness', *PLoS currents* **9**.

Narasimhan, L., Srinivasan, K. and Sudhir, K. (2015), 'Marketing science in emerging markets', *Marketing Science* **34**(4), 473–479.

Nielsen, L. (2013), *Personas-user focused design*, Vol. 15, Springer.

Orozco-Olvera, V. H. and Malhotra, N. (2023), *Media and Health Outcomes*, Springer International

Publishing, Cham, pp. 1–38.
URL: *https://doi.org/10.1007/978-3-319-57365-6*$_3$85 − 1

Rafieian, O. and Yoganarasimhan, H. (2021), ‘Targeting and privacy in mobile advertising’, *Marketing Science* **40**(2), 193–218.

Rao, N., Donati, D. and Orozco, V. (2020), ‘Conducting surveys and interventions entirely online: a virtual lab practitioner’s manual’, *World Bank: Washington, DC, USA* .

Rosenstock, I., Stretcher, V. and Becker, M. (1988), ‘Social learning theory and the health belief method’, *Health Education Quarterly* **13**, 73–92.

Sapiezynski, P., Ghosh, A., Kaplan, L., Rieke, A. and Mislove, A. (2022), Algorithms that" don’t see color" measuring biases in lookalike and special ad audiences, *in* ‘Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society’, pp. 609–616.

Schwenk, C. R. (1984), ‘Cognitive simplification processes in strategic decision-making’, *Strategic management journal* **5**(2), 111–128.

Sharma, A., Rana, S. K., Prinja, S. and Kumar, R. (2016), ‘Quality of health management information system for maternal & child health care in haryana state, india’, *PLoS One* **11**(2), e0148449.

Sharot, T. (2011), ‘The optimism bias’, *Current biology* **21**(23), R941–R945.

Shawky, S., Kubacki, K., Dietrich, T. and Weaven, S. (2019), ‘Using social media to create engagement: a social marketing review’, *Journal of Social Marketing* **9**(2), 204–224.

Singh, M. P., Saha, K. B., Chand, S. K. and Sabin, L. L. (2019), ‘The economic cost of malaria at the household level in high and low transmission areas of central india’, *Acta tropica* **190**, 344–349.

Singh, M. P., Saha, K. B., Chand, S. K. and Savargaonkar, D. (2020), ‘Socioeconomic determinants of community knowledge and practice in relation to malaria in high- and low-transmission areas of central india’, *Journal of Biosocial Science* **52**(3), 317–329.

Sun, J. J., Bellezza, S. and Paharia, N. (2021), ‘Buy less, buy luxury: Understanding and overcoming product durability neglect for sustainable consumption’, *Journal of Marketing* **85**(3), 28–43.

Tarozzi, A., Mahajan, A., Blackburn, B., Kopf, D., Krishnan, L. and Yoong, J. (2014), ‘Micro-loans, insecticide-treated bednets, and malaria: evidence from a randomized controlled trial in orissa, india’, *American Economic Review* **104**(7), 1909–1941.

Thompson, E. B., Heley, F., Oster-Aaland, L., Stastny, S. N. and Crawford, E. C. (2013), ‘The impact of a student-driven social marketing campaign on college student alcohol-related beliefs and behaviors’, *Social Marketing Quarterly* **19**(1), 52–64.

Tucker, C., Agrawal, A., Gans, J. and Goldfarb, A. (2018), ‘Privacy, algorithms, and artificial intelligence’, *The economics of artificial intelligence: An agenda* pp. 423–437.

Tversky, A. and Kahneman, D. (1974), ‘Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty.’, *science* **185**(4157), 1124–1131.

Viswanathan, M., Umashankar, N., Sreekumar, A. and Goreczny, A. (2021), ‘Marketplace literacy as a pathway to a better world: evidence from field experiments in low-access subsistence marketplaces’, *Journal of Marketing* **85**(3), 113–129.

Wang, Y., Lewis, M. and Singh, V. (2021), 'Investigating the effects of excise taxes, public usage restrictions, and antismoking ads across cigarette brands', *Journal of Marketing* **85**(3), 150–167.

Wanzirah, H., Tusting, L. S., Arinaitwe, E., Katureebe, A., Maxwell, K., Rek, J., Bottomley, C., Staedke, S. G., Kamya, M., Dorsey, G. et al. (2015), 'Mind the gap: house structure and the risk of malaria in uganda', *PLoS One* **10**(1), e0117396.

We Are Social (2023), *Digital Global Overview Report*, We Are Social.
URL: *https://wearesocial.com*

Wernerfelt, N., Tuchman, A., Shapiro, B. and Moakler, R. (2022), 'Estimating the value of offsite data to advertisers on meta', *University of Chicago, Becker Friedman Institute for Economics Working Paper* (114).

WHO (2018), *World Malaria Report 2018*, World Health Organization.

WHO (2019), *World Malaria Report 2019*, World Health Organization.

WHO (2021), *World Malaria Report 2021*, World Health Organization.

World Bank (2022), *World Bank Data*, World Bank, Washington, DC.
URL: *https://data.worldbank.org/indicator/SH.MLR.INCD.P3*

Yadav, K., Dhiman, S., Rabha, B., Saikia, P. and Veer, V. (2014), 'Socio-economic determinants for malaria transmission risk in an endemic primary health centre in assam, india', *Infectious diseases of poverty* **3**(1), 1–8.

Zhang, W., Chintagunta, P. K. and Kalwani, M. U. (2021), 'Social media, influencers, and adoption of an eco-friendly product: Field experiment evidence from rural china', *Journal of Marketing* **85**(3), 10–27.

Zimmerman, R. S., Xiao, Z., Mehrotra, P. and Roy, C. (2016), 'Models of health behavior change', *Introduction to Global Health Promotion* **65**.

Can Facebook Ads Prevent Malaria?
Two Field Experiments in India

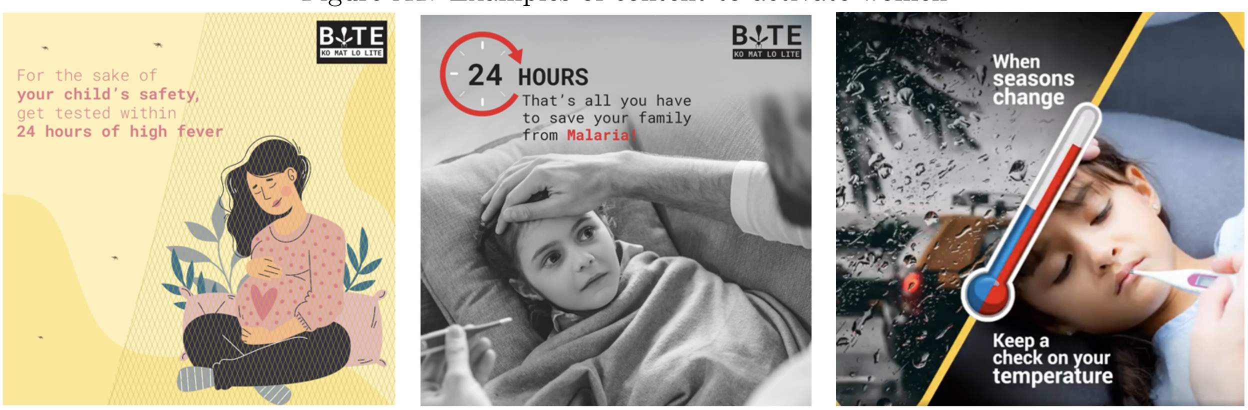
Appendix

A Content Development and Campaign Management

A.1 Developing content for different personas

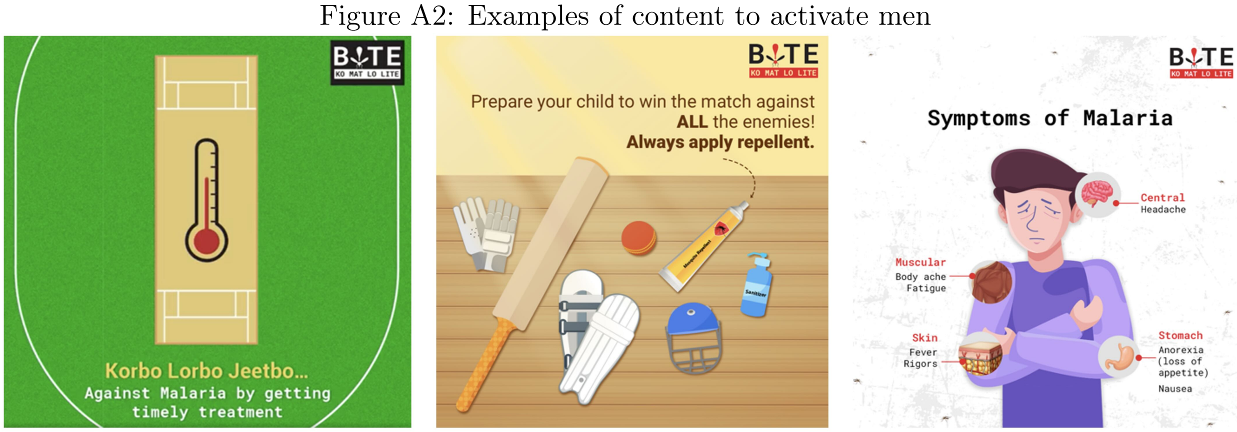
Data insights from public posts revealed that women tended to emphasize family and protection against mosquitoes. Hence, the content development team worked to develop content that women would find engaging; this included content highlighting children and motherhood. Some ads outlined the detrimental effects that malaria can have on pregnant women.

Figure A1: Examples of content to activate women



Men, on the other hand, tended to engage more on non-personal aspects of campaigns, such as government, healthcare, symptoms, treatment, and awareness days. Content was also developed to engage male audiences including content which featured a popular sport cricket.

Figure A2: Examples of content to activate men



Younger audiences, especially in the 18-24 age bucket, were posting about malaria and other mosquito-borne diseases at a lower rate when compared to older people. The MNM team made a conscious effort to use meme-styled content and humor to engage this audience.

Figure A3: Examples of content to activate younger audiences

The data insights demonstrated that older people were disproportionately more likely to post about malaria and other mosquito-borne disease; this made older people important advocates for promoting the campaign's objectives. Thus, content was developed specifically for grandparents encouraging them to have their grandchildren tested if they experience malaria symptoms.

Figure A4: Examples of content to activate older audiences

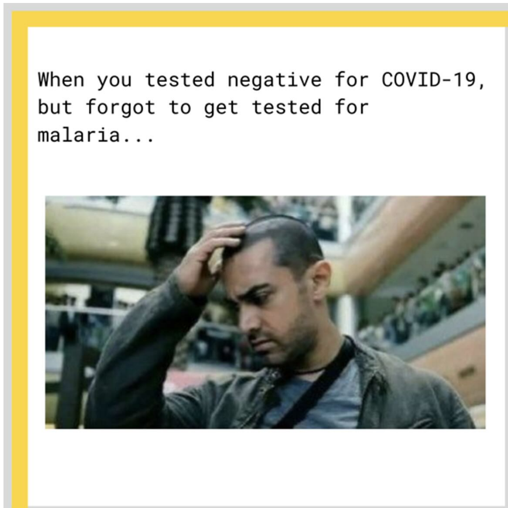
The insights revealed that in rural areas, people were posting about malaria and other mosquito-borne diseases at a lower rate than people in urban areas; this finding underscored the need to produce resonant content for rural audiences since malaria disproportionately occurs in rural areas. Furthermore, less urbanized areas were found to post less about mosquito-borne disease in English; this demonstrated the need to have content in other languages including Hindi and Hin-glish. Moreover, analysis of public posts revealed that rural residents highlighted an appreciation for ASHAs (community health workers).

The campaign occurred during the context of the COVID-19 pandemic and one of the main goals was to limit the rise of malaria fatalities during COVID-19 by encouraging people to engage

Figure A5: Examples of content for rural audiences

in preventative behavior and to seek testing. Moreover, several symptoms of COVID-19, such as fever, are similar to symptoms for malaria; this further underscored the need for people to get tested. Analysis of posts showed that a large number of mosquito-borne disease posts also mentioned COVID-19, with many posts mentioning the continued threat of malaria and other mosquito-borne diseases. Acknowledging the pandemic, the creative team wanted to highlight the continued threat of mosquito-borne disease and developed content encouraging people to get tested for malaria and take protective measures.

Figure A6: Examples of COVID-19-related content



A.2 Campaign dissemination via Facebook and Instagram ads

The dissemination of the content occurred in two phases (or campaigns). For the first campaign, from July to November in 2020, the MNM team used Facebook Ad Manager to create one campaign with the objective of maximizing engagement. In that campaign, the team created six ad sets, each one targeting the demographic of a specific persona. The team limited the geographic targeting by excluding the treatment areas identified by the research team. The creative designed for each persona was placed in the appropriate ad set. Ads were turned off after they reached a frequency over five or noticed a significant drop in ad engagement. The team relied on the Facebook Ad Algorithm to identify the highest-performing ad in each ad set. Figure A7 reports these ads.

During a second campaign in December 2020 and January 2021, the team used Facebook Ad Manager to create one campaign also with the objective of maximizing engagement. In that campaign, the team created six ad sets, each one targeting the demographic of a specific persona. They then filled each ad set with the top-performing ad creatives from the first campaign.

A.3 Facebook and Instagram community management

Instagram and Facebook channels require daily engagement. Malaria No More employed a community manager who oversaw comment threads across both channels. This person used a management guide equipped with pre-written responses for frequently asked questions. The following activities were carried out at least twice a day:

1. Hiding all negative or nonconstructive (spam-like) comments
2. Liking all positive comments
3. Responding to all Direct Messages
4. Inviting all the people who have engaged with the MNM page to also like the page

Figure A7: Top performing content

B Survey Questions

B.1 Longitudinal Survey

a) *Did you sleep under a mosquito net last night?* This question was coded as 1 if the respondent reported to have slept under a bed net the previous night and 0 otherwise. Households without mosquito nets were excluded.

b) *How many family members that live in your house (including yourself) slept under a mosquito net last night?* This question required respondents to enter a number. We used this and family size to construct a new variable reflecting the share of family members who used a bed net the night before the interview. Households without mosquito nets were excluded.

c) *Thinking about yesterday, did you/your family use long-sleeve clothing to prevent mosquito bites?* and *Thinking about yesterday, did you/ your family use body/wall sprays to prevent mosquito bites?* These two questions were coded as 1 if the respondent answered affirmatively, and 0 otherwise.

d) *Have you/someone in your family had a fever (100.4°F / 38°C or above) / malaria in the last two weeks?* This question was asked separately for fever and malaria, and was coded as 1 if the respondent answered affirmatively, and 0 otherwise.

e) Conditional on respondent reporting a fever: *Did you seek medical help?* This question was coded as 1 if answered affirmatively, and 0 otherwise. If yes: *How long after the onset of the fever?* We coded as 1 those reporting help seeking within 24 hours.

f) Conditional on respondent reporting malaria: *Were you or your family members tested for Malaria?* This question was coded as 1 if answered affirmatively, and 0 otherwise.

B.2 Cross-sectional Survey

a) *Have you/someone in your family had Malaria since last August?* This question was coded as 1 if answered affirmatively, and 0 otherwise.

b) Conditional on respondent reporting malaria: *How many family members (including yourself) have had Malaria since last August?* This question required respondents to enter a number. We used this and the family size to construct a new variable reflecting the share of family members who have had malaria since August 2020.

c) *How worried are you that you/someone in your family will get COVID-19/Malaria in the next year?* This question was asked separately for COVID-19 and malaria. Respondents could choose among four options, from "Not at all worried" to "Extremely worried".

d) *Would you consider buying a mosquito net for your house?* This question was asked only to respondents without mosquito nets, and was coded as 1 if answered affirmatively, and 0 otherwise.

e) *Has your mosquito net been re-treated with an insecticide since last August?* This question was asked only to respondents with mosquito nets, and was coded as 1 if answered affirmatively, and 0 otherwise.

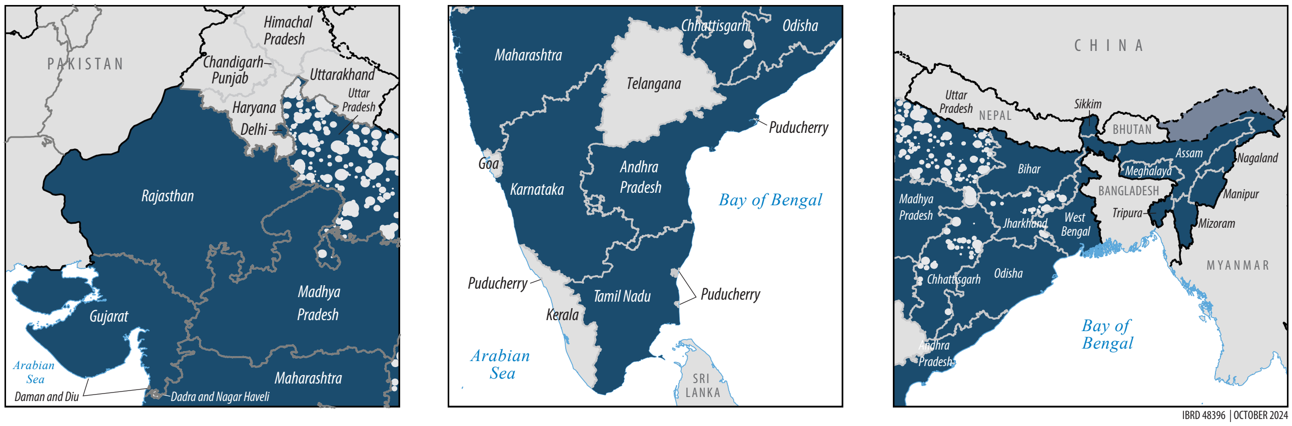
B.3 Individual-level Study

a) *Did you sleep under a mosquito net last night?* This question was coded as 1 if the respondent reported to have slept under a bed net the previous night and 0 otherwise. Households without mosquito nets were excluded.

b) *What would you do if tomorrow you or a family member got a fever (100.4°F / 38°C or above) and there was a doctor or a healthcare facility near your home?* This question was coded as 1 if the respondent would seek treatment right away or within 24 hours of the onset of a fever.

C Figures

Figure C1: Potential reach of the MNM campaign

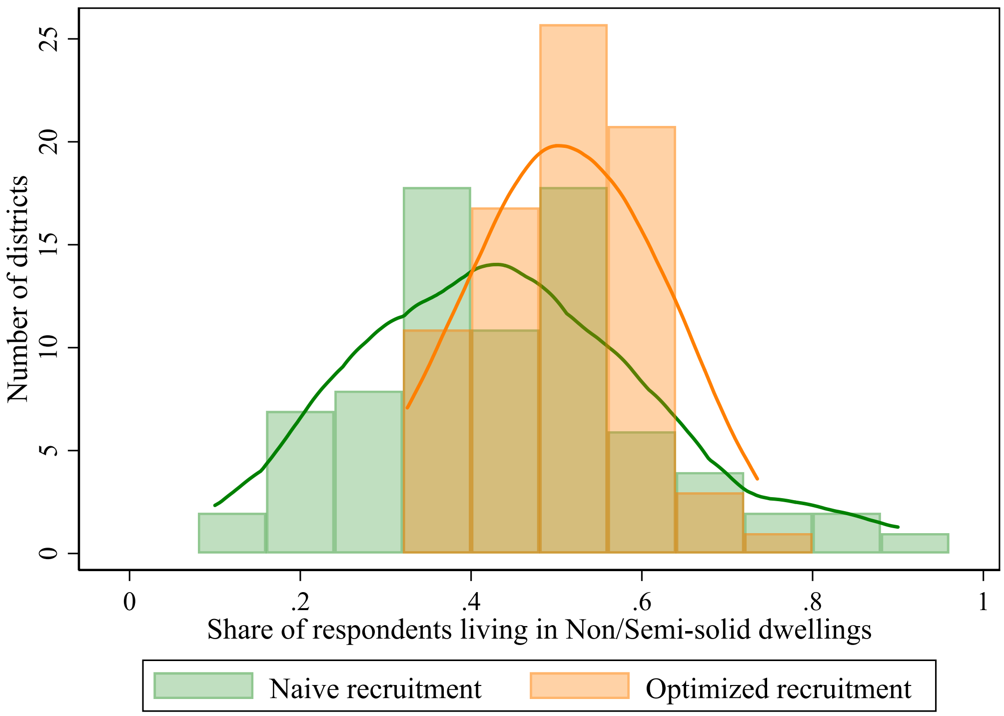


This map was produced by the Cartography Unit of the World Bank Group. The boundaries, colors, denominations and any other information shown on this map do not imply, on the part of the World Bank Group, any judgment on the legal status of any territory, or any endorsement or acceptance of such boundaries.

Notes: The MNM campaign targeted individuals residing in the 22 states in blue. These are: Manipur, Assam, Andhra Pradesh, Delhi, Gujarat, Maharashtra, Meghalaya, Karnataka, Nagaland, Odisha, Rajasthan, Tamil Nadu, Tripura, West Bengal, Sikkim, Arunachal Pradesh, Mizoram, Bihar, Madhya Pradesh, Uttar Pradesh, Chhattisgarh, Jharkhand. The last three states were part of the evaluation exercise and the gray circles in the map represent the control districts, which were excluded from the targeting of the campaign.

Figure C2: Examples of Recruitment ads

Figure C3: Survey-sample optimization



D Tables

Table D1: Pairwise correlations between malaria incidence and household characteristics

	(1) Member had malaria last 5 years	(2) Member had malaria last 2 weeks
Solid house	-0.056***	-0.036***
	(0.0000)	(0.0016)
Household size	0.048***	0.044***
	(0.0000)	(0.0001)
Distance to medical center	0.043***	0.036***
	(0.0002)	(0.0016)
Education	0.045***	-0.001
	(0.0001)	(0.9621)
Backward caste	0.043***	0.027**
	(0.0002)	(0.0190)
Hindu	0.031***	-0.007
	(0.0072)	(0.5690)
Observations	7548	7548

The table reports the Pearson correlation coefficients between malaria incidence and household characteristics. Baseline responses collected from July 23 to August 19, 2020 are included in the sample. *p-values* in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D2: Household-level predictors of malaria incidence

	Member had malaria last 5 years	Member had malaria last 2 weeks	Solid House
	(1)	(2)	(3)
Solid house	-0.022***	-0.004**	
	(0.005)	(0.001)	
Household size	0.017***	0.005**	-0.015
	(0.005)	(0.002)	(0.013)
Distance to medical center	0.012**	0.003**	-0.173***
	(0.005)	(0.002)	(0.010)
Education	0.021***	0.001	0.162***
	(0.005)	(0.002)	(0.011)
Backward caste	0.013***	0.003*	-0.085***
	(0.004)	(0.001)	(0.014)
Hindu	0.011**	-0.001	0.003
	(0.005)	(0.001)	(0.011)
Constant	0.184***	0.016***	-0.000
	(0.011)	(0.001)	(0.019)
Observations	7548	7548	7548
Adj. R-squared	0.011	0.004	0.068
Mean of Y	0.184	0.016	
Mean of Y: Solid house	0.163	0.012	
Mean of Y: Non/Semi-solid house	0.206	0.021	

The table reports the OLS coefficients from a multivariate regression of malaria incidence on standardized household characteristics (columns 1-2) and of dwelling type on the remaining household characteristics (column 3). Baseline responses collected from July 23 to August 19, 2020 are included in the sample. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D3: Longitudinal Sample - Baseline covariates and outcomes by experimental condition

	Treatment *Mean* (1)	Control *Mean* (2)	Normalized Difference (3)	*T-test (p-value)* (4)
Panel A: Covariates				
N. of respondents	2524	2681		
N. of total survey waves per respondent	3.037	2.991	0.015	0.658
Female	0.084	0.088	-0.010	0.639
N. of HH members	6.540	6.697	-0.031	0.326
Solid house	0.469	0.455	0.020	0.517
Non/Semi-solid house	0.531	0.545	-0.020	0.517
Age	27.266	26.974	0.024	0.372
Has completed university	0.594	0.598	-0.005	0.860
Has completed secondary school	0.273	0.266	0.011	0.664
Has completed primary school	0.097	0.103	-0.014	0.563
Unemployed	0.485	0.502	-0.024	0.256
Student	0.195	0.181	0.025	0.250
OBC caste	0.372	0.390	-0.026	0.319
SC/Dalit caste	0.121	0.130	-0.019	0.448
ST caste	0.064	0.052	0.036	0.502
General caste	0.399	0.385	0.021	0.496
Medical center distance $<$ 15min	0.275	0.270	0.008	0.780
Medical center distance 15-30min	0.266	0.267	-0.001	0.974
Medical center distance 30-60min	0.211	0.217	-0.010	0.640
Has mosquito net	0.761	0.778	-0.028	0.522
Has air conditioning	0.091	0.076	0.038	0.114
Pregnant woman in house	0.133	0.128	0.011	0.662
Hindu	0.843	0.865	-0.045	0.230
Muslim	0.116	0.099	0.038	0.374
Any member had malaria in last 5 years	0.206	0.209	-0.004	0.935
Worried/Very worried about malaria	0.357	0.370	-0.020	0.307
Worried/Very worried about COVID-19	0.449	0.448	0.001	0.974
Panel B: Outcomes				
Used bednet the previous night	0.710	0.694	0.024	0.515
Share of HH members who used bednet the previous night	0.708	0.685	0.045	0.221
Used long-sleeve clothing the previous day	0.561	0.581	-0.029	0.387
Used body/wall sprays the previous day	0.351	0.369	-0.027	0.287
Any member had malaria in last 2 weeks	0.019	0.019	-0.002	0.914
Any member had a fever in last 2 weeks	0.063	0.067	-0.011	0.666
Any member tested for malaria	0.596	0.569	0.038	0.785
Any member sought medical help if fever	0.792	0.849	-0.105	0.224

The table reports covariates and outcomes measured at baseline for individuals' and households' in the experience sample, by experimental condition. The normalized difference is the difference between the two means divided by the square root of the sum of the variances of the two variables.

Table D4: Cross-sectional Sample - Covariates by experimental condition

	Treatment *Mean* (1)	Control *Mean* (2)	Normalized Difference (3)	*T-test* *(p-value)* (4)
Covariates				
N. of respondents	1476	1576		
Female	0.121	0.115	0.013	0.669
N. of HH members	6.524	6.442	0.016	0.668
Solid house	0.425	0.437	-0.016	0.633
Non/Semi-solid house	0.575	0.563	0.016	0.633
Age	27.238	27.699	-0.034	0.300
Has completed university	0.566	0.560	0.009	0.798
Has completed secondary school	0.255	0.278	-0.036	0.195
Has completed primary school	0.121	0.114	0.015	0.505
Unemployed	0.545	0.496	0.071	0.008
Student	0.180	0.201	-0.037	0.204
OBC caste	0.341	0.345	-0.006	0.869
SC/Dalit caste	0.122	0.121	0.002	0.955
ST caste	0.060	0.055	0.015	0.715
General caste	0.438	0.438	-0.000	0.995
Medical center distance $<$ 15min	0.307	0.280	0.041	0.132
Medical center distance 15-30min	0.280	0.282	-0.002	0.934
Medical center distance 30-60min	0.192	0.199	-0.013	0.589
Has mosquito net	0.701	0.711	-0.016	0.763
Mosquito nets per HH member	0.499	0.511	-0.015	0.739
Has air conditioning	0.100	0.096	0.007	0.788
Pregnant woman in house	0.158	0.133	0.049	0.085
Hindu	0.862	0.865	-0.005	0.902
Muslim	0.102	0.101	0.003	0.937

The table reports covariates for individuals' and households' in the cross-sectional sample by experimental condition. The normalized difference is the difference between the two means divided by the square root of the sum of the variances of the two variables.

Table D5: Longitudinal Sample - Baseline covariates and outcomes by dwelling type

	Solid House *Mean* (1)	Non/Semi-solid House *Mean* (2)	Normalized Difference (3)	*T-test* *(p-value)* (4)
Panel A: Covariates				
N. of respondents	2402	2803		
T	0.493	0.478	0.020	0.517
N. of total survey waves per respondent	3.173	2.877	0.100	0.000
Female	0.099	0.074	0.063	0.001
N. of HH members	6.537	6.694	-0.031	0.160
Age	28.567	25.871	0.223	0.000
Has completed university	0.695	0.511	0.271	0.000
Has completed secondary school	0.218	0.314	-0.154	0.000
Has completed primary school	0.067	0.127	-0.144	0.000
Unemployed	0.365	0.603	-0.347	0.000
Student	0.204	0.173	0.055	0.007
OBC caste	0.380	0.382	-0.002	0.908
SC/Dalit caste	0.101	0.147	-0.098	0.000
ST caste	0.037	0.076	-0.122	0.000
General caste	0.454	0.338	0.169	0.000
Medical center distance $<$ 15min	0.338	0.216	0.196	0.000
Medical center distance 15-30min	0.290	0.247	0.069	0.002
Medical center distance 30-60min	0.195	0.230	-0.060	0.001
Has mosquito net	0.761	0.777	-0.027	0.204
Has air conditioning	0.127	0.046	0.204	0.000
Pregnant woman in house	0.107	0.151	-0.094	0.000
Hindu	0.862	0.848	0.029	0.218
Muslim	0.102	0.112	-0.023	0.318
Any member had malaria in last 5 years	0.181	0.230	-0.086	0.000
Worried/Very worried about malaria	0.330	0.392	-0.093	0.000
Worried/Very worried about COVID-19	0.429	0.466	-0.052	0.010
Panel B: Outcomes				
Used bednet the previous night	0.681	0.719	-0.059	0.021
Share HH members using bednet previous night	0.682	0.709	-0.053	0.026
Used long-sleeve clothing the previous day	0.531	0.606	-0.108	0.000
Used body/wall sprays the previous day	0.389	0.335	0.079	0.001
Any member had malaria in last 2 weeks	0.015	0.022	-0.041	0.037
Any member had a fever in last 2 weeks	0.062	0.067	-0.013	0.496
Any member tested for malaria	0.543	0.603	-0.085	0.540
Any member sought medical help if fever	0.827	0.819	0.014	0.865

The table reports covariates and outcomes measured at baseline for individuals' and households' in the experience sample, by dwelling type. The normalized difference is the difference between the two means divided by the square root of the sum of the variances of the two variables.

Table D6: Cross-sectional Sample - Covariates by dwelling type

	Solid House *Mean* (1)	Non/Semi-solid House *Mean* (2)	Normalized Difference (3)	*T-test* *(p-value)* (4)
Covariates				
N. of respondents	1316	1736		
Female	0.144	0.098	0.099	0.669
N. of HH members	6.221	6.679	-0.091	0.668
Age	29.454	25.977	0.258	0.300
Has completed university	0.647	0.500	0.212	0.798
Has completed secondary school	0.231	0.294	-0.102	0.195
Has completed primary school	0.092	0.137	-0.100	0.505
Unemployed	0.404	0.607	-0.293	0.008
Student	0.195	0.187	0.015	0.204
OBC caste	0.325	0.356	-0.046	0.869
SC/Dalit caste	0.083	0.151	-0.151	0.955
ST caste	0.042	0.069	-0.083	0.715
General caste	0.524	0.372	0.219	0.995
Medical center distance < 15min	0.363	0.240	0.191	0.132
Medical center distance 15-30min	0.306	0.262	0.069	0.934
Medical center distance 30-60min	0.176	0.211	-0.063	0.589
Has mosquito net	0.694	0.715	-0.033	0.763
Mosquito nets per HH member	0.494	0.514	-0.026	0.739
Has air conditioning	0.144	0.063	0.191	0.788
Pregnant woman in house	0.110	0.172	-0.125	0.085
Hindu	0.876	0.854	0.045	0.902
Muslim	0.100	0.103	-0.008	0.937

The table reports individuals' and households' characteristics of the cros-sectional sample, by dwelling type. The normalized difference is the difference between the two means divided by the square root of the sum of the variances of the two variables.

Table D7: Administrative Data - Mean outcomes across subdistricts

	All subdistricts			Subdistricts with >50% population within circles		
	All health facilities	Urban	Rural	All health facilities	Urban	Rural
N. of cases	9.63	3.65	9.37	10.05	4.76	8.93
N. of cases per M people	51.23	7.33	53.80	20.89	9.45	18.87
N. of tests per M people	3278.63	1050.34	3267.20	2689.72	1221.98	2425.33
Test positivity rate (%)	1.50	1.45	1.43	1.24	1.71	1.10
Population (Thousand)	550.35	905.92	509.51	716.82	960.55	656.96
N. of subdistricts	395	111	392	119	65	116

Notes: The table reports the means of the variables across the different subsamples. Only subdistricts with non-missing entries are included. Subdistricts where the number of positive malaria cases was larger than the number of malaria tests conducted are dropped. Not all the subdistricts contain both urban and rural health facilities. Urban health facilities are present only in a subset of subdistricts.

Table D8: Wearing long sleeves and using sprays (during the campaign)

	Share of times any member used long sleeves			Share of times any member used sprays		
	(1)	(2)	(3)	(4)	(5)	(6)
T	-0.005	0.001	-0.001	-0.020	-0.035	-0.017
	(0.019)	(0.018)	(0.023)	(0.023)	(0.022)	(0.029)
T × Solid house			0.004			-0.037
			(0.028)			(0.046)
Solid house	-0.040***	-0.033**	-0.035	0.031	0.033	0.050
	(0.015)	(0.016)	(0.022)	(0.023)	(0.025)	(0.038)
Individual controls		✓	✓		✓	✓
Cluster controls		✓	✓		✓	✓
Observations	5174	5174	5174	2351	2351	2351
Adj. R-squared	0.014	0.049	0.049	0.002	0.015	0.015
Mean in control	0.642	0.642	0.642	0.410	0.419	0.410
SD in control	0.404	0.404	0.404	0.485	0.489	0.485

Notes: Each observation is an individual. Responses collected from September 1, 2020 to January 31, 2021 are included and are weighted by the number of survey waves the individual completed in this period. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and air conditioner, and 5-year malaria incidence at baseline. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D9: Incidence of malaria and testing behaviors (during the campaign)

	Any member has had malaria				Any member was tested for malaria			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full	Full	Full	Solid house	Full	Full	Full	Solid house
T	-0.004	-0.006	-0.019	0.009	0.003	-0.026	-0.048	-0.006
	(0.010)	(0.007)	(0.012)	(0.010)	(0.061)	(0.048)	(0.066)	(0.088)
T × Solid house			0.027				0.060	
			(0.018)				(0.102)	
Solid house		-0.014	-0.027*			0.110*	0.079	
		(0.009)	(0.014)			(0.057)	(0.072)	
Individual controls		✓	✓	✓		✓	✓	✓
Cluster controls		✓	✓	✓		✓	✓	✓
Observations	5141	5141	5141	2373	324	324	324	120
Adj. R-squared	0.003	0.035	0.035	0.031	-0.008	0.104	0.102	0.116
Mean in control	0.062	0.062	0.062	0.046	0.679	0.679	0.679	0.704
SD in control	0.241	0.241	0.241	0.209	0.468	0.468	0.468	0.461

Each observation is an individual. Responses collected from September 1, 2020 to January 31, 2021 are included and are weighted by the number of survey waves the individual completed in this period. In columns (1-4), the dependent variable takes value 1 if the respondent/family member has had malaria at least once in the period under consideration. In columns (5-8), the dependent variable takes value 1 if the respondent/family member with malaria was tested at least once in the period under consideration. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and air conditioner, and 5-year malaria incidence at baseline. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D10: Bed net ownership, willingness to purchase and re-treatment with insecticide (after the campaign)

	1 if respondent has a net			1 if respondent would buy net			1 if net retreated with insecticide		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All	Solid house	Non-solid house	All	Solid house	Non-solid house	All	Solid house	Non-solid house
T	-0.008	-0.003	-0.010	-0.007	-0.042	0.017	0.046*	0.067**	0.026
	(0.027)	(0.033)	(0.031)	(0.031)	(0.041)	(0.048)	(0.025)	(0.033)	(0.031)
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cluster controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	3052	1316	1736	844	378	466	2017	855	1162
Adj. R-squared	0.054	0.068	0.046	0.130	0.052	0.064	0.019	0.014	0.021
Mean in control	0.711	0.686	0.730	0.389	0.269	0.498	0.409	0.394	0.420
SD in control	0.454	0.464	0.444	0.488	0.444	0.501	0.492	0.489	0.494

Each observation is an individual. Responses collected from January 7, 2021 to March 31, 2021 are included. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and of an air conditioner. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D11: Concerns about getting malaria next year (after the campaign)

Y=1 if respondent is extremely worried about any family member getting malaria

	Solid house			Non-solid house		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Has bednet	Has no bednet	All	Has bednet	Has no bednet
T	0.034	0.031	0.067*	0.026	0.024	0.013
	(0.024)	(0.031)	(0.037)	(0.020)	(0.023)	(0.031)
Individual controls	✓	✓	✓	✓	✓	✓
Cluster controls	✓	✓	✓	✓	✓	✓
Observations	1316	913	403	1736	1241	495
Adj. R-squared	0.013	0.018	0.026	0.012	0.018	0.014
Mean in control	0.240	0.252	0.213	0.262	0.265	0.254
SD in control	0.427	0.435	0.410	0.440	0.442	0.436

Each observation is an individual. Responses collected from January 7, 2021 to March 31, 2021 are included. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and of an air conditioner. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D12: Logit estimates: Incidence of fever and timely help-seeking (during the campaign)

	Any member has had a fever				Any member sought for help within 24h			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Full	Full	Full	Solid house	Full	Full	Full	Solid house
T	0.045	0.049	0.095	0.041	0.044	0.115	-0.195	0.950***
	(0.102)	(0.088)	(0.110)	(0.132)	(0.211)	(0.192)	(0.286)	(0.302)
T × Solid house			-0.101				0.864*	
			(0.182)				(0.444)	
Solid house		-0.108	-0.058			0.382	-0.061	
		(0.101)	(0.129)			(0.277)	(0.354)	
Individual controls		✓	✓	✓		✓	✓	✓
Cluster controls		✓	✓	✓		✓	✓	✓
Observations	5158	5158	5158	2381	724	724	724	301
Pseudo R-squared	0.003	0.039	0.039	0.050	0.004	0.086	0.092	0.162
Marginal effect	0.007	0.008	0.015	0.006	0.006	0.015	-0.025	0.101
Mean in control	0.156	0.156	0.156	0.145	0.797	0.797	0.797	0.808
SD in control	0.363	0.363	0.363	0.352	0.403	0.403	0.403	0.395

Each observation is an individual. Responses collected from September 1, 2020 to January 31, 2021 are included and are weighted by the number of survey waves the individual completed in this period. In columns (1-4), the dependent variable takes value 1 if the respondent/family member has had a fever (100.4°F / 38°C or above) at least once in the period under consideration. In columns (5-8), the dependent variable takes value 1 if the respondent/family member with fever has sought for help within 24 hours at least once in the period under consideration. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and air conditioner, and 5-year malaria incidence at baseline. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D13: Logit estimates: Incidence of malaria and testing behaviors (during the campaign)

	Any member has had malaria				Any member was tested for malaria			
	(1) Full	(2) Full	(3) Full	(4) Solid house	(5) Full	(6) Full	(7) Full	(8) Solid house
T	-0.061	-0.078	-0.221	0.202	0.017	-0.202	-0.309	1.195
	(0.146)	(0.100)	(0.146)	(0.181)	(0.294)	(0.272)	(0.362)	(0.768)
T × Solid house			0.354				0.317	
			(0.262)				(0.593)	
Solid house		-0.210	-0.383*			0.756**	0.598	
		(0.142)	(0.208)			(0.356)	(0.435)	
Individual controls		✓	✓	✓		✓	✓	✓
Cluster controls		✓	✓	✓		✓	✓	✓
Observations	5141	5141	5141	2373	324	324	324	120
Pseudo R-squared	0.008	0.071	0.072	0.088	0.001	0.191	0.192	0.517
Marginal effect	-0.004	-0.005	-0.015	0.011	0.003	-0.033	-0.050	0.103
Mean in control	0.062	0.062	0.062	0.046	0.679	0.679	0.679	0.704
SD in control	0.241	0.241	0.241	0.209	0.468	0.468	0.468	0.461

Each observation is an individual. Responses collected from September 1, 2020 to January 31, 2021 are included and are weighted by the number of survey waves the individual completed in this period. In columns (1-4), the dependent variable takes value 1 if the respondent/family member has had malaria at least once in the period under consideration. In columns (5-8), the dependent variable takes value 1 if the respondent/family member with malaria was tested at least once in the period under consideration. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and air conditioner, and 5-year malaria incidence at baseline. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D14: Logit estimates: Incidence of malaria (since campaign launch)

	1 if any HH member had malaria				
	(1) Full	(2) Full	(3) Full	(4) Solid house	(5) Non-solid house
T	0.053	0.025	0.266	-0.616**	0.286
	(0.196)	(0.152)	(0.192)	(0.289)	(0.188)
T × Solid house			-0.743**		
			(0.357)		
Solid house		-0.378*	-0.041		
		(0.205)	(0.273)		
Individual controls		✓	✓	✓	✓
Cluster controls		✓	✓	✓	✓
Observations	3052	3052	3052	1316	1736
Pseudo R-squared	0.001	0.079	0.082	0.112	0.105
Marginal effect	0.003	0.001	0.014	-0.025	0.018
Mean in control	0.058	0.058	0.058	0.055	0.061
SD in control	0.235	0.235	0.235	0.229	0.239

Each observation is an individual. Responses collected from January 7, 2021 to March 31, 2021 are included. The dependent variable takes value 1 if the respondent/family member has had malaria since August 2020. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and of an air conditioner. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D15: Placebo regressions: Sleeping under bed net at baseline (before the campaign)

	Share of respondents who used bednet				Share of HH members who used bednet			
	(1) Full	(2) Full	(3) Full	(4) Solid house	(5) Full	(6) Full	(7) Full	(8) Solid house
T	0.010	0.013	0.013	0.015	0.004	0.009	0.004	0.014
	(0.019)	(0.018)	(0.021)	(0.022)	(0.017)	(0.016)	(0.019)	(0.021)
T × Solid house			-0.000				0.010	
			(0.023)				(0.024)	
Solid house		-0.039^{***}	-0.039^{**}			-0.049^{***}	-0.054^{***}	
		(0.013)	(0.017)			(0.013)	(0.017)	
Individual controls		✓	✓	✓		✓	✓	✓
Cluster controls		✓	✓	✓		✓	✓	✓
Observations	5544	5544	5544	2790	5457	5457	5457	2762
Adj. R-squared	0.002	0.015	0.015	0.013	0.002	0.022	0.022	0.020
Mean in control	0.690	0.690	0.690	0.662	0.684	0.684	0.684	0.654
SD in control	0.463	0.463	0.463	0.473	0.374	0.374	0.374	0.390

Notes: Each observation is an individual. Responses collected in the baseline survey from July 1, 2020 to August 19, 2021 are included. Only households with at least one bednet are included. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of an air conditioner and 5-year malaria incidence at baseline. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D16: Concerns about getting COVID-19 next year (after the campaign)

Y=1 if respondent is extremely worried about any family member getting COVID-19

	Solid house			Non-solid house		
	(1) All	(2) Has bednet	(3) Has no bednet	(4) All	(5) Has bednet	(6) Has no bednet
T	-0.003	-0.012	0.028	0.005	-0.002	0.030
	(0.024)	(0.030)	(0.043)	(0.022)	(0.029)	(0.036)
Individual controls	✓	✓	✓	✓	✓	✓
Cluster controls	✓	✓	✓	✓	✓	✓
Observations	1316	913	403	1734	1240	494
Adj. R-squared	0.007	0.004	0.060	-0.001	-0.006	0.005
Mean in control	0.318	0.339	0.273	0.317	0.323	0.300
SD in control	0.466	0.474	0.447	0.465	0.468	0.459

Each observation is an individual. Responses collected from January 7, 2021 to March 31, 2021 are included. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and of an air conditioner. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D17: Sensitivity of administrative estimates to different population thresholds

	Overall incidence				
	>10%	>20%	>30%	>40%	>50%
Treated × Post	-2.87	2.24	1.88	2.37	-4.68
	(7.94)	(7.62)	(7.11)	(8.49)	(5.37)
Observations	3337	2784	2328	1812	1498
Adj. R-squared	0.292	0.274	0.194	0.218	0.187
	Urban incidence				
	>10%	>20%	>30%	>40%	>50%
Treated × Post	-6.64**	-4.82*	-4.00	-5.79**	-6.21**
	(2.64)	(2.54)	(2.60)	(2.85)	(2.86)
Observations	1199	1155	1072	887	791
Adj. R-squared	0.175	0.423	0.449	0.458	0.461
	Rural incidence				
	>10%	>20%	>30%	>40%	>50%
Treated × Post	-0.35	6.30	5.96	6.58	-2.24
	(8.81)	(8.81)	(8.39)	(9.44)	(5.79)
Observations	3102	2563	2131	1641	1338
Adj. R-squared	0.297	0.262	0.169	0.187	0.148
Mean incidence at baseline	26.94	24.29	19.22	21.26	20.89

Notes: Each column restricts the attention to a subsample of subdistricts whose share of the population living within our study circles is above the reported threshold. All controls and fixed effects of Table 5 are included here. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D18: Malaria monthly incidence (N. of cases)

	log(N. of cases)			Poisson(N. of cases)		
	(1) All	(2) Urban	(3) Rural	(4) All	(5) Urban	(6) Rural
Treated × Post	-0.125	-0.296**	-0.048	0.304	-1.601***	0.905
	(0.110)	(0.145)	(0.094)	(0.446)	(0.499)	(0.666)
Subdistrict FE	✓	✓	✓			
District FE				✓	✓	✓
Month and year FE	✓	✓	✓	✓	✓	✓
Subdistrict controls	✓	✓	✓	✓	✓	✓
District controls	✓	✓	✓	✓	✓	✓
Observations	1498	791	1338	1498	791	1338
Adj. R-squared	0.535	0.572	0.492			
Log-likelihood				-15318.3	-2439.6	-12498.1
Mean N. of cases at baseline	10.049	10.049	10.049	10.049	10.049	10.049

Notes: Each observation is a subdistrict-month-year. The panel spans April 2020 to May 2021. Post takes value 1 after August 2020, i.e. when the Facebook campaign started. Subdistrict controls include: subdistrict population, area, share of urban population, elevation, terrain ruggedness and distance to the state capital, all interacted with Post. District controls include total population from the Census as well as the shares of baseline survey respondents living in kutcha dwellings, with university degree, sleeping under mosquito nets, and reporting any malaria cases in the last 5 years and 2 weeks, all interacted with Post. All regressions include state indicators interacted with Post. Standard errors are clustered at the district level. Data come from the Indian Health Management Information System. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D19: Tests conducted for malaria and positivity rates

	Tests per M people			Positive cases/Tests		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Urban	Rural	All	Urban	Rural
Treated × Post	444.807	265.388	325.588	-0.017**	-0.022*	-0.003
	(338.641)	(232.203)	(461.989)	(0.006)	(0.013)	(0.010)
Subdistrict FE	✓	✓	✓	✓	✓	✓
Month and year FE	✓	✓	✓	✓	✓	✓
Subdistrict controls	✓	✓	✓	✓	✓	✓
District controls	✓	✓	✓	✓	✓	✓
Observations	1498	791	1338	1498	791	1338
Adj. R-squared	0.563	0.350	0.665	0.085	0.080	0.061

Notes: Each observation is a subdistrict-month-year. The panel spans April 2020 to May 2021. Post takes value 1 after August 2020, i.e. when the Facebook campaign started. Subdistrict controls include: subdistrict population (only in columns 4-6), area, share of urban population, elevation, terrain ruggedness and distance to the state capital, all interacted with Post. District controls include total population from the Census as well as the shares of baseline survey rsepondents living in kutcha dwellings, with university degree, sleeping under mosquito nets, and reporting any malaria cases in the last 5 years and 2 weeks, all interacted with Post. All regressions include state indicators interacted with Post. Standard errors are clustered at the district level. Data come from the Indian Health Management Information System. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D20: Logit estimates: Ad recall at endline

	All respondents		Solid house		Non-solid house	
	(1)	(2)	(3)	(4)	(5)	(6)
T	0.266**	0.297**	0.452**	0.483**	0.098	0.047
	(0.124)	(0.120)	(0.223)	(0.233)	(0.199)	(0.205)
Individual controls		✓		✓		✓
Cluster controls		✓		✓		✓
Observations	1492	1492	728	728	764	764
Pseudo R-squared	0.003	0.035	0.009	0.069	0.004	0.059
Marginal effect	0.030	0.032	0.050	0.050	0.011	0.005
Mean in control	0.114	0.114	0.101	0.101	0.126	0.126

Each observation is an individual. Responses collected in the endline survey between December 15, 2020 and January 31, 201 are included. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and air conditioner, and 5-year malaria incidence at baseline. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D21: Ad recall at endline by age and levels of concern

	Solid house				Non-solid house			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Age<26	Age≥26	Concerned	Not concerned	Age<26	Age≥26	Concerned	Not concerned
T	0.159***	-0.030	0.086***	0.000	0.018	0.006	0.026	-0.018
	(0.040)	(0.028)	(0.032)	(0.037)	(0.032)	(0.031)	(0.031)	(0.036)
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Cluster controls	✓	✓	✓	✓	✓	✓	✓	✓
Observations	350	378	385	343	464	300	449	315
Adj. R-squared	0.077	-0.017	0.013	-0.037	-0.006	-0.035	-0.001	-0.017
Mean in control	0.089	0.113	0.087	0.118	0.143	0.097	0.120	0.133

Each observation is an individual. Responses collected in the endline survey between December 15, 2020 and January 31, 201 are included. Concerned individuals are those who at baseline reported being at least somewhat worried that someone in their family will get malaria in the next year. T is the treatment indicator. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. State fixed-effects are included in all specifications. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of bednet and air conditioner, and 5-year malaria incidence at baseline. Cluster controls are reported in Table 1. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D22: Individual-level study - Covariates by experimental condition

	Ind. Treatment *Mean* (1)	Control *Mean* (2)	Normalized Difference (3)	*T-test* *(p-value)* (4)
Covariates				
N. of respondents	771	771		
In treatment district	0.471	0.486	-0.022	0.541
Female	0.256	0.235	0.034	0.344
N. of HH members	6.341	6.721	-0.079	0.029
Solid house	0.543	0.503	0.057	0.114
Semi-solid house	0.309	0.336	-0.041	0.253
Age	30.399	30.026	0.024	0.509
Has completed university	0.652	0.651	0.002	0.957
Has completed secondary school	0.258	0.248	0.017	0.640
Has completed primary school	0.069	0.080	-0.031	0.383
Unemployed	0.503	0.547	-0.062	0.083
Student	0.158	0.156	0.005	0.889
OBC caste	0.340	0.320	0.029	0.417
SC/Dalit caste	0.106	0.105	0.003	0.934
ST caste	0.043	0.053	-0.034	0.341
General caste	0.486	0.488	-0.002	0.959
Medical center distance < 15min	0.291	0.318	-0.042	0.245
Medical center distance 15-30min	0.310	0.267	0.067	0.064
Medical center distance 30-60min	0.195	0.192	0.005	0.897
Has mosquito net	0.693	0.660	0.049	0.174
Has air conditioning	0.137	0.130	0.016	0.654
Pregnant woman in house	0.095	0.096	-0.003	0.931
Hindu	0.881	0.853	0.057	0.115
Muslim	0.084	0.097	-0.032	0.376

The table reports covariates for individuals' and households' characteristics in the individual study sample by experimental condition. The normalized difference is the difference between the two means divided by the square root of the sum of the variances of the two variables.

Table D23: Logit estimates: Sleeping under bed net (individual-level study)

Y=1 if respondent has slept under bed net the previous night

	Any dwelling type			Solid house	Non-Solid house
	(1)	(2)	(3)	(4)	(5)
Individual T	0.359***	0.387***	0.424**	0.345*	0.465**
	(0.136)	(0.141)	(0.208)	(0.192)	(0.218)
Individual T × Solid house			-0.068		
			(0.281)		
Solid house	-0.545***	-0.396***	-0.363*		
	(0.137)	(0.153)	(0.203)		
T	0.092	0.103	0.102	0.266	-0.080
	(0.137)	(0.141)	(0.141)	(0.196)	(0.213)
Individual controls		✓	✓	✓	✓
Observations	1043	1043	1043	513	530
Pseudo R-squared	0.017	0.044	0.044	0.051	0.063
Marginal effect	0.075	0.078	0.086	0.075	0.082
Mean in control	0.654	0.654	0.654	0.595	0.706
SD in control	0.476	0.476	0.476	0.492	0.456

Each observation is an individual. Indivitual T is the individual treatment indicator for the remarketing campaign. T is the indicator of the district-level treatment assignment. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house and ownership of an air conditioner. Heteroskedasticity-robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D24: Logit estimates: Timely treatment seeking (individual-level study)

Y=1 if respondent would seek treatment within 24 hours of the onset of a high fever

	Any dwelling type			Solid house	Non-Solid house
	(1)	(2)	(3)	(4)	(5)
seek_treatment					
Individual T	0.422**	0.403*	0.469	0.372	0.476
	(0.208)	(0.209)	(0.344)	(0.257)	(0.353)
Individual T × Solid house			-0.105		
			(0.427)		
Solid house	-0.499**	-0.505**	-0.462		
	(0.211)	(0.240)	(0.287)		
T	0.094	0.101	0.100	0.003	0.186
	(0.206)	(0.209)	(0.209)	(0.269)	(0.346)
Individual controls		✓	✓	✓	✓
Observations	1542	1542	1542	807	735
Pseudo R-squared	0.013	0.037	0.038	0.042	0.085
Marginal effect	0.026	0.025	0.029	0.027	0.022
Mean in control	0.921	0.921	0.921	0.905	0.937
SD in control	0.270	0.270	0.270	0.294	0.243

Each observation is an individual. Indivitual T is the individual treatment indicator for the remarketing campaign. T is the indicator of the district-level treatment assignment. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of a bednet and of an air conditioner. Heteroskedasticity-robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D25: Sleeping under bed net (individual-level study) by treated and control districts

Y=1 if respondent has slept under bed net the previous night

	Respondents in control districts			Respondents in treated districts		
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Solid house	Non-solid house	All	Solid house	Non-solid house
Individual T	0.104^{***}	0.102^{*}	0.112^{**}	0.065	0.056	0.057
	(0.040)	(0.062)	(0.053)	(0.042)	(0.060)	(0.063)
Solid house	-0.086^{*}			-0.077^{*}		
	(0.044)			(0.045)		
Individual controls	✓	✓	✓	✓	✓	✓
Observations	551	252	299	492	261	231
Adj. R-squared	0.049	0.051	0.039	0.006	0.035	0.001
Mean in control	0.642	0.557	0.702	0.667	0.626	0.711
SD in control	0.480	0.499	0.459	0.472	0.486	0.455

Each observation is an individual. Indivitual T is the individual treatment indicator for the remarketing campaign. Solid house is a binary indicator equal to 1 if the individual lives in a dwelling made of stones, bricks, cement or concrete. Individual controls include gender, age, household size, caste, education, distance to the medical center, employment and student status, religion, presence of pregnant woman in the house, ownership of a bednet and of an air conditioner. Heteroskedasticity-robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$