MONASH University

MONASH BUSINESS SCHOOL

CENTRE FOR HEALTH ECONOMICS

# WORKING PAPERS

## Does Word-of-Mouth Lead to Better Care? Evidence from Odisha, India

Centre for Health Economics, Monash University
no. 2026-08

**Kai Shen Lim, Bijetri Bose, Terence C. Cheng**

**Keywords:** Word-of-mouth; health care markets; limited attention; social learning; India

**JEL Classification:** D12, D83, I11, I15

# Does Word-of-Mouth Lead to Better Care? Evidence from Odisha, India*

Kai Shen Lim[1], Bijetri Bose[2], and Terence C. Cheng[3]

[1]Harvard Kennedy School
[2]Azim Premji University
[3]Monash Business School

25 July 2026

**Abstract**

Word-of-mouth (WOM) recommendations from family and friends can substitute for absent quality certification and weak regulatory oversight. Whether such informal information improves or distorts patient decisions remains unclear. We study this question using original linked household and facility survey data from Odisha, India, covering 3,090 patient visits across 373 villages. Patients who choose providers based on recommendations spend more out-of-pocket on drugs and transportation. Yet, this additional spending does not translate into better outcomes: WOM is not associated with higher perceived care quality, and WOM users are significantly less likely to return to the same provider. To understand these patterns, we estimate an attentive logit model that separately identifies consideration and preference stages of provider choice. We find that WOM narrows attention to a smaller set of providers that are not higher-quality. Counterfactual simulations show that removing WOM-induced attention distortions does not significantly improve the quality of providers patients choose, while raising awareness towards other providers yields modest gains. These findings suggest that informal networks cannot reliably substitute for formal quality information systems in fragmented healthcare markets.

*JEL classifications*: D12, D83, I11, I15
*Keywords*: Word-of-mouth; health care markets; limited attention; social learning; India

*We gratefully acknowledge the funding from the Bill and Melinda Gates Foundation and the Tata Trust. We are grateful to the Government of Odisha for their support and encouragement for undertaking this assessment of the state's health system and for granting us the necessary permissions for this study. We express our heartfelt thanks to the thousands of people in Odisha who participated in this study and generously responded to our surveys. Without their cooperation, this study would not have been possible. We gratefully acknowledge the consultations with our partner organizations, the Health Systems Transformation Platform and the Indian Institute of Public Health-Bhubaneswar. We thank the team at Oxford Policy Management who collected and cleaned the data for this study. We are also grateful for the support of Winnie Yip, Anuska Kalita, Annie Haakenstaad, Jan Cooper, Elizabeth Osborn, and all other members of the India Health System Project.

# 1 Introduction

How do patients choose providers when quality is unobservable and institutions are weak? In many low- and middle-income countries, formal quality certification is absent, public reporting is limited, and regulatory oversight is minimal (Das and Hammer, 2014; Das et al., 2016). Patients experiencing illness navigate fragmented systems with substantial heterogeneity in costs, treatment appropriateness, and provider competence; none of which are easily observable prior to use. In these settings, word-of-mouth (WOM) recommendations from family, friends, and neighbors may serve as a source of information guiding provider choice.

The conventional view holds that social learning mitigates information frictions and improves welfare. If individuals share experiences and transmit accurate signals about quality, then informal networks can substitute for missing institutions, helping patients avoid low-quality providers and identify effective treatments. Social learning has been shown to reduce uncertainty and improve decisions in settings from agricultural technology adoption (Foster and Rosenzweig, 1995; Conley and Udry, 2010) to health service utilization (Deri, 2005). From this perspective, WOM reduces search costs, corrects misperceptions, and directs demand toward better care. However, an alternative view suggests that social information may transmit biased or incomplete signals, generating information cascades that lead to a suboptimal equilibrium (Banerjee, 1992; Bikhchandani et al., 1998). If recommenders have misaligned incentives, limited experience, or systematically misperceive quality, WOM could steer patients toward more expensive or lower-quality care.

Understanding WOM's role is especially important in low-income settings, where the combination of information asymmetries and weak institutions creates ground for social information to shape behavior. The policy implications are significant: if WOM effectively guides patients toward higher-quality care, then the low capacity of governments in these settings to provide formal quality certification or regulation may be less problematic than commonly assumed, and policies could focus on facilitating peer information exchange. Conversely, if WOM steers patients toward costlier or lower-quality providers, relying on informal networks as a substitute for institutional quality assurance could reduce welfare. This would suggest that investments in formal information systems and quality regulation may be essential.

This paper provides new descriptive evidence on how recommendations influence provider choice, health care utilization, and health cost using rich and uniquely detailed linked primary data from Odisha, India, one of the country's poorest states. We combine an original household survey covering 30,645 individuals across 7,567 households with comprehensive surveys of public secondary and tertiary health care facilities in sampled areas. The household survey records detailed information on illness episodes, provider choice, and the reasons for those choices, including whether providers were selected based on recommendations from relatives or friends, alongside granular measures of out-of-pocket (OOP) expenditure, treatments received, and patient-reported experiences. We link these data, via geospatial coordinates, to facility-level information on quality infrastructure, service availability, and medicine stock-outs.

This data structure allows us to observe both demand-side choices and supply-side characteristics, and to assess how WOM affects the entire care pathway from provider selection, cost drivers and their composition, and patient-assessed care quality. Contrasting with studies using randomized formal information interventions (Björkman and Svensson, 2009; Björkman Nyqvist et al., 2017), our survey approach allows us to identify how informal recommendations through friends and family affect choices. Linked patient-facility data are also uncommon in low-income settings, where administrative records on individual provider choice are typically unavailable.

We document three sets of findings on care-seeking behavior for outpatient care. First, relying on WOM is associated with higher OOP expenditures. Our estimates show that patients who choose providers based on social recommendations spend an additional ₹790 to ₹820, depending on specification (76 to 86 percent of the sample mean). These higher costs are concentrated in spending on medicines (₹262) and transportation (₹199), with smaller contributions from consultation fees and diagnostic tests. This increase in drug spending does not reflect higher prices in public pharmacies. Instead, it reflects greater reliance on private pharmacies, where WOM users spend an additional 72.3 (40.9% of the mean) on medications. These patterns are consistently observed across OLS and coarsened exact matching specifications.

Second, this additional spending does not translate into better outcomes. WOM has no statistically significant association with the number of diagnostic tests received or with the probability that patients rate their care as excellent. More strikingly, WOM users are significantly less likely to report being extremely likely to return to the same provider: a 6 to 7 percentage point decline, depending on specification (roughly 40 percent of the sample mean). These findings suggest that WOM use is associated with costlier treatment pathways without commensurate improvements in perceived quality or satisfaction.

Third, to understand the mechanisms underlying these patterns, we estimate a discrete choice model in which patients may not consider all providers in their choice set. Using the attentive logit framework of Abaluck and Adams-Prassl (2021), we model hospital choice as a two-stage process in which observable facility attributes shape both the probability that a provider enters a patient's consideration set and the probability of choice conditional on consideration. This framework allows us to distinguish limited attention (an inability to attend to all choices available) from preferences using asymmetric responses of demand to facility characteristics (Sims, 2006).

Our results show that WOM acts primarily through the consideration stage rather than through preferences. Exposure to WOM does not systematically alter how patients value quality or availability once a provider is considered. Instead, WOM attenuates responsiveness to objective measures of accessibility, such as distance and operating hours, anchoring attention to a narrower subset of providers. As a result, patients with network exposure may choose from a narrower set of providers without corresponding gains in provider quality. Social networks influence where patients look, but they do not reliably improve the quality of care patients ultimately receive.

Using the estimates from the discrete choice model, we conduct counterfactual simulations to assess whether removing WOM-induced attention distortions, or raising patient awareness more broadly, would improve provider choice. The first exercise asks how much consideration sets would expand if WOM-exposed patients faced no attention distortion. The second simulates a broader awareness intervention, such as a facility directory or outreach program, applied to the full sample. Removing WOM-induced attention distortions expands consideration sets and shifts choices toward hospitals with longer hours, but does not significantly improve the quality of providers patients choose. A broader awareness intervention yields small quality gains, suggesting that limited baseline awareness is the more important friction constraining provider choice.

There are two important limitations to highlight. First, WOM use is self-reported and chosen by patients rather than assigned. Our descriptive estimates address selection on observable characteristics, including anything constant within a patient's own village via village fixed effects, but cannot rule out selection on unobservables that vary within village. We therefore interpret the descriptive estimates as robust conditional associations rather than causal effects. Second, referral status is recorded only for the chosen provider, so the WOM measure is a person-level indicator rather than an alternative-specific one. The choice model identifies how WOM users screen hospital attributes in general, but not the direct inclusion of a specific recommended provider into the consideration set. The results in this paper should be viewed with these two key limitations in mind.

Our findings contribute to several related literatures, but our most direct contribution is the work on information frictions and provider choice in low-income health markets. Audit and standardized-patient studies have documented substantial heterogeneity in provider effort and treatment quality across both public and private providers in India (Das et al., 2016; Mohanan et al., 2015; Das et al., 2012), and field experiments have shown that providing credible information, through community-based monitoring (Björkman and Svensson, 2009; Björkman Nyqvist et al., 2017), on bed-net effectiveness (Dupas and Miguel, 2017), or on treatment efficacy (Ashraf et al., 2013) can shift behavior. We complement this work by studying informal information transmitted through social networks rather than formal interventions, and by linking individual-level WOM exposure to facility-level attributes in a choice model framework. Our finding that WOM is associated with higher out-of-pocket costs without improving perceived care quality is consistent with the broader pattern in this literature that, in the absence of credible quality signals, patient choice does not reliably converge on higher-quality care.

We also extend the attentive logit framework of Abaluck and Adams-Prassl (2021) to a low-income healthcare setting. Their original application was to Medicare prescription drug plan choice, where consumers face a structured exchange with publicly available attributes; we apply the framework to a setting where information is informal, opaque, and largely transmitted through social networks. We also speak to the growing literature on social learning and information diffusion in Indian villages (Banerjee et al., 2013), where WOM has been shown to spread information about microfinance and other novel products through networks with identifiable structural features. Our results suggest that WOM operates differently in healthcare than in some of these contexts: rather than substituting effectively for missing institutions, it appears to direct patients toward costlier treatment pathways without improving care quality, consistent with the cautionary findings of Li and Zhang (2022) on biased advice transmission.

The remainder of the paper proceeds as follows. Section 2 describes the institutional context, data collection, and measurement. Section 3 presents descriptive estimates of the association between WOM and costs, utilization, and patient ratings, using OLS and CEM-matched OLS. Section 4 specifies the choice model and contains the results. Section 5 concludes by discussing our findings, implications for policy.

# 2 Context and Data

Odisha's health system combines a three-tiered public network, consisting of primary care facilities (health and wellness centers (HWCs), sub-centers, and primary health centers (PHCs)), secondary care facilities (community health centers (CHCs) and sub-district and district hospitals), and tertiary care hospitals, with a large and heterogeneous private sector. Public healthcare facilities in the state, funded and run by the central and state's health departments, provide outpatient care at nominal or zero user fees, including clinical consultations, essential medicines, and diagnostic services.[1] Although the public health facilities dispense generic essential medicines free of cost, stock-outs are pervasive: facility surveys conducted during our study period found that, on average, only 81.1 percent of a standard basket of 41 essential medicines were in stock at rural public facilities, and 64 percent at urban facilities (Bose et al., 2024). These shortages, together with staffing gaps averaging 24 to 42 percent of sanctioned positions, create strong incentives for patients to purchase drugs from private pharmacies after a public-sector consultation, a pattern documented in linked patient-facility-pharmacy data from the same setting (Haakenstad et al., 2025). Private pharmacies predominantly stock branded and branded-generic medicines at higher prices than their public-sector equivalents, though systematic price comparisons are not available.

## 2.1 Household and Facility Surveys

The data for this study come from two surveys conducted in Odisha, India, between August 2019 and March 2020. The first is a household survey that collected detailed information on health system encounters for every individual in the household. The

[1]The state-led Niramaya (2015) and Nidan (2018) schemes provide free essential medicines and free diagnostic services across public health facilities.

data records rich visit-level data for each patient encounter, making it particularly well suited to examining the effects of WOM referrals on a wide range of cost, utilization, and patient experience outcomes. The survey includes detailed demographic and economic characteristics, measures of illness severity and diagnoses, and village identifiers. This information collectively allows us to isolate the role of WOM in shaping provider choice, while controlling for observed patient need and local health system characteristics.

The second survey is on public secondary and tertiary healthcare facilities. These surveys were conducted through structured questionnaires administered to officers-in-charge and, where necessary, other senior staff members. Reliable publicly available facility-level data are uncommon in India, hence these surveys provide a novel and unusually detailed dataset on health services offered, staffing, infrastructure, and facility operations. A distinctive feature of these facility surveys is that they can be directly linked to individual patient visits using facility names, addresses, and geographic identifiers reported by patients. This linkage allows us to identify the actual facilities patients visited and to observe the corresponding facility characteristics.

Data collection was conducted in six districts across the state, selected to capture variation in administrative regions and levels of development. Within these districts, 30 blocks (the sub-district administrative unit) were randomly selected using probability-proportional-to-size sampling. Under this sampling design, more populous districts contributed a larger number of blocks to the sample. Institutional Review Board (IRB) approval for this study was obtained from the Harvard Chan School of Public Health (IRB18-1675), an independent IRB (SIGMA) in India, and the Health Research Approval Committee of the Odisha Government.

The household survey collected data from 30,645 individuals across 7,567 households in six districts of the state. Households were selected using a multistage, stratified, probability-proportional-to-size sampling design, with oversampling of households that included individuals with chronic conditions or recent interactions with the health system. Survey weights were constructed based on selection probabilities, and iterative proportional fitting (raking) was used to calibrate the weights to known population totals for the state of Odisha. Additional details on the sampling design and survey implementation are provided in Appendix A.1. The analysis sample used in this study comprises 3,090 individuals, and includes both family and non-family members residing in the sampled households.

For the facility surveys, different data collection designs were used depending on the type of public facilities. A census of all public secondary and tertiary healthcare facilities in the 30 sampled blocks was conducted using government records. The final sample comprised 122 public secondary and tertiary healthcare facilities.

### 2.2 Measures

Survey respondents were asked if they were ill or injured, and had sought treatment or medical advice from a healthcare provider in the 15 days preceding the survey. They were also asked about the reason for the visit, the type of facility or provider visited, and the reasons they chose their providers. Respondents may report multiple reasons for their choice of provider. We define WOM referrals as cases in which respondents cite being "referred by relatives or friends" among the reasons for choosing their provider. This measure captures informal recommendations transmitted through personal networks.

Figure 1 shows that WOM is a meaningful determinant of provider choice: 12% of respondents cite recommendations by relatives or friends among their reasons for choosing a provider. Accessibility and competence dominate stated motives: 71% cite convenient location, 45% provider competence, and 42% convenient hours. WOM is cited about as often as provider friendliness (14%) and more often than drug availability (10%) or low cost (5%). WOM referrals also vary across provider types: it is highest for private hospitals and pharmacies (approximately 14%), public hospitals (12%) and lower for public primary health centres (9%) and AYUSH providers (6%). Taken together, these patterns show that word-of-mouth plays an important role, especially in care decisions involving a hospital or pharmacy encounter.

Respondents were also asked detailed questions about their care experience, including the nature of treatment received, diagnostic tests undertaken, the number of medicines prescribed, and OOP expenditures associated with the visit. OOP spending was recorded separately for physician fees, diagnostic tests, medications, transportation, and other healthcare-related costs. Information on insurance reimbursements was also collected and netted out from total OOP expenditures. Finally, respondents were asked to provide an overall rating of the quality of care received, taking all aspects of the visit into account.

## 3 Descriptive Evidence

We begin by describing differences between individuals who do and do not rely on WOM recommendations. Table 1 reports respondent characteristics by whether provider choice was informed by WOM. Approximately 12% of the sample sought care through a recommendation. Women are significantly less likely to rely on WOM, contrasting with individuals belonging to other backward classes; there is no statistically significant difference for scheduled tribe or scheduled caste status. Reliance on WOM is weakly associated with socioeconomic status: individuals without insurance and those with lower incomes are somewhat more likely to seek recommendations. Individuals with chronic conditions are less likely to rely on WOM, although this association is only marginally significant, and those seeking treatment for heart disease are also less likely to use recommendations. Finally, WOM users are more likely to seek care for acute respiratory infections and localized pain or weakness, and are significantly less likely to report fever as the primary diagnosis.

### 3.1 Word-of-Mouth, Costs, and Patient Outcomes

To assess the effects WOM has on the outcomes of interest, we implemented the following empirical strategy. Let $i$ index individuals, $v$ villages (PSUs), $f$ facility type, and $g$ diagnosis category. Our outcome of interest is $Y_i$, denoting costs, procedures, utilization, or patient-reported outcomes. The treatment variable $WOM_i$ equals one if respondent $i$ cites a referral by relatives or friends among the reasons for choosing a provider, and zero otherwise. We begin with the following OLS specification:

$$Y_i \;=\; \beta \cdot \mathrm{WOM}_i \;+\; \mu_{v(i)} \;+\; \delta_{f(i)} \;+\; \theta_{g(i)} \;+\; \varepsilon_i, \tag{1}$$

where $\mu_{v(i)}$ are village fixed effects, $\delta_{f(i)}$ facility-type fixed effects, and $\theta_{g(i)}$ diagnosis fixed effects. Adjustment for individual characteristics occurs through the matched specification: the CEM estimator restricts comparisons to WOM and non-WOM users matched on age, chronic-disease status, fever, and log income. Standard errors are clustered at the village level.

To assess the effects of WOM on the outcomes of interest, we estimate two complementary specifications. The baseline OLS specification controls for observable characteristics via diagnosis, facility-type and village fixed effects. A matched OLS is also used to restrict comparisons to observably similar WOM and non-WOM users via CEM, improving balance on pre-treatment characteristics. Village fixed effects absorb any time-invariant local characteristic, such as remoteness or the reliability of the public drug supply, that is common to all patients in the same village.

Table 2 presents estimated associations between WOM and total OOP costs. Both specifications yield positive, large, and statistically significant estimates: OLS (₹790; 76% of the sample mean) and matched OLS (₹820; 86%). These estimates show that individuals who chose their providers based on recommendations incurred significantly higher OOP costs.

What types of care account for the higher OOP spending? Figure 2 decomposes the OLS estimates across expenditure categories and shows that the increase in OOP costs is concentrated in spending on drugs (₹261.9) and transportation (₹199.2), with smaller contributions from doctor fees (₹81.4) and medical tests (₹61.2). The matched estimates in the same figure are broadly similar. To further unpack the role of drug costs, Table 3 focuses on patients who were prescribed medication. Individuals who chose their provider based on recommendations spent on average ₹297.2 more on drugs (53.3% of the mean) and were prescribed a similar number of drugs. This increase in drug spending does not reflect higher prices in public pharmacies; instead, it reflects greater reliance on private pharmacies, where WOM users spend an additional ₹72.3 (40.9% of the mean) on medications. Together, these results show that WOM use is associated with providers and treatment pathways that involve higher downstream drug and travel costs.

Does higher spending translate into better patient-related outcomes? Table 4 shows that the additional spending associated with WOM does not translate into improved patient-reported outcomes or greater treatment intensity. WOM has no statistically significant association with the probability that patients rate care as excellent; its association with the number of diagnostic tests received, which we perceive as reflecting providers' effort, is not significant in the OLS specification and only marginally significant (at the 10% level) in the matched specification. By contrast, WOM users are significantly less likely to report being extremely likely to return to the same provider in both specifications, indicating lower subsequent satisfaction despite substantially higher expenditures.

One caveat applies to the return-intention result. Patients following a recommendation are plausibly visiting a provider for the first time, whereas non-WOM patients may include repeat visitors with established provider relationships. Because the survey does not record prior use of the chosen provider, we cannot separate lower satisfaction from lower baseline familiarity, and the return-intention estimate should be read as combining both margins.

The preceding analysis pools all provider types. As we will elaborate in Section 4.1, because the choice model is limited to public hospitals and CHCs, a natural question is whether WOM's association with costs holds within this subsample. Table A2 reports OLS estimates of WOM interacted with facility type, both unwinsorized and winsorized at the 99th percentile. The main effect, corresponding to public hospitals, is positive throughout and statistically significant in the winsorized specification (₹482, $p < 0.05$), suggesting that WOM is associated with higher costs among the patients for whom we estimate the choice demand model. We now turn to this model to investigate the mechanisms underlying this effect.

## 4 Word-of-Mouth and Consumer Choice

The descriptive analyses document that WOM is associated with higher costs without improving quality, but they cannot distinguish whether WOM alters costs by changing which providers patients are aware of, by shifting how patients evaluate providers conditional on awareness, or through some combination. To investigate the provider-selection channel, we estimate a choice model of provider choice that separately identifies two stages: (1) which providers enter a patient's consideration set, and (2) how patients evaluate providers conditional on consideration. The distinction matters because the policy implications differ: if WOM primarily affects consideration, then interventions that broaden awareness (e.g., facility directories, community health worker referrals) are more relevant; if WOM alters preferences, then the content of recommendations is what matters.

## 4.1 Data Construction

In low-income settings, facility-level data are uncommon and hence linking objective measures of service availability to individual patient choices is often unfeasible. We address this limitation using a dedicated facility survey that links each outpatient visit to the specific public hospital or community health center (CHC) visited, allowing us to construct individual-level choice sets over facilities within a district.

Facility attributes are collected through a dedicated facility survey administered to the senior-most clinical or administrative officer responsible for facility operations, such as the medical superintendent, hospital manager, or chief medical officer. Distance between each patient and facility is computed using the Haversine formula applied to GPS coordinates recorded independently during the household survey and the facility survey. Outpatient department (OPD) operating hours are reported directly by facility staff. Quality improvement is measured as a binary indicator equal to one if the facility survey respondent reports the presence of any quality improvement activities at the facility. Drug availability is measured by directly auditing medicine stock on the day of the survey against a standardized list of essential medicines: enumerators recorded whether each drug was physically present, and we aggregate these item-level audits into a facility-level measure equal to the fraction of listed drugs available.

The unit of observation is a single outpatient visit by an individual. Markets are defined at the district level as districts are used as administrative units for health planning. The estimation sample includes 1,033 visits to hospitals and CHCs. Of these, 643 (62 percent) were successfully matched to facility-level survey data; the remaining 38 percent could not be linked, primarily because respondents provided insufficient identifying information. After dropping observations with missing facility attributes, the final estimation sample contains 533 individuals.

We construct each individual's choice set within their district market. Because some districts contain many hospitals, evaluating the choice probability over every facility is impractical, so we approximate the feasible set by sampling alternatives (McFadden, 1978). For each visit, the choice set comprises the chosen hospital, the default hospital (the closest facility by Haversine distance), and six additional hospitals drawn uniformly at random from the remaining facilities in the district, yielding choice sets of 2 to 8 alternatives. We sample uniformly so that, under the conditional logit, no sampling correction is required. Visits lacking either a chosen or a default hospital are dropped. The choice of six sampled rivals is assessed in Appendix A7, which re-estimates the model for sampled-rival counts from 4 to 10.

The generalizability of the choice estimates depends on how well the linked sample represents the broader population of hospital users. Table A3 compares linked and unlinked hospital users. The matched sample is less likely to be married, has larger households, lower chronic disease prevalence (24 vs. 35 percent), and better self-reported health (20 vs. 29 percent reporting poor or fair health). Linked patients are also more likely to present with fever and acute respiratory infections and less likely to present with heart disease, childbirth, or cancer. These differences suggest that the choice estimates generalize most directly to younger, healthier patients seeking care for common acute conditions, and may not extend to sicker patients with more complex care needs. Appendix A.4 shows an inverse-probability-weighted bootstrap that adjusts for these linked-sample differences and shows that the qualitative attention-stage findings are similar under this correction.

### 4.2 Choice Model

A standard discrete choice model assumes patients evaluate all available hospitals and select the utility-maximizing option. This assumption is unlikely in Odisha's fragmented health system, where patients have limited access to systematic information on provider cost and quality. Instead, we model hospital choice as a two-stage process: first, a consideration set forms as each hospital either enters or does not enter the patient's awareness; second, the patient maximizes utility over considered hospitals. This nests the standard model as a special case in which all hospitals are considered with probability one.

To implement this framework, we use the hybrid consideration model of Abaluck and Adams-Prassl (2021).[2] This model combines two features: a default option that is always available, and independent consideration probabilities for non-default alternatives. Each patient faces a designated default hospital, which we define as the closest facility by Haversine distance.[3] Each non-default hospital $j$ enters the consideration set independently with a probability that depends on its observable attributes, while the default is always considered. Patients for whom no non-default hospital enters the consideration set visit the default without deliberation. This implies that the attention constant $\gamma_0$ determines how frequently this occurs. This structure captures a behaviorally intuitive idea: in the absence of active search, patients default to the nearest provider, and only those who engage in some form of deliberation, possibly triggered by a referral or prior experience, look beyond the closest option.

**Consideration stage.** Formally, let $a_{ijm} \in \{0, 1\}$ indicate whether non-default hospital $j$ in district $m$ enters the consideration set of individual $i$. Conditional on the patient being attentive, the probability that a non-default hospital is considered is given by

[2] We use 'hybrid' to denote the two-stage structure: default-specific consideration followed by utility maximization, rather than the wake-up probability that appears in some specifications of Abaluck and Adams-Prassl (2021). In our implementation, the attention constant $\gamma_0$ absorbs any baseline propensity to consider non-default alternatives; we do not separately identify the probability of being attentive versus inattentive.

[3] We use Haversine distance rather than road distance because road network data are incomplete in rural Odisha, particularly for smaller routes connecting villages to block-level facilities.

$$\Pr(a_{ijm} = 1 \mid \text{attentive}) = \Lambda\Bigg( \gamma_0 + \gamma_1 \,\text{Distance}_{ij} + \gamma_2 \,\text{Hours}_j \\ + W_i \times \big(\delta_1 \,\text{Distance}_{ij} + \delta_2 \,\text{Hours}_j\big) \Bigg), \tag{2}$$

where $\Lambda(\cdot)$ denotes the logistic CDF, $\gamma_0$ is a constant governing the baseline propensity to consider non-default alternatives, and $W_i$ indicates whether individual $i$ reported using WOM to choose her provider (a person-level indicator, since referral status is observed only for the chosen facility). The parameters $\gamma_1$ and $\gamma_2$ capture how accessibility, measured by distance and operating hours, influences whether a hospital enters consideration. The interactions with $W_i$ allow WOM exposure to alter how strongly patients emphasize these attributes when forming their consideration sets.

A measurement issue warrants discussion. Our WOM indicator is constructed from a post-visit survey question asking respondents their reasons for choosing the facility they visited. Because this question refers to the chosen provider, we observe referral status only for the facility actually visited, not for unchosen alternatives. Consequently, $W_i$ cannot be made alternative-specific: we cannot construct a variable indicating which hospital in the choice set was recommended. Instead, $W_i$ enters as a person-level indicator interacted with hospital attributes, identifying whether WOM users exhibit systematically different attention patterns across all hospitals in their choice set. This captures the narrowing or broadening of attention to hospital characteristics, but cannot directly identify the primary channel through which referrals may operate, namely, placing a specific recommended hospital into the consideration set. The choice model thus identifies the following: conditional on having followed a referral from friends or family members, how does the patient's attention to other hospitals' attributes change?

A related concern is that the survey's response categories could mechanically generate the attention patterns we estimate: if citing WOM precluded citing accessibility, WOM users would appear less responsive to distance and hours by construction. Given that multiple responses are permitted, the data rule this out: 68 percent of WOM users also cite convenient location or convenient hours among their reasons for choosing the provider, compared with 77 percent of non-WOM users. The attention-stage interactions therefore do not simply restate the survey's reason categories.

The hybrid model rests on two key assumptions regarding consideration. First, conditional on observable attributes, whether hospital $j$ enters the consideration set is independent of whether hospital $k$ does. This rules out correlated consideration, for example, a neighbor mentioning two hospitals in the same conversation, causing both to enter awareness simultaneously. In a setting where WOM operates through personal networks, some correlation in awareness is plausible, and we note this as a limitation. Second, consideration probabilities for non-default alternatives depend only on a hospital's own attributes, not on the characteristics of rival hospitals. This is standard in the attentive logit literature and is practically motivated here by the absence of data on the full information environment patients face.

**Utility stage** Conditional on consideration, individual $i$ derives indirect utility from hospital $j$ given by

$$u_{ijm} = \beta_1 \text{Distance}_{ij} + \beta_2 \text{Hours}_j + \beta_3 \text{Drugs}_j + \beta_4 \text{Quality}_j + \beta_5 \text{CHC}_j + \varepsilon_{ijm} \tag{3}$$

where $\varepsilon_{ijm}$ is an i.i.d. Type I extreme value error. This captures preferences over travel costs, medicine availability, quality, and facility tier, among facilities that the patient actually considers. In our preferred specification, OPD hours enters both equations (2) and (3), allowing hours to shape both which hospitals are noticed and how they are evaluated. We discuss the rationale for this specification choice below.

## 4.3 Estimation and Identification

We estimate the model by maximum likelihood using the exact hybrid attentive logit likelihood described in Abaluck and Adams-Prassl (2021), designating the closest hospital as the default option. Standard errors are heteroskedasticity-robust. The estimation sample consists of 533 unique individuals, each contributing one choice occasion; the 3,233 observations reported in the tables are person-by-alternative rows entering the conditional likelihood.[4]

Choice sets are constructed by sampling alternatives within each district market, as described in Section 4.1. For the conditional logit, uniform sampling of alternatives delivers consistent estimates of the full-choice-set parameters (McFadden, 1978). We do not claim that this guarantee extends to the limited-consideration model: to our knowledge, sampling of alternatives has no established consistency result for attentive logit models, whose identification rests on departures from the symmetry that the conditional logit imposes. The model's consideration-set quantities, such as the expected consideration set size, are accordingly defined relative to the sampled set rather than the full district. We therefore treat the number of sampled rivals as a modelling choice and assess its influence directly: Appendix A7 re-estimates the model varying the number of sampled rivals from 4 to 10, and the estimates are stable.

Our preferred specification (column 3 of Table 5) is an attentive logit in which OPD hours enters both the attention and the utility stages, and in which WOM interacts with distance and hours in the attention equation. The rationale for letting hours enter both stages is that hours plausibly affects whether patients notice a hospital (facilities with short or inconvenient hours may never register as options) and how they evaluate it conditional on awareness (longer hours reduce scheduling constraints

[4]Each individual contributes one choice problem (one chosen hospital, one default, and up to six randomly sampled rivals). The 3,233 choice occasions refer to the person-by-alternative observations that enter the conditional likelihood, averaging $3{,}233/533 \approx 6$ alternatives per individual; this is consistent with choice sets of size 2 to 8 across the 533 individuals.

and are directly valued). The WOM interactions in the attention equation identify whether WOM users exhibit systematically different choice behavior. Drug availability and quality improvement appear only in utility; alternative specifications that allow WOM to interact with utility shifters are reported in Appendix Table A4 and yield no indication that WOM alters preferences. Distance enters both equations because it plausibly affects both stages and because the default is defined as the nearest hospital, making cross-stage variation in distance important to the model.

Identification follows from asymmetric responses of choice probabilities to facility attributes. The key insight from Abaluck and Adams-Prassl (2021) is that when all consumers consider all options, cross-attribute derivatives must be symmetric, a discrete-choice analog of the Slutsky conditions. Violations of this symmetry constructively identify consideration probabilities without exclusion restrictions. In our application, the gap between $\frac{\partial \Pr(\text{choose } A)}{\partial \text{hours}_B}$ and $\frac{\partial \Pr(\text{choose } B)}{\partial \text{hours}_A}$ directly identifies how operating hours affect whether hospitals enter the consideration set. The intuition is as follows: if increasing hospital $A$'s operating hours raises the probability of choosing $A$ more than increasing hospital $B$'s hours reduces the probability of choosing $A$, this asymmetry can only arise if hours affects whether hospitals enter the consideration set, not just their utility conditional on consideration.

In our application, distance and hours influence both stages, while drug availability, quality improvements, and facility type enter only utility. Variation in WOM exposure identifies how social networks reshape attention by altering sensitivity to distance and hours. We interpret the resulting estimates as descriptive rather than causal: the model recovers attention and preference parameters under its maintained assumptions but does not address endogeneity of WOM exposure itself. Appendix A.6 reports a descriptive test of this asymmetry. The pooled data do not reject the symmetric-response restriction, but choice sensitivity varies systematically with WOM exposure: WOM users place significantly less weight on default-hospital attributes and more weight on rival attributes, consistent with WOM narrowing the consideration set.

To assess the sensitivity of our estimates to the modelling choices above, we also report two alternative specifications. Column 1 of Table 5 reports a conditional logit assuming full consideration of every hospital in the choice set; comparison with the attentive logit estimates reveals how inferences change when limited attention is accommodated. Column 2 reports a more restrictive attentive logit that imposes the tighter exclusion of OPD hours from utility (hours enters only the attention equation) and excludes the WOM interactions. If hours genuinely enters utility, column 2's attention coefficient on hours will absorb preference-relevant variation, and relaxing the exclusion in the preferred specification will shrink it. We document this pattern in Section 4.4.

## 4.4 Results

### 4.4.1 Choice coefficient estimates

Table 5 reports the estimates from the attentive logit model under alternative specifications and choice set constructions. Column (1) gives a conditional logit benchmark assuming full consideration. Comparing it to the attentive logit specifications reveals how inferences change when limited attention is accommodated. Columns (2) and (3) report attentive logit estimates: a restricted specification without WOM interactions and with hours excluded from utility, and our preferred specification that includes OPD hours in both the attention and utility stages and allows WOM to interact with distance and hours in the attention equation.[5] Panel A presents attention (consideration) parameters, while Panel B reports utility parameters governing choice conditional on consideration.

Distance and operating hours are strong predictors of consideration across both attentive logit specifications. Hospitals that are closer and have longer operating hours are substantially more likely to enter individuals' consideration sets. Comparing the restricted specification (column 2) and the preferred specification (column 3) quantifies the gain from relaxing the hours-utility exclusion. The attention coefficient on hours falls from 0.84 in column 2 to 0.44 in the preferred specification, and a positive utility coefficient on hours emerges (1.62). The reallocation suggests that under the baseline exclusion, roughly half of the apparent attention effect of hours was in fact preference-relevant variation that the model was forced to attribute to consideration. The fact that the residual attention coefficient on hours remains large and significant (0.44) shows that hours genuinely operates through both channels: longer hours raise both the probability of considering a hospital and the utility of choosing it conditional on consideration.

Word-of-mouth exposure affects this screening process. The WOM $\times$ OPD Hours interaction is negative and marginally significant, indicating that WOM exposure dampens responsiveness to operating hours in the consideration stage: WOM users are less likely to attend to hospitals with longer hours than observationally comparable non-users. The point estimate on WOM $\times$ Distance is negative (-0.56), but imprecisely estimated. While the sign is consistent with WOM anchoring attention to nearby providers, the confidence interval includes zero, and we cannot reject the null of no differential distance screening. Importantly, this dampened responsiveness contrasts with stated motives: convenient hours is among the most frequently cited reasons for provider choice in the full sample, and WOM users cite accessibility nearly as often as non-users, yet their choices load significantly less on hours in the consideration stage. Stated accessibility motives are similar across groups, while revealed attention differs.

[5]Model fit statistics in Table A5 support the relaxation of the hours exclusion: relative to the restricted specification, the preferred specification lowers AIC by 23.95 points and BIC by 5.71 points. Further relaxation, in particular adding WOM interactions with drug availability and quality in the attention equation, produces unidentified coefficients with standard errors exceeding point estimates by orders of magnitude (see Table A4), indicating that the data do not support these additional interactions.

Utility parameters are stable across specifications and economically intuitive. Conditional on consideration, patients strongly dislike distance and value quality improvements. Drug availability enters positively but is imprecisely estimated across all specifications. Allowing WOM to interact with utility shifters in an alternative specification (reported in the appendix) yields no evidence that WOM systematically alters preferences once hospitals enter the consideration set. These patterns suggest that WOM primarily affects which hospitals are considered, rather than how patients evaluate observable attributes conditional on consideration.

#### 4.4.2 Marginal effects

Table 6 translates these choice estimates into interpretable magnitudes by reporting average marginal effects (AMEs) on unconditional choice probabilities. Each AME is decomposed into a consideration channel and a conditional choice channel. Distance has the largest effect: a 0.1 standard deviation increase reduces the probability of choosing a hospital by 0.010, approximately 7 percent of the mean choice probability of 0.153. This occurs entirely through consideration. OPD hours also works primarily through the consideration channel, with a smaller total AME of about 2 percent of the mean. Utility-stage variables, that is quality improvement, drug availability, and facility type, have negligible marginal effects on unconditional choice probabilities, suggesting that consideration is the primary channel in this setting. WOM interactions functions exclusively through consideration. WOM $\times$ OPD Hours lowers unconditional choice probabilities by 4 percent of the mean, while WOM $\times$ Distance lowers them by 2.5 percent.

The estimated attention parameters imply narrow consideration sets. To quantify this, we compute the expected number of hospitals considered by each patient conditional on being attentive. Because consideration is latent, we recover this as an expected value from the estimated attention probabilities. For each choice occasion $t$ of individual $i$, the expected consideration set size is

$$\mathrm{E}\left[\,|\mathcal{C}_{it}|\,\right] \;=\; 1 \;+\; \sum_{j \neq \text{default}} \hat{\pi}_{ijt}, \tag{4}$$

where $\hat{\pi}_{ijt} \;=\; \Lambda(\hat{\gamma}_0 + \hat{\gamma}_1 \text{Distance}_{ij} + \hat{\gamma}_2 \text{Hours}_j + W_i[\hat{\delta}_1 \text{Distance}_{ij} + \hat{\delta}_2 \text{Hours}_j])$ is the estimated consideration probability for non-default hospital $j$ from equation (2), and the leading 1 reflects the default hospital, which is always considered when the patient is attentive.[6]

Under the preferred specification, the mean of (4) across the 533 individuals in the estimation sample is 2.26, the default plus roughly 1.3 additional hospitals out of up to seven non-default alternatives. The baseline attentive logit without WOM interactions

[6]The expected consideration set size is computed over the sampled choice set (chosen hospital, default, and six randomly drawn alternatives) and therefore depends on the specific draw. Because the seed is fixed and a single draw is used for all specifications, the numbers are internally consistent across models, but the levels should be interpreted with this caveat in mind.

implies an even narrower mean of 2.06. These figures are consistent with the low baseline consideration probability implied by the attention constant ($\hat{\gamma}_0 = -0.99$, corresponding to $\Lambda(-0.99) \approx 0.27$ before distance and hours penalties are applied), and with the baseline value in Table 7 showing that WOM users choose the default hospital with probability 0.90.

### 4.4.3 Counterfactual simulations

We conduct two counterfactual exercises to quantify the mechanisms implied by the attention estimates. In both, choice probabilities are predicted by enumerating all possible consideration sets per individual. Table 7 reports the results.

The first exercise isolates the WOM attention channel. We set the WOM interaction terms in the attention index to zero for the 53 WOM-exposed individuals, holding preferences fixed. Removing WOM-induced attention distortions expands the expected consideration set by 0.16 hospitals (8 percent) and reduces reliance on the closest hospital by 1.9 percentage points. Choices shift toward hospitals with longer operating hours and expected travel distance rises modestly. Expected quality improves by less than one percentage point. The dominant pattern is that WOM narrows consideration, but because the default hospital is strongly preferred on distance grounds (the utility coefficient on distance is $-3.85$), the reallocation of choice probability away from the default is modest even when the consideration set expands.

The second exercise asks whether a feasible information intervention could broaden consideration enough to improve the quality of providers patients choose. We raise the attention constant $\gamma_0$ by $\Delta \in \{0.5, 1.0\}$ for all 533 individuals (the full estimation sample, not just WOM users), holding all other parameters fixed (including WOM distortions). This simulates a facility directory, community health worker outreach program, or mobile application that increases baseline awareness of non-default hospitals. Raising $\gamma_0$ by 0.5 moves the baseline consideration probability from 0.27 to 0.38 and expands consideration sets by 0.38 hospitals (17 percent). The larger intervention ($\Delta = 1.0$, raising the baseline probability to 0.50) roughly doubles this effect. Both interventions reduce P(default), shift expected choices toward hospitals with longer hours and higher quality, and modestly increase expected travel distance. The quality gains are small in absolute terms (0.6 to 1.0 percentage points) because most hospitals in the choice set already report quality improvements, limiting the scope for reallocation. The effects are broadly similar for WOM and non-WOM patients, suggesting that the primary constraint on provider matching is low baseline awareness rather than WOM-specific distortion.

## 5 Discussion and concluding remarks

This paper investigates how WOM recommendations shape healthcare decisions in a low-income setting where formal quality information is scarce. We document three main findings. First, WOM is substantially associated with higher OOP costs. Our estimates indicate that WOM users spend approximately ₹790-820 more per visit than non-users, a 76-86 percent increase relative to the sample mean. This gap is concentrated in higher drug and transport expenditures, and is robust to alternative specifications and winsorization. Second, WOM is not associated with better care quality or patient satisfaction. WOM users are no more likely to rate care as excellent or to receive additional diagnostic tests, yet they are significantly less likely to return to the same provider.

Third, choice estimates reveal that WOM operates primarily through the consideration stage of provider choice. WOM users exhibit dampened responsiveness to operating hours when screening hospitals, and consideration sets are narrow: on average the default plus only one additional hospital out of up to seven alternatives. Counterfactual simulations suggest that removing WOM-induced attention distortions would modestly expand consideration and shift choices toward hospitals with longer hours, though strong distance preferences limit the reallocation of choice probability away from the default provider.

How do these descriptive and choice findings fit together? The descriptive and choice analyses answer related but distinct questions using overlapping but different samples, and it is worth being explicit about how they connect. The descriptive estimates in Section 3.1 identify the total association between WOM and costs across all provider types: WOM is associated with high OOP spending, concentrated in higher transportation and drug costs. The choice model identifies one mechanism, namely attention narrowing, among public hospital and CHC users only.

The two sets of results are therefore complementary rather than competing. The choice model explains part of the descriptive finding: among public hospital users, WOM dampens responsiveness to operating hours in the consideration stage and may tighten distance-based screening, anchoring patients to a narrower set of familiar providers. The counterfactual exercises confirm the direction: removing WOM-induced attention distortions expands consideration and shifts choices toward hospitals with longer hours, though the reallocation of choice probability is modest because the default hospital dominates on distance. An information intervention that raises baseline awareness produces larger effects, but even doubling the consideration probability yields quality gains of only one percentage point.

The choice and descriptive findings present an apparent puzzle. The choice model suggests that WOM narrows consideration toward closer hospitals, yet the descriptive estimates show that WOM users incur higher travel-related costs. Several mechanisms could explain this pattern. WOM-referred patients may choose nearby public hospitals that experience frequent stock-outs, necessitating secondary trips to private pharmacies, consistent with the elevated cost per drug at private pharmacies among WOM users (Table 3). Alternatively, the transport gap may be concentrated in facility types excluded from the choice model, particularly private pharmacies, where the estimated association is largest in the descriptive analysis.

The drug cost evidence is consistent with a stock-out channel: WOM users pay more per drug at private pharmacies and are more likely to purchase there, a pattern that would arise if referrals to nearby public facilities are followed by secondary trips to private pharmacies when stock-outs occur. Because the transport gap partly reflects other medical expenses, and because direct identification requires linking the public-sector visit, the stock-out event, and the downstream pharmacy purchase within the same episode, we treat this as one plausible channel rather than a definitive explanation.

A related limitation arises from the measurement of WOM itself, which suggests yet another channel through which transport costs could rise. Because referral status is observed only for the chosen provider, the choice model captures how WOM users screen *all* hospitals' attributes differently, not the direct placement of a specific recommended hospital into the consideration set. In principle, a friend's recommendation could place a distant hospital into the consideration set (raising transport costs) while simultaneously narrowing the patient's attention to other hospitals' characteristics (the mechanism the model identifies). These two channels – direct placement of a recommended provider and generalized attention narrowing – are not mutually exclusive, but we can identify only the latter.

These findings highlight the need to strengthen mandatory quality-regulation frameworks and investments in integrated, patient-facing information systems. The National Health Policy of 2017 emphasizes high-quality patient-centered care, and India now has several public quality-certification programs, including the National Quality Assurance Standards, Kayakalp, LaQshya, and MusQan. These programs remain fragmented, disease- or service-specific, and burdened by overlapping standards and logistical hurdles (Gandhi et al., 2025; Jayanthi and Jain, 2026).

Similarly, the Ayushman Bharat Digital Mission (ABDM), specifically the Health Facility Registry (HFR) and Healthcare Professionals Registry (HPR), is a step in the right direction. By providing a verified, geospatial repository of public and private facilities, the HFR increases institutional visibility and streamlines regulatory processes like empanelment, certification, and license renewals. Crucially, mandating HFR registration could compel widespread compliance with the Clinical Establishments Act (CEA) of 2010, bridging the current enforcement gap to ensure that facilities meet standardized quality and safety benchmarks. Despite the progress of the ABDM, efforts to reduce inter-state variation in HPR and HFR registrations have been slow, particularly in the private sector (Mishra et al., 2024).

## References


Abaluck, J. and Adams-Prassl, A. (2021). What do consumers consider before they choose? identification from asymmetric demand responses. *The Quarterly Journal of Economics*, 136(3):1611–1663.

Ashraf, N., Jack, B. K., and Kamenica, E. (2013). Information and subsidies: Complements or substitutes? *Journal of Economic Behavior & Organization*, 88:133–139.

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

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives*, 12(3):151–170.

Björkman, M. and Svensson, J. (2009). Power to the people: Evidence from a randomized field experiment on community-based monitoring in Uganda. *Quarterly Journal of Economics*, 124(2):735–769.

Björkman Nyqvist, M., de Walque, D., and Svensson, J. (2017). Experimental evidence on the long-run impact of community-based monitoring. *American Economic Journal: Applied Economics*, 9(1):33–69.

Bose, B., Cheng, T. C., Kalita, A., Haakenstad, A., and Yip, W. (2024). Resource shortage in public health facilities and private pharmacy practices in Odisha, India. *Health Policy and Planning*, 39(10):1074–1086.

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

Das, J. and Hammer, J. (2014). Quality of primary care in low-income countries: Facts and economics. *Annual Review of Economics*, 6(1):525–553.

Das, J., Holla, A., Das, V., Mohanan, M., Tabak, D., and Chan, B. (2012). In urban and rural India, a standardized patient study showed low levels of provider training and huge quality gaps. *Health Affairs*, 31(12):2774–2784.

Das, J., Holla, A., Mohpal, A., and Muralidharan, K. (2016). Quality and accountability in health care delivery: Audit-study evidence from primary care in India. *American Economic Review*, 106(12):3765–3799.

Deri, C. (2005). Social networks and health service utilization. *Journal of Health Economics*, 24(6):1076–1107.

Dupas, P. and Miguel, E. (2017). Impacts and determinants of health levels in low-income countries. In *Handbook of Economic Field Experiments*, volume 2, pages 3–93. Elsevier.

Foster, A. D. and Rosenzweig, M. R. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 103(6):1176–1209.

Gandhi, B. P., Bhanderi, D., and Parmar, H. N. (2025). To gain insights and compare different quality certification programs for public health facilities in India. *Journal of Family Medicine and Primary Care*, 14(10):4117–4124.

Haakenstad, A., Kalita, A., Bose, B., Chakraborty, A., Gupta, K., Tsuei, S. H.-T., Woskie, L. R., and Yip, W. (2025). Public patient forwarding to private pharmacies: an analysis of data linking patients, facilities and pharmacies in the state of odisha, india. *BMJ Global Health*, 10(2):e017788.

Jayanthi, T. and Jain, T. (2026). Assets and obstacles in obtaining national quality assurance standards certification: Lessons learnt from certified public health facilities in Tamil Nadu. *Indian Journal of Community Health*, 38(1):96–100.

Li, J. and Zhang, X. (2022). Do people provide good advice? *Journal of Economic Behavior & Organization*, 199:44–64.

McFadden, D. (1978). Modelling the choice of residential location. In Karlqvist, A., Lundqvist, L., Snickars, F., and Weibull, J., editors, *Spatial Interaction Theory and Planning Models*, pages 75–96. North-Holland, Amsterdam.

Mishra, U. S., Yadav, S., and Joe, W. (2024). The ayushman bharat digital mission of india: an assessment. *Health Systems & Reform*, 10(2):2392290.

Mohanan, M., Vera-Hernandez, M., Das, V., Giardili, S., Goldhaber-Fiebert, J. D., Rabin, T. L., Raj, S. S., Schwartz, J. I., and Seth, A. (2015). The know-do gap in quality of health care for childhood diarrhea and pneumonia in rural India. *JAMA Pediatrics*, 169(4):349–357.

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163.

# Figures

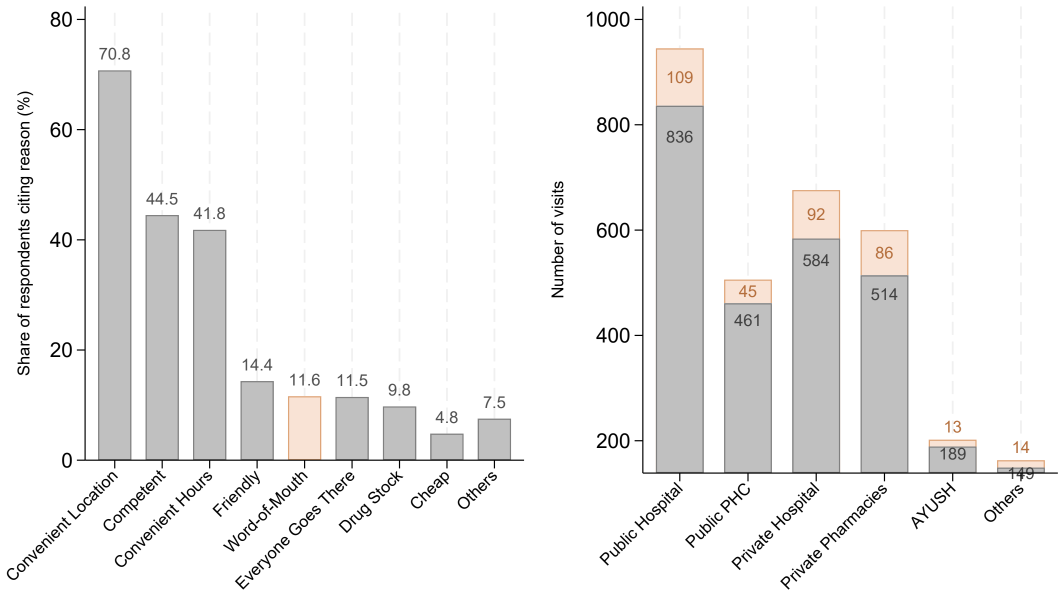


Figure 1. Prevalence of Word-of-Mouth (WOM)

*Notes:* Data are from consumer surveys on illness or injury-related events over the last 15 days. The sample is restricted to an individual's first visit to a healthcare provider, comprising 3,090 individuals across 373 villages. In Panel A, the height of each bar depicts the share of respondents citing each reason for choosing their provider; multiple responses were permitted, so shares sum to more than 100 percent. "Word-of-Mouth" corresponds to informal referrals by relatives or friends; formal clinical referrals by public or private providers are recorded separately in the survey and grouped under 'Others' in the figure. The 'Others' category aggregates facility cleanliness, being the only provider in the area, and formal referrals by public or private providers..

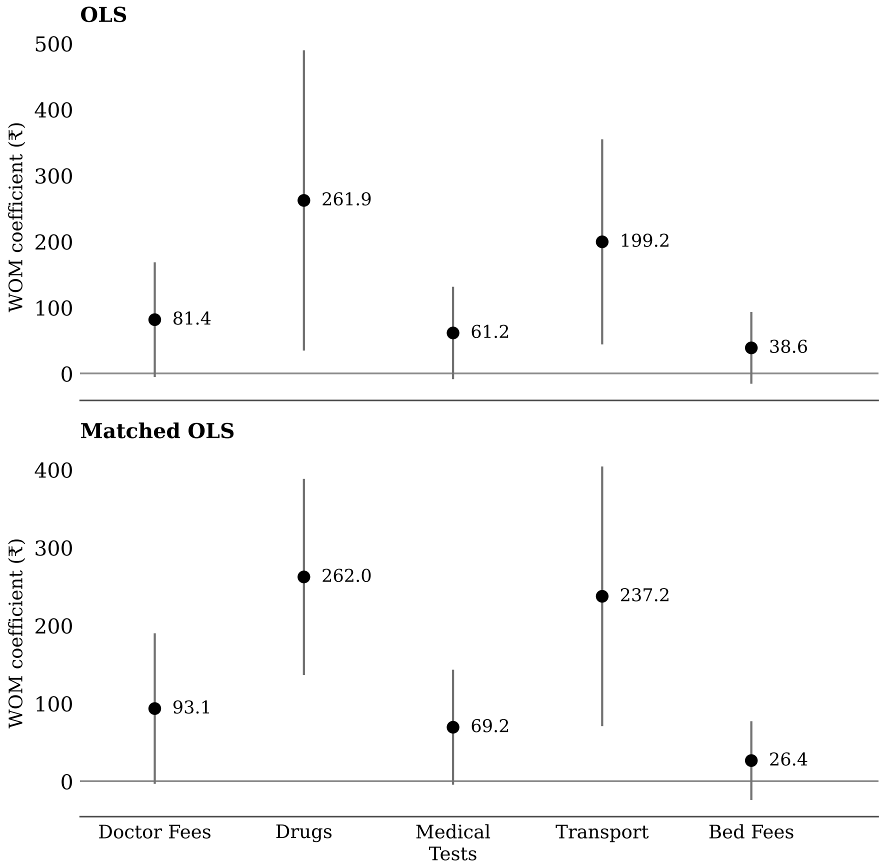


Figure 2. WOM and Components of OOP Costs

*Notes:* The figure shows OLS estimates (top panel) and CEM-matched OLS estimates (bottom panel) of the association between word-of-mouth (WOM) and five components of out-of-pocket (OOP) costs, based on a sample of up to N = 3,090 provider visits (N = 1,690 for the matched panel). The dots represent point estimates, and the vertical lines represent 95% confidence intervals. All models include fixed effects for village, diagnosis, and facility type, with standard errors clustered at the village level. Drug costs include expenditures both within the initial healthcare facility and at nearby pharmacies. We exclude other cost categories included in total costs such as physiotherapy, disposable appliances, personal medical devices, blood, and other miscellaneous items. The "Transport" category combines transportation costs and other medical expenses associated with the visit.

Table 1. Sample Characteristics by Word-of-Mouth Use

| | Word-of-Mouth | | Difference | P-value |
|---|---|---|---|---|
| | Yes (1) | No (0) | (Yes - No) | |
| **A. Demographics** | | | | |
| N | 359 (11.6%) | 2,731 (88.4%) | | |
| Age | 34.786 (21.877) | 33.031 (23.016) | 1.755 | 0.172 |
| Female (=1) | 0.460 (0.499) | 0.523 (0.500) | -0.063 | 0.025** |
| Married (=1) | 0.661 (0.474) | 0.622 (0.485) | 0.039 | 0.170 |
| Household Size | 4.432 (1.794) | 4.387 (1.744) | 0.045 | 0.650 |
| Schedule Tribe (=1) | 0.148 (0.355) | 0.179 (0.384) | -0.031 | 0.138 |
| Schedule Caste (=1) | 0.198 (0.399) | 0.201 (0.401) | -0.003 | 0.890 |
| Other Backward Class (=1) | 0.479 (0.500) | 0.400 (0.490) | 0.079 | 0.004*** |
| Other (=1) | 0.175 (0.381) | 0.218 (0.413) | -0.043 | 0.062* |
| **B. Wealth** | | | | |
| Any Insurance (=1) | 0.117 (0.322) | 0.152 (0.359) | -0.035 | 0.077* |
| BPL Card (=1) | 0.271 (0.445) | 0.245 (0.430) | 0.026 | 0.288 |
| Log Annual Income | 11.313 (0.778) | 11.400 (0.885) | -0.087 | 0.079* |
| **C. Health Status** | | | | |
| Chronic Disease (=1) | 0.203 (0.403) | 0.248 (0.432) | -0.045 | 0.065* |
| Poor/Fair Health (=1) | 0.201 (0.401) | 0.194 (0.395) | 0.007 | 0.766 |
| **D. Post-Visit Health Diagnosis** | | | | |
| Fever (=1) | 0.496 (0.501) | 0.583 (0.493) | -0.087 | 0.002*** |
| Childbirth (=1) | 0.047 (0.213) | 0.030 (0.170) | 0.017 | 0.072* |
| Diarrhea (=1) | 0.031 (0.173) | 0.038 (0.190) | -0.007 | 0.505 |
| Heart Disease (=1) | 0.025 (0.157) | 0.045 (0.207) | -0.020 | 0.079* |
| Injury (=1) | 0.022 (0.148) | 0.018 (0.134) | 0.004 | 0.601 |
| Diabetes (=1) | 0.022 (0.148) | 0.016 (0.126) | 0.006 | 0.392 |
| Acute Respiratory Infection (=1) | 0.114 (0.319) | 0.078 (0.269) | 0.036 | 0.020** |
| Local Pain/Weakness (=1) | 0.114 (0.319) | 0.068 (0.253) | 0.046 | 0.002*** |
| Skin Infection (=1) | 0.014 (0.117) | 0.025 (0.157) | -0.011 | 0.187 |
| Cancer (=1) | 0.006 (0.075) | 0.002 (0.043) | 0.004 | 0.161 |

*Notes:* The table presents summary statistics comparing individuals who chose their provider based on word-of-mouth (WOM) to those who did not. The sample is restricted to an individual's first visit to a healthcare provider for an illness or injury in the last 15 days. For continuous variables (e.g., Age), the table shows the mean with the standard deviation in parentheses. For binary variables (e.g., Female=1), the table shows the proportion. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2. WOM and Total Out-of-Pocket Costs

| | Total OOP Costs | |
|---|---|---|
| | (1) OLS | (2) Matched OLS |
| WOM | 789.7*** | 819.8*** |
| | (298.6) | (238.5) |
| Observations | 3,090 | 1,690 |
| Mean of dep. var. | ₹1035.8 | ₹954.2 |
| $R^2$ | 0.207 | 0.333 |
| Village FE | Yes | Yes |
| Diagnosis FE | Yes | Yes |
| Facility Type FE | Yes | Yes |

Notes: The sample includes individuals who reported seeking healthcare in the last 15 days at the time of the interview. Total out-of-pocket cost is the sum of expenses for doctor fees, tests, drugs, physiotherapy, blood, oxygen, medical appliances, disposables, transport, bed fees, and other medical/non-medical items. All models include fixed effects for village, diagnosis, and facility type. Standard errors, clustered at the village level, are in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3. Word-of-Mouth, Drug Costs, and Utilization for Patients Prescribed Medicine

| | (1) OLS | (2) Matched OLS |
|---|---|---|
| *Panel A. Total Drug Cost* | | |
| WOM | 297.198** | 289.480*** |
| | (131.679) | (77.370) |
| Observations | 2,531 | 1,380 |
| Mean of dep. var. | 557.34 | 513.94 |
| $R^2$ | 0.205 | 0.320 |
| *Panel B. Number of Drugs* | | |
| WOM | 0.130 | 0.150 |
| | (0.107) | (0.135) |
| Observations | 2,531 | 1,380 |
| Mean of dep. var. | 3.32 | 3.29 |
| $R^2$ | 0.229 | 0.311 |
| *Panel C. Cost per Drug, Public Pharmacy* | | |
| WOM | -19.973 | -2.289 |
| | (13.659) | (6.005) |
| Observations | 264 | 114 |
| Mean of dep. var. | 21.84 | 11.64 |
| $R^2$ | 0.585 | 0.751 |
| *Panel D. Cost per Drug, Private Pharmacy* | | |
| WOM | 72.302*** | 52.741*** |
| | (26.508) | (19.808) |
| Observations | 2,118 | 1,148 |
| Mean of dep. var. | 176.56 | 171.52 |
| $R^2$ | 0.249 | 0.349 |
| Village FE | Yes | Yes |
| Diagnosis FE | Yes | Yes |
| Facility Type FE | Yes | Yes |

Notes: Each panel reports a separate regression of the outcome on an indicator for word-of-mouth (WOM). The analysis is restricted to patients who were prescribed at least one drug. Panels C and D additionally condition on having purchased from a public or private pharmacy, respectively. All models include fixed effects for village, diagnosis, and facility type. Standard errors, clustered at the village level, are in parentheses. Column (2) reweights by coarsened-exact-matching weights. Panel C, column (2) is estimated on only 44 village clusters, significant fewer than the roughly 350 elsewhere in this table, and should be read with additional caution. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4. Word-of-Mouth, Treatment Intensity, and Patient-Reported Outcomes

| | (1) OLS | (2) Matched OLS |
|---|---|---|
| *Panel A. Overall Quality (=1 if 'Excellent')* | | |
| WOM | -0.020 | -0.018 |
| | (0.016) | (0.016) |
| Observations | 3,090 | 1,690 |
| Mean of dep. var. | 0.065 | 0.069 |
| $R^2$ | 0.205 | 0.394 |
| *Panel B. Likely to Return (=1 if 'Extremely likely')* | | |
| WOM | -0.070*** | -0.059** |
| | (0.022) | (0.027) |
| Observations | 3,090 | 1,690 |
| Mean of dep. var. | 0.171 | 0.156 |
| $R^2$ | 0.197 | 0.382 |
| *Panel C. Number of Diagnostic Tests* | | |
| WOM | 0.078 | 0.107* |
| | (0.049) | (0.056) |
| Observations | 3,090 | 1,690 |
| Mean of dep. var. | 0.331 | 0.327 |
| $R^2$ | 0.229 | 0.383 |
| Village FE | Yes | Yes |
| Diagnosis FE | Yes | Yes |
| Facility Type FE | Yes | Yes |

Notes: Each panel reports a separate regression of the named outcome on an indicator for word-of-mouth. The dependent variable in Panel A is an indicator equal to 1 if the patient rated quality as 'Excellent'. The dependent variable in Panel B is an indicator equal to 1 if the patient was 'Extremely likely' to return. The dependent variable in Panel C is the number of diagnostic tests received. All models include fixed effects for village, diagnosis, and facility type. Standard errors, clustered at the village level, are in parentheses. Column (2) reweights by coarsened-exact-matching weights. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5. Choice Demand Estimates

| | (1) Cond. Logit | (2) Restricted Alogit | (3) Preferred Alogit |
|---|---|---|---|
| *Panel A: Attention Parameters ($P(A)$)* | | | |
| Distance (Std.) | — | $-1.501^{***}$ | $-1.436^{***}$ |
| | | (0.130) | (0.127) |
| OPD Hours (Std.) | — | $0.843^{***}$ | $0.438^{**}$ |
| | | (0.185) | (0.187) |
| WOM $\times$ Distance | — | — | $-0.564$ |
| | | | (0.465) |
| WOM $\times$ OPD Hours | — | — | $-0.901^{*}$ |
| | | | (0.464) |
| Constant $\gamma_0$ | — | $-1.483^{***}$ | $-0.989^{***}$ |
| | | (0.480) | (0.349) |
| *Panel B: Utility Parameters ($P(Y \mid A)$)* | | | |
| Distance (Std.) | $-2.093^{***}$ | $-3.031^{**}$ | $-3.851^{***}$ |
| | (0.126) | (1.258) | (1.294) |
| OPD Hours (Std.) | $0.678^{***}$ | — | $1.624^{**}$ |
| | (0.114) | | (0.666) |
| Drug Availability (%) | $1.450^{**}$ | 3.245 | 4.353 |
| | (0.573) | (1.993) | (2.693) |
| Quality Improvement | $0.997^{***}$ | $2.017^{**}$ | $2.500^{**}$ |
| | (0.185) | (0.830) | (1.117) |
| CHC (Dummy) | $-0.364^{*}$ | $-0.890$ | $-0.378$ |
| | (0.210) | (0.622) | (0.874) |
| Observations | 3,233 | 3,233 | 3,233 |
| Log likelihood | $-325.96$ | $-343.68$ | $-328.70$ |

*Notes:* Column (1) tabulates a conditional logit assuming full consideration. Columns (2)–(3) report estimates from the attentive logit model of Abaluck and Adams-Prassl (2021), computed with the exact likelihood. Panel A reports attention parameters influencing whether a hospital enters the consideration set. Panel B reports utility parameters governing choice conditional on consideration. In column (2) OPD hours enters only through attention; column (3) allows hours to enter both stages. WOM denotes individual-level word-of-mouth use and enters only through interactions with hospital attributes in the attention equation. Distance and OPD hours are standardized. Choice sets include the chosen hospital, the closest hospital (default), and six randomly sampled alternatives within the district. Log likelihoods for the conditional logit and the attentive logit are computed on different scales and are not comparable across models. Standard errors in parentheses; column (1) uses standard errors clustered on the individual, columns (2)–(3) report conventional attentive-logit standard errors. $^{***}$ $p < 0.01$, $^{**}$ $p < 0.05$, $^{*}$ $p < 0.1$.

Table 6. Average Marginal Effects on Hospital Choice Probability

| | | Decomposition of Total AME | |
|---|---|---|---|
| Variable | Total AME | Consideration | Choice |
| Distance (0.1 SD) | -0.0102 | -0.0102 | -0.0000 |
| OPD Hours (0.1 SD) | 0.0024 | 0.0024 | 0.0000 |
| Quality Improvement (0–1) | 0.0012 | 0.0000 | 0.0012 |
| WOM $\times$ Distance (0.1 SD) | -0.0038 | -0.0038 | 0.0000 |
| WOM $\times$ OPD Hours (0.1 SD) | -0.0061 | -0.0061 | 0.0000 |

*Notes:* Average marginal effects computed from the preferred attentive logit specification (Table 5, column 3). Each AME is decomposed into a consideration channel (change in the probability that the hospital enters the choice set, weighted by conditional choice probability) and a choice channel (change in conditional choice probability, weighted by attention probability). Effects are scaled by a 0.1 standard deviation change for continuous standardized variables and a discrete 0–1 change for binary variables.

Table 7. Counterfactual Simulations

| | WOM-exposed (N=53) | | Full sample (N=533) | | |
|---|---|---|---|---|---|
| Outcome | Baseline | No WOM | Baseline | $\gamma_0$+0.5 | $\gamma_0$+1.0 |
| Consid. set size | 2.019 | 2.177 | 2.264 | 2.647 | 3.074 |
| Prob. choose default | 0.903 | 0.884 | 0.859 | 0.841 | 0.826 |
| Exp. quality (0–1) | 0.903 | 0.911 | 0.940 | 0.946 | 0.950 |
| Exp. OPD hours (SD) | 0.538 | 0.591 | 0.581 | 0.599 | 0.612 |
| Exp. distance (SD) | −1.539 | −1.523 | −1.525 | −1.514 | −1.505 |

*Notes:* "Baseline" under WOM-exposed is the observed average among the 53 WOM-exposed individuals; "No WOM" recomputes their outcomes with WOM's estimated effect on attention set to zero, preferences unchanged. "Baseline" under Full sample is the observed average among all 533 individuals; the two $\gamma_0$ columns simulate an information intervention raising baseline awareness of non-default hospitals for everyone, holding all other parameters (including WOM interactions) fixed. Distance and OPD hours are expressed as standard deviations from the sample mean.

# A Appendix

## A.1 Survey Design

### A.1.1 Household Sampling

To sample households, first, all households in each sampled village were listed. From the listing data, the following information about each household in a village was collected: (i) Number of households with no event (NE); (ii) Number of households with an outpatient visit in the last 15 days (O); (iii) Number of households with a chronic illness diagnosed by a health provider (C); and (iv) Number of households with hospitalization in the last one year (H). In addition to the four points of information above, the listing tool also collected information on the 'preferred provider of the household for outpatient care' and 'preferred provider of the household for inpatient care'. While this information was not used for household sampling, this was used for provider sampling (described later in this report). An absolute precision of 0.007 and a design effect of 2.5 were assigned to arrive at a household sample size of 7500. Considering the rarest event, hospitalization (0.04 or 4.4 percent based on NSSO data), for a sample of 7500 households, the confidence interval (CI) is 95 percent and design effect is 2.5. The number of households sampled from each category were 3000 for no event (NE), and 1500 households each from the outpatient visit (O), chronic illness (C) and hospitalization (H) categories. The oversampling of these latter categories (O, C and H) in comparison to the NE group aims to provide enough precision to stratified estimates of these sub-samples, such as disease-specific expenditures or gender-specific care-seeking frequency. These sample sizes were based on the highest possible number of households that could have these events, as well as the margin of error estimates for variables of interest, for example, expenditure per hospitalization, expenditure per outpatient visit, or percentage referred to chemists. Therefore, the overall margin of error for the estimates of these groups grew even smaller, ranging from 0.000 to 0.001 percent. The total number of households sampled was 7567. Data was collected about each member of each of these households, so the total number of individuals in the sample was approximately 30,645.

## A.2 Appendix Figures

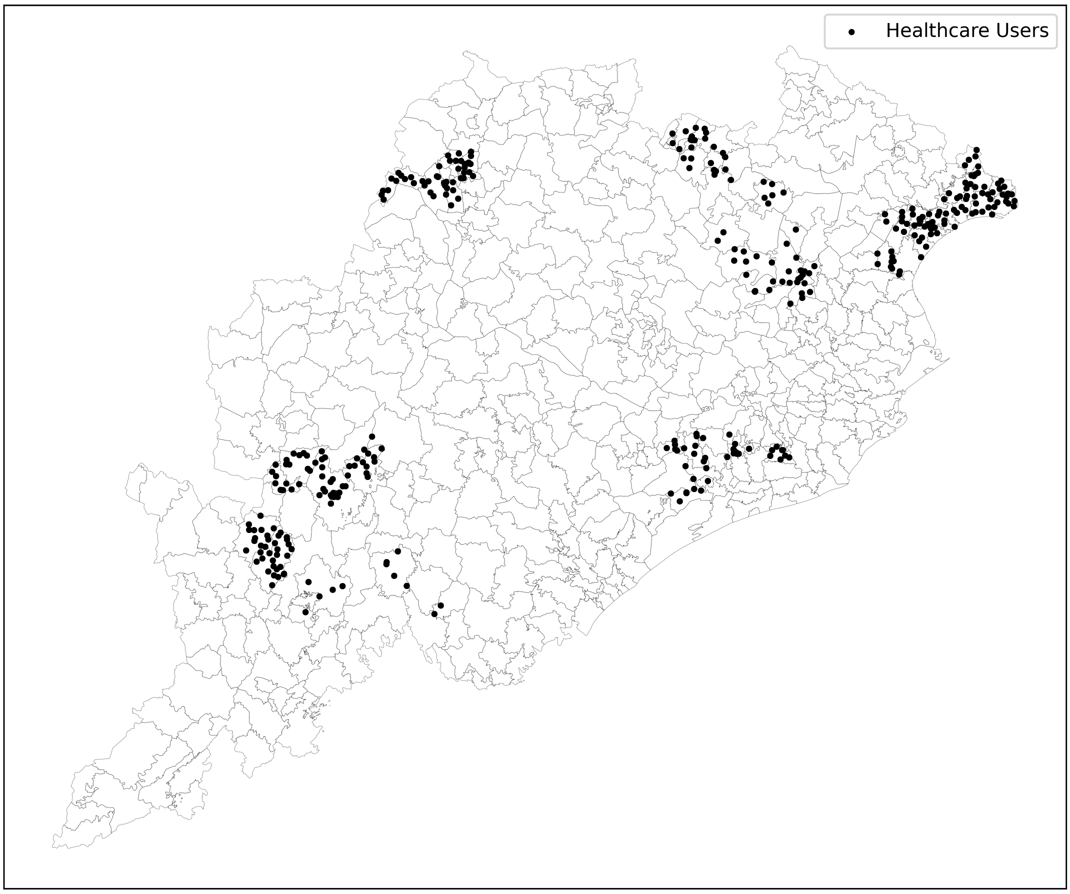


Figure A1. Location of Individuals Who Sought Health Care in the last 15 days

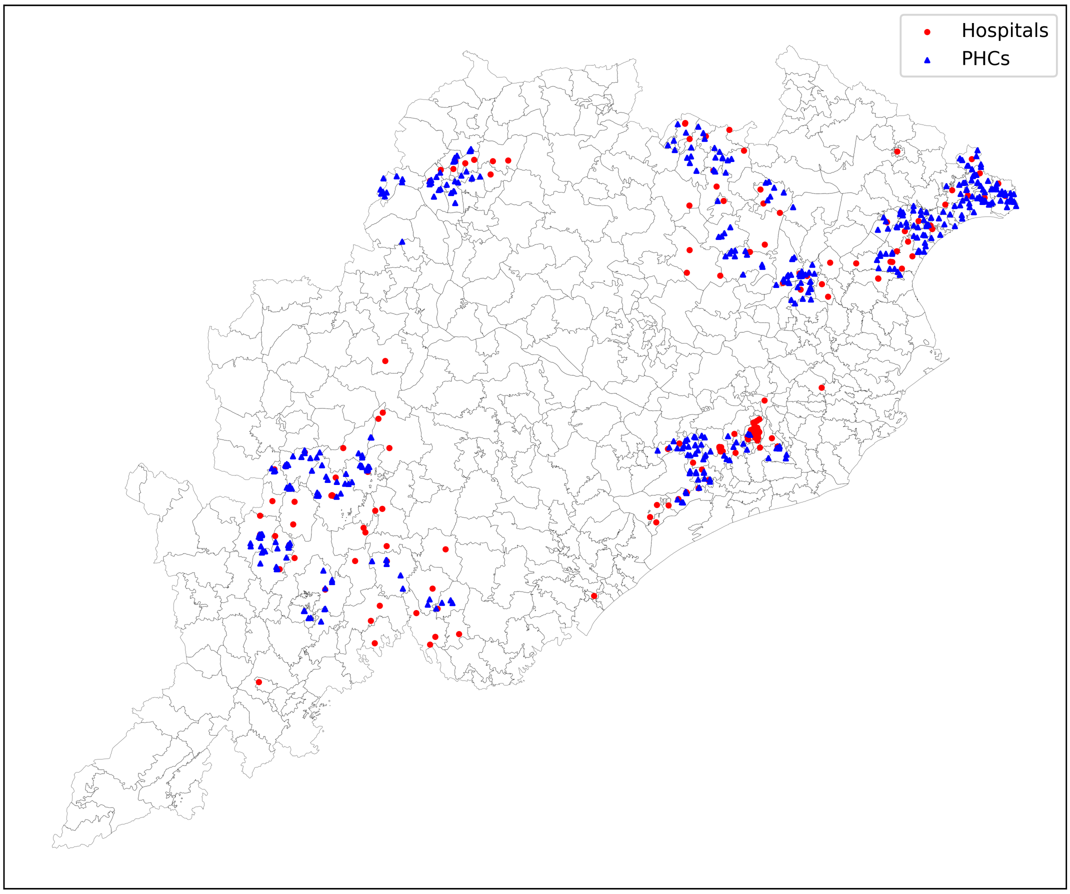


Figure A2. Location of Surveyed Hospitals and PHCs

### A.3 Appendix Tables

Table A1. Post-Match Balance Table: Individuals Who Use Word-of-Mouth

| | Word-of-Mouth | | Difference | P-value |
|---|---|---|---|---|
| | Yes (1) | No (0) | (Yes - No) | |
| **A. Demographics** | | | | |
| N | 323 (18.9%) | 1,387 (81.1%) | | |
| Age | 33.433 (21.518) | 33.468 (21.369) | -0.035 | 0.979 |
| Female (=1) | 0.461 (0.499) | 0.540 (0.499) | -0.079 | 0.011** |
| Married (=1) | 0.643 (0.480) | 0.627 (0.484) | 0.016 | 0.627 |
| Household Size | 4.430 (1.722) | 4.337 (1.641) | 0.093 | 0.360 |
| Schedule Tribe (=1) | 0.155 (0.362) | 0.183 (0.386) | -0.028 | 0.239 |
| Schedule Caste (=1) | 0.198 (0.399) | 0.206 (0.404) | -0.008 | 0.758 |
| Other Backward Class (=1) | 0.483 (0.500) | 0.396 (0.489) | 0.087 | 0.004*** |
| Other (=1) | 0.164 (0.371) | 0.215 (0.411) | -0.051 | 0.041** |
| **B. Wealth** | | | | |
| Any Insurance (=1) | 0.118 (0.323) | 0.178 (0.383) | -0.060 | 0.009*** |
| BPL Card (=1) | 0.280 (0.449) | 0.291 (0.454) | -0.011 | 0.695 |
| Log Annual Income | 11.338 (0.627) | 11.334 (0.625) | 0.004 | 0.928 |
| **C. Health Status** | | | | |
| Chronic Disease (=1) | 0.186 (0.390) | 0.235 (0.424) | -0.049 | 0.055* |
| Poor/Fair Health (=1) | 0.186 (0.390) | 0.181 (0.385) | 0.005 | 0.850 |
| **D. Post-Visit Health Diagnosis** | | | | |
| Fever (=1) | 0.511 (0.501) | 0.511 (0.500) | 0.000 | 1.000 |
| Childbirth (=1) | 0.050 (0.217) | 0.029 (0.168) | 0.021 | 0.061* |
| Diarrhea (=1) | 0.031 (0.173) | 0.035 (0.185) | -0.004 | 0.695 |
| Heart Disease (=1) | 0.025 (0.156) | 0.038 (0.192) | -0.013 | 0.233 |
| Injury (=1) | 0.025 (0.156) | 0.025 (0.156) | 0.000 | 0.989 |
| Diabetes (=1) | 0.025 (0.156) | 0.013 (0.115) | 0.012 | 0.141 |
| Acute Respiratory Infection (=1) | 0.121 (0.326) | 0.098 (0.297) | 0.023 | 0.218 |
| Local Pain/Weakness (=1) | 0.105 (0.307) | 0.086 (0.281) | 0.019 | 0.286 |
| Skin Infection (=1) | 0.012 (0.111) | 0.025 (0.156) | -0.013 | 0.172 |
| Cancer (=1) | 0.003 (0.056) | 0.001 (0.029) | 0.002 | 0.299 |

*Notes:* The table presents summary statistics for the post-match sample, comparing individuals who chose their provider based on word-of-mouth (WOM) to those who did not. The sample is restricted to an individual's first visit to a healthcare provider. For continuous variables, the table shows the mean with the standard deviation in parentheses. For binary variables, the table shows the proportion. Significance levels are denoted as follows: *** p<0.01, ** p<0.05, * p<0.1.

Table A2. Heterogeneity of WOM Effects by Facility Type (OLS)

| | (1)<br>Total OOP | (2)<br>Total OOP (Win. $p_{99}$) |
|---|---|---|
| WOM (= effect at Public Hospital) | 503.2 | 482.2** |
| | (328.3) | (209.1) |
| *WOM × Facility Type Interactions* | | |
| × Public PHC | −40.5 | −85.6 |
| | (426.8) | (327.0) |
| × Private Hospital | 1,000.1 | 159.3 |
| | (982.1) | (371.2) |
| × Private Pharmacy | 258.6 | −49.8 |
| | (544.9) | (276.0) |
| × AYUSH | −101.6 | −119.5 |
| | (700.7) | (613.3) |
| × Other | −278.0 | −448.8* |
| | (406.2) | (258.3) |
| Observations | 3,090 | 3,090 |
| Mean of dep. var. | ₹1,036 | ₹905 |
| $R^2$ | 0.209 | 0.290 |
| Village FE | Yes | Yes |
| Diagnosis FE | Yes | Yes |

Notes: OLS estimates from a pooled regression of out-of-pocket costs on WOM interacted with facility-type indicators. The base category is public hospitals; interaction terms give the differential WOM effect at each other facility type relative to public hospitals. The WOM main effect (first row) is the estimated effect at public hospitals. Column (1) uses unwinsorized costs; column (2) winsorizes at the 99th percentile. Facility-type main effects are included in both models but not reported. The joint test evaluates the null that all interaction terms are zero, i.e., that WOM effects are equal across facility types. All specifications include village and diagnosis fixed effects. Standard errors, clustered at the village level, are in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3. Balance Table: Unlinked vs. Linked Hospital Choice Sample

| | Sample | | Difference | P-value |
|---|---|---|---|---|
| | Unmatched (0) | Matched (1) | (1 - 0) | |
| **A. Demographics** | | | | |
| N | 390 (37.8%) | 643 (62.2%) | | |
| Age | 35.067 (23.148) | 32.876 (23.277) | -2.191 | 0.142 |
| Female (=1) | 0.513 (0.500) | 0.476 (0.500) | -0.037 | 0.250 |
| Married (=1) | 0.668 (0.472) | 0.596 (0.491) | -0.072 | 0.030** |
| Household Size | 4.318 (1.889) | 4.558 (1.666) | 0.240 | 0.033** |
| **B. Wealth** | | | | |
| Any Insurance (=1) | 0.164 (0.371) | 0.162 (0.369) | -0.002 | 0.921 |
| BPL Card (=1) | 0.219 (0.414) | 0.283 (0.451) | 0.064 | 0.023** |
| Log Annual Income | 11.375 (0.995) | 11.397 (0.789) | 0.022 | 0.698 |
| **C. Health Status** | | | | |
| Chronic Disease (=1) | 0.349 (0.477) | 0.240 (0.427) | -0.109 | <0.001*** |
| Poor/Fair Health (=1) | 0.285 (0.452) | 0.201 (0.401) | -0.084 | 0.002*** |
| **D. Post-Visit Diagnosis** | | | | |
| Fever (=1) | 0.469 (0.500) | 0.586 (0.493) | 0.117 | <0.001*** |
| Childbirth (=1) | 0.054 (0.226) | 0.033 (0.178) | -0.021 | 0.095* |
| Diarrhea (=1) | 0.031 (0.173) | 0.044 (0.204) | 0.013 | 0.303 |
| Heart Disease (=1) | 0.051 (0.221) | 0.037 (0.190) | -0.014 | 0.282 |
| Injury (=1) | 0.031 (0.173) | 0.031 (0.174) | 0.000 | 0.976 |
| Diabetes (=1) | 0.028 (0.166) | 0.022 (0.146) | -0.006 | 0.515 |
| Acute Respiratory (=1) | 0.049 (0.216) | 0.084 (0.278) | 0.035 | 0.032** |
| Local Pain/Weakness (=1) | 0.100 (0.300) | 0.075 (0.263) | -0.025 | 0.155 |
| Skin Infection (=1) | 0.031 (0.173) | 0.026 (0.161) | -0.005 | 0.683 |
| Cancer (=1) | 0.005 (0.072) | 0.002 (0.039) | -0.003 | 0.301 |

*Notes:* The table presents summary statistics comparing unlinked individuals to linked individuals in the hospital choice sample. For continuous variables, the table shows the mean with the standard deviation in parentheses. For binary variables, the table shows the proportion. P-values are calculated from the difference in means. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Table A4. Robustness: Alternative Attentive Logit Specifications

| | (1) Hours excluded from utility | (2) All WOM × Attn | (3) Fully Loaded |
|---|---|---|---|
| *Panel A: Attention Parameters ($P(A)$)* | | | |
| Distance (Std.) | $-1.478^{***}$ | $-1.490^{***}$ | $-1.473^{***}$ |
| | (0.125) | (0.130) | (0.128) |
| OPD Hours (Std.) | $0.957^{***}$ | $0.987^{***}$ | $0.946^{***}$ |
| | (0.191) | (0.194) | (0.193) |
| WOM × Distance | $-0.477$ | n.i. | n.i. |
| | (0.458) | | |
| WOM × OPD Hours | $-1.182^{*}$ | n.i. | n.i. |
| | (0.483) | | |
| WOM × Drug Availability | | n.i. | n.i. |
| WOM × Quality Improvement | | n.i. | n.i. |
| *Panel B: Utility Parameters ($P(Y \mid A)$)* | | | |
| Distance (Std.) | $-3.122^{**}$ | $-2.615^{***}$ | $-2.923^{**}$ |
| | (1.017) | (0.455) | (0.939) |
| Drug Availability (%) | 3.182 | 2.701 | 2.989 |
| | (2.035) | (1.487) | (1.915) |
| Quality Improvement | $2.088^{**}$ | $1.702^{***}$ | $2.073^{**}$ |
| | (0.724) | (0.422) | (0.715) |
| CHC (Dummy) | $-0.972$ | $-0.923^{*}$ | $-0.745$ |
| | (0.570) | (0.405) | (0.416) |
| WOM × Distance | | | 0.504 |
| | | | (1.081) |
| WOM × OPD Hours | | | 0.200 |
| | | | (0.389) |
| WOM × Drug Availability | | | 2.711 |
| | | | (4.281) |
| WOM × Quality Improvement | | | $-2.325$ |
| | | | (1.223) |
| Observations | 3,233 | 3,233 | 3,233 |
| Log likelihood | $-340.27$ | $-336.04$ | $-336.13$ |

*Notes:* Alternative specifications of the hybrid attentive logit model. Column (1) restricts OPD hours to the attention equation only (excluding it from utility); the preferred specification in the main table allows hours to enter both stages. Column (2) adds WOM interactions with drug availability and quality improvement to the attention equation; the resulting coefficients are not identified (denoted "n.i."). This motivates restricting WOM interactions to distance and operating hours in the preferred specification. Column (3) additionally allows WOM to interact with both attention and utility-stage variables; attention interactions remain unidentified (n.i.), and none of the WOM × utility interactions are statistically significant, supporting the restriction of WOM to the attention equation. Robust standard errors in parentheses. $^{***}$ $p < 0.01$, $^{**}$ $p < 0.05$, $^{*}$ $p < 0.1$.

Table A5. Model Fit Comparison: Attentive Logit Specifications

| Specification | LL | $k$ | AIC | BIC |
|---|---|---|---|---|
| Baseline | $-343.68$ | 7 | 701.36 | 743.93 |
| Hours excl. utility | $-340.27$ | 9 | 698.53 | 753.26 |
| **Preferred** | **$-328.70$** | **10** | **677.41** | **738.22** |
| All WOM × Attn | $-336.04$ | 11 | 694.09 | 760.98 |
| Fully loaded | $-336.13$ | 15 | 702.27 | 793.49 |

*Notes:* All specifications are estimated using the hybrid attentive logit likelihood of Abaluck and Adams-Prassl (2021) with the closest hospital as default. $k$ is the number of estimated parameters.

## A.4 Robustness to Linked-Sample Selection: IPW-Adjusted Bootstrap

The choice estimation sample consists of 533 individuals: 643 of the 1,033 hospital and CHC visits in the household survey (62 percent) were linked to facility-level survey data, of which 533 (52 percent of all visits) have complete facility attributes. Table A3 documents systematic differences between matched and unmatched patients on marital status, household size, BPL status, chronic disease prevalence, and self-reported health. A natural concern is whether the choice estimates generalize beyond the subsample or instead reflect selection on these observable differences. We address this concern by re-estimating the preferred specification under inverse probability weighting (IPW), where weights $\hat{w}_i = 1/\hat{p}_i$ adjust each linked individual's contribution to the likelihood by the inverse of their estimated probability of being linked.

Propensity scores $\hat{p}_i$ are estimated from logit models predicting match status on demographics (age, gender, marital status, household size), wealth (insurance, BPL card, log income), health (chronic disease, self-reported health), presenting condition fixed effects, and district fixed effects, using the household-survey sample of public hospital and CHC users. Weights are stabilized by the marginal match rate and winsorized at the first and ninety-ninth percentiles. We consider three propensity specifications: a baseline that excludes WOM exposure from the propensity model, an alternative that includes WOM as a predictor of being matched, and a third that excludes WOM and restricts the propensity-fitting sample to non-WOM users.

Inference is conducted via nonparametric bootstrap. We resample unique individuals with replacement and re-estimate the IPW-weighted attentive logit on each draw ($B = 200$). Because the attentive logit likelihood is not globally concave, the optimizer occasionally fails to find a sensible interior solution in resampled data, particularly in draws containing few WOM users where the attention-stage interactions are weakly identified. Following standard practice for inference in non-linear models with non-concave likelihoods, we exclude any bootstrap draw in which a reported coefficient takes an absolute value exceeding 100, a threshold roughly twenty times the largest sensible coefficient magnitude observed in the headline specification.

The exclusion rate is 23 percent for the baseline propensity specification and 42–43 percent for the two alternative specifications, reflecting the small bootstrap sample size (526 individuals, 53 WOM users) and the parameter-rich attention stage. Table A6 reports the bootstrap medians and 2.5–97.5 percentile intervals from the remaining draws, alongside the headline point estimates from the main text. The IPW-adjusted bootstrap medians are broadly consistent with the headline estimates across all three propensity specifications. Utility-stage coefficients on distance and quality preserve their signs and approximate magnitudes; attention-stage coefficients on distance and the WOM $\times$ OPD Hours interaction are similar in sign and magnitude to the headline. Drug availability in utility is more sensitive to reweighting, with bootstrap medians moving from 4.35 in the headline to 1.47–3.75 across IPW specifications, consistent with the marginal significance documented in the main text. The qualitative findings of the paper, namely narrow consideration sets, distance-anchored screening, and WOM-driven attention narrowing through OPD hours, are robust to IPW correction for linked-sample selection.

Table A6. IPW-Adjusted Bootstrap Estimates

| | (1)<br>Headline | (2)<br>IPW: baseline | (3)<br>IPW: with WOM | (4)<br>IPW: non-WOM only |
|---|---|---|---|---|
| *Panel A: Attention parameters* | | | | |
| Distance | −1.436 | −1.506 | −1.453 | −1.448 |
| | (0.127) | [−2.887, −1.291] | [−1.845, −1.266] | [−1.745, −1.257] |
| OPD Hours | 0.438 | 1.067 | 1.143 | 1.160 |
| | (0.187) | [0.747, 7.592] | [0.691, 2.654] | [0.799, 1.702] |
| WOM × Distance | −0.564 | −0.886 | −1.105 | −0.938 |
| | (0.465) | [−14.143, 0.129] | [−6.202, 0.036] | [−11.464, 0.151] |
| WOM × OPD Hours | −0.901 | −1.189 | −1.165 | −1.111 |
| | (0.464) | [−22.048, 0.167] | [−4.649, −0.210] | [−2.856, 0.130] |
| Constant | −0.989 | −1.458 | −1.986 | −2.059 |
| | (0.349) | [−2.343, 5.349] | [−2.916, 2.136] | [−2.863, −1.039] |
| *Panel B: Utility parameters* | | | | |
| Distance | −3.851 | −3.236 | −3.999 | −3.696 |
| | (1.294) | [−6.309, −1.860] | [−8.484, −1.888] | [−10.920, −2.347] |
| Drug Availability | 4.353 | 3.750 | 1.466 | 1.795 |
| | (2.693) | [1.234, 7.236] | [−9.350, 6.744] | [−14.126, 9.657] |
| Quality Improvement | 2.500 | 2.081 | 2.810 | 2.668 |
| | (1.117) | [0.935, 5.445] | [1.421, 22.023] | [1.676, 15.420] |
| CHC | −0.378 | −0.970 | −0.610 | −0.793 |
| | (0.874) | [−2.776, 0.007] | [−2.323, 2.715] | [−2.501, 0.726] |
| $N$ (individuals) | 533 | 526 | 526 | 526 |
| Retained bootstrap draws | — | 154 | 117 | 115 |
| Excluded draws (%) | — | 23.0 | 41.5 | 42.5 |

*Notes:* Column (1) reproduces the headline point estimates from Table 5, column (3), with robust standard errors in parentheses. Columns (2)–(4) report bootstrap medians under IPW reweighting with three alternative propensity specifications: (2) propensity model excludes WOM; (3) propensity model includes WOM; (4) propensity model excludes WOM and is fit on the non-WOM subsample. Brackets in columns (2)–(4) contain 2.5–97.5 percentile intervals from the retained bootstrap draws. A bootstrap draw is excluded if any estimated coefficient has absolute value exceeding 100, a threshold approximately twenty times the largest coefficient magnitude in the headline specification; this rule isolates draws in which the optimizer failed to find a sensible interior maximum. Retention counts and exclusion rates are reported at the bottom of each column. The IPW estimation sample of 526 matched individuals is slightly smaller than the headline sample of 533 because seven linked individuals lack one or more covariates required for propensity-score estimation. Number of bootstrap draws: $B = 200$.

## A.5 Choice Model Robustness to Choice Set Sampling

Table A7. Attentive Logit Estimates by Choice-Set Size

| | Number of sampled rivals, $n$ | | | |
|---|---|---|---|---|
| | 4 | 6 | 8 | 10 |
| *Panel A: Attention parameters* | | | | |
| Distance | $-1.282^{***}$ | $-1.436^{***}$ | $-1.514^{***}$ | $-1.535^{***}$ |
| | (0.126) | (0.127) | (0.128) | (0.128) |
| OPD Hours | $0.504^{***}$ | $0.438^{**}$ | $0.509^{***}$ | $0.486^{***}$ |
| | (0.187) | (0.187) | (0.189) | (0.178) |
| WOM × Distance | $-0.428$ | $-0.564$ | $-0.573$ | $-0.571$ |
| | (0.468) | (0.465) | (0.434) | (0.428) |
| WOM × OPD Hours | $-0.570$ | $-0.901^{*}$ | $-0.881^{*}$ | $-0.881^{*}$ |
| | (0.522) | (0.464) | (0.463) | (0.454) |
| *Panel B: Utility parameters* | | | | |
| Distance | $-3.804^{***}$ | $-3.851^{***}$ | $-4.260^{***}$ | $-4.570^{***}$ |
| | (1.327) | (1.294) | (1.331) | (1.383) |
| Drug avail. | 3.447 | 4.353 | 4.201 | 4.327 |
| | (2.421) | (2.693) | (2.845) | (2.924) |
| Quality | $2.648^{**}$ | $2.500^{**}$ | $2.651^{**}$ | $2.854^{**}$ |
| | (1.179) | (1.117) | (1.109) | (1.182) |
| OPD Hours | $1.670^{***}$ | $1.624^{**}$ | $1.537^{**}$ | $1.644^{**}$ |
| | (0.600) | (0.666) | (0.692) | (0.716) |
| CHC | $-0.734$ | $-0.378$ | $-0.636$ | $-0.714$ |
| | (0.762) | (0.874) | (0.886) | (0.908) |
| $E[\|\mathcal{C}\|]$ (WOM users) | 1.93 | 2.02 | 2.10 | 2.18 |
| $E[\|\mathcal{C}\|]$ (non-WOM users) | 2.16 | 2.29 | 2.35 | 2.41 |
| Observations | 2,411 | 3,233 | 3,794 | 3,990 |

*Notes:* Preferred attentive logit (Table 5, column 3) estimated by the exact likelihood at varying choice-set sizes $n$, the number of randomly sampled rivals added to each individual's chosen and default alternatives. The $n = 6$ column reproduces the preferred specification in Table 5. Standard errors in parentheses. $E[|\mathcal{C}|]$ is computed in closed form from the estimated attention probabilities; the default is always included. $^{***}$ $p < 0.01$, $^{**}$ $p < 0.05$, $^{*}$ $p < 0.1$.

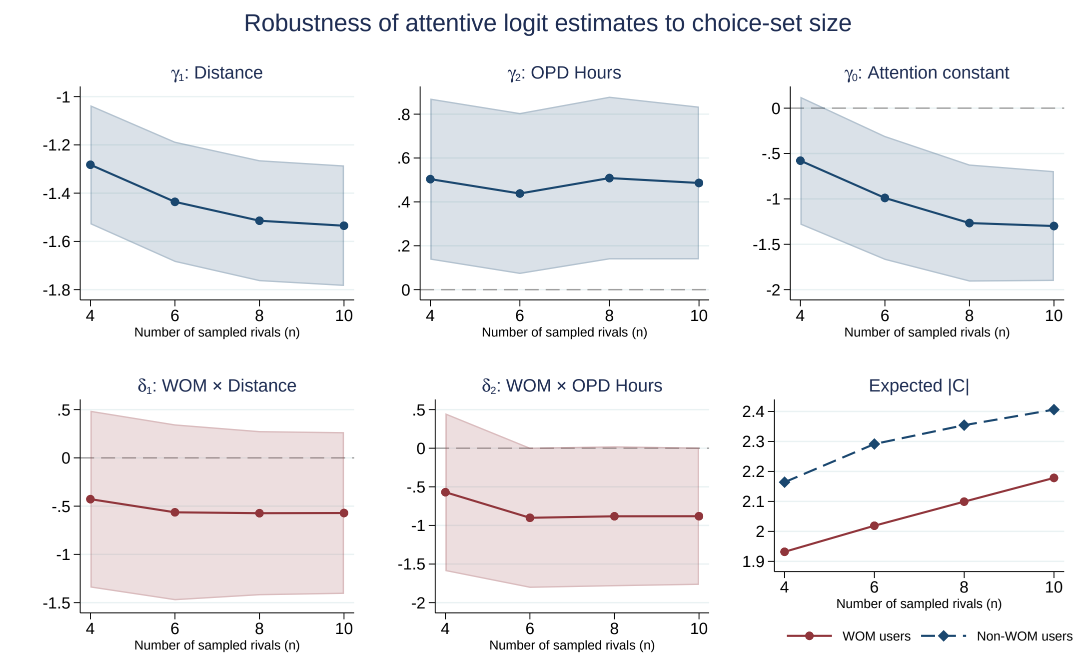


Figure A3. Robustness of Attentive Logit Estimates to Choice-Set Size

*Notes:* Each panel reports an attention-stage parameter (or implied $E[|\mathcal{C}|]$) from the preferred attentive logit specification estimated at varying choice-set sizes $n$, where $n$ is the number of randomly sampled non-default rivals per individual. Shaded bands are 95% confidence intervals based on robust standard errors. The estimation sample, default hospital, and chosen alternative are held fixed; only the number of additional sampled rivals varies. Each individual's potential rivals are ranked once by a fixed random draw, so the choice set at smaller $n$ is nested within the choice set at larger $n$. Expected consideration set size in the bottom-right panel is computed in closed form from the estimated attention probabilities and includes the default hospital.

## A.6 Descriptive Evidence on Asymmetric Attribute Responses

The hybrid consideration model is identified from the asymmetric response of choice probabilities to attributes of considered versus unconsidered alternatives (Abaluck and Adams-Prassl, 2021). Under full consideration, choice responds symmetrically to a hospital's attributes regardless of which hospital they belong to. Under limited consideration, choice is more sensitive to attributes of the default (always-considered) alternative, because rival attributes affect choice only when the patient is attentive.

Formally, let $D_i \in \{0, 1\}$ indicate that individual $i$ chose the default (nearest) hospital, and let $k \in \{\text{Hours}, \text{Drugs}, \text{Quality}\}$ index the three facility attributes examined in Table A8. Let $X^d_{ik}$ denote attribute $k$ at individual $i$'s default hospital, and let $\bar{X}^r_{ik}$ denote the mean of attribute $k$ across the non-default (rival) hospitals in $i$'s choice set. We estimate the linear probability model

$$D_i \;=\; \alpha \;+\; \sum_k \gamma_{d,k}\, X^d_{ik} \;+\; \sum_k \gamma_{r,k}\, \bar{X}^r_{ik} \;+\; \delta_1\, \text{Distance}^d_i \;+\; \delta_2\, \overline{\text{Distance}}^r_i \;+\; \varepsilon_i, \qquad (5)$$

where $\gamma_{d,k}$ is the coefficient on the default hospital's attribute $k$ and $\gamma_{r,k}$ is the coefficient on the rival-mean of attribute $k$. Columns 2–3 of Table A8 augment equation (5) with $W_i$, our word-of-mouth indicator, entered both on its own and interacted with each of the six attribute terms, allowing all of $\gamma_{d,k}$ and $\gamma_{r,k}$ to differ for WOM users; column 3 additionally includes district fixed effects. Attributes (hours, drug availability, quality) are standardized so that default and rival coefficients are directly comparable. Distance enters as a control ($\delta_1$, $\delta_2$) but is not tested, since the default is defined as the nearest hospital. For a given attribute $k$, we write $\gamma_d \equiv \gamma_{d,k}$ and $\gamma_r \equiv \gamma_{r,k}$ below; the symmetry restriction implied by full consideration is $\gamma_d = -\gamma_r$, which we test attribute-by-attribute and, in the "Joint asymmetry restriction" row of Table A8, jointly across all three attributes.

To examine this pattern in our data, we collapse the choice sample to one observation per individual and regress an indicator for choosing the default (nearest) hospital on the default's attributes and the mean of the corresponding rival attributes. Attributes (hours, drug availability, quality) are standardized so that default and rival coefficients are directly comparable. Distance enters as a control but is not tested, since the default is defined as the nearest hospital. The symmetry restriction implied by full consideration is $\gamma_d = -\gamma_r$.

Table A8 reports the results. In the pooled specification (column 1), the signs are in the predicted direction (default attributes affect positively on choice of default while rival attributes affect negatively), but the joint test of $\gamma_d = -\gamma_r$ does not reject ($p = 0.51$). The pooled data are not informative enough on their own to discriminate full from limited consideration. A more informative pattern emerges when responses are allowed to vary with WOM (columns 2 and 3).

The default $\times$ WOM and rival $\times$ WOM interaction terms are jointly significant (column 3). The default-attribute interactions are uniformly negative: WOM users place less weight on the attributes of the default hospital across all three dimensions. The rival-attribute interactions are mixed. Rival hours enters positively, as predicted if WOM users' choices are less anchored to the default; rival drug availability is null; and rival quality improvement enters negatively and significantly, indicating that WOM users are less responsive to rival quality than non-users. This last coefficient is inconsistent with a uniform reweighting toward rival attributes, but consistent with the broader pattern of narrowed attention: WOM users respond less to observable quality signals wherever they appear.

For example, a one standard deviation increase in default OPD hours raises the probability that a non-WOM user chooses the default by 9.6 percentage points but only by 0.1 percentage points for a WOM user. The choice model in Section 4 quantifies this WOM-driven narrowing of consideration within a unified framework.

Table A8. Descriptive Test of Asymmetric Attribute Responses

| | (1) Pooled | (2) + WOM Interactions | (3) + District FE |
|---|---|---|---|
| *Panel A. Default and rival attributes* | | | |
| OPD Hours (default) | 0.085*** | 0.092*** | 0.096*** |
| | (0.020) | (0.022) | (0.024) |
| Drug Availability (default) | -0.009 | -0.007 | -0.017 |
| | (0.020) | (0.021) | (0.021) |
| Quality Improvement (default) | 0.054** | 0.068*** | 0.051** |
| | (0.024) | (0.025) | (0.024) |
| OPD Hours (rival mean) | -0.057*** | -0.066*** | -0.054*** |
| | (0.014) | (0.014) | (0.017) |
| Drug Availability (rival mean) | -0.019 | -0.017 | -0.035 |
| | (0.024) | (0.025) | (0.033) |
| Quality Improvement (rival mean) | -0.020 | -0.011 | -0.063*** |
| | (0.017) | (0.017) | (0.023) |
| *Panel B. WOM interactions* | | | |
| OPD Hours (default) × WOM | | -0.088** | -0.095** |
| | | (0.042) | (0.042) |
| Drug Availability (default) × WOM | | -0.076 | -0.082 |
| | | (0.056) | (0.054) |
| Quality Improvement (default) × WOM | | -0.097* | -0.106** |
| | | (0.056) | (0.052) |
| OPD Hours (rival mean) × WOM | | 0.130* | 0.126** |
| | | (0.067) | (0.062) |
| Drug Availability (rival mean) × WOM | | 0.037 | 0.038 |
| | | (0.059) | (0.057) |
| Quality Improvement (rival mean) × WOM | | -0.093** | -0.099** |
| | | (0.043) | (0.042) |
| *Panel C. WOM main effect* | | | |
| WOM | | 0.067 | 0.070 |
| | | (0.050) | (0.048) |
| *Hypothesis tests* ($p$-values) | | | |
| Hours: $\gamma_d = -\gamma_r$ | 0.258 | 0.304 | 0.202 |
| Drugs: $\gamma_d = -\gamma_r$ | 0.309 | 0.394 | 0.160 |
| Quality: $\gamma_d = -\gamma_r$ | 0.295 | 0.079 | 0.732 |
| Joint asymmetry restriction | 0.506 | 0.266 | 0.447 |
| Default × WOM interactions = 0 | | 0.005 | 0.003 |
| Rival × WOM interactions = 0 | | 0.039 | 0.027 |
| $N$ | 533 | 533 | 533 |
| $R^2$ | 0.314 | 0.337 | 0.362 |
| District fixed effects | No | No | Yes |

*Notes:* Linear probability model. The unit of observation is an individual in the choice estimation sample. The outcome equals one if the individual chose the default (nearest) hospital. Default and rival attributes (OPD hours, drug availability, quality improvement) are standardized to z-scores; rival attributes are computed as the mean across non-default alternatives in the individual's choice set. Distance of the default and mean distance of rivals are included as controls but are not reported. The asymmetry restriction $\gamma_d = -\gamma_r$ implied by full consideration is reported in the lower panel both attribute-by-attribute and jointly. Significant individual coefficients are consistent with non-zero attribute responses; the asymmetry test evaluates the restriction $\gamma_d + \gamma_r = 0$. Heteroskedasticity-robust (HC1) standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.