Quality Consideration and Universal Healthcare*

Dibya Deepta Mishra

August 2, 2026

Preliminary Draft. Please do not cite or circulate.

Abstract

This study examines barriers to universal health coverage in India, focusing on limited information as a key obstacle. Using extensive healthcare insurance claims data and a discrete choice model with limited consideration, I find that patients overlook high-quality hospitals despite valuing both quality and proximity. I measure quality by risk-adjusted survival, using a cross-fitted neural network to predict survival from patient and case mix and treating the hospital-level residual as quality. Patient choices respond to this measure and to Google ratings independently, but the two are uncorrelated with each other, so the publicly available signal does not track clinical outcomes. My results show significant heterogeneity across demographic groups, with minority caste and female patients operating under more restricted choice sets. Through counterfactual analyses, I evaluate two policy interventions: strategic expansion of specialized cardiac services and targeted information provision through mobile technology. Simulations suggest that ensuring patients consider the highest quality hospital in their region would raise risk-adjusted survival by 0.11 percentage points, a relative reduction of roughly 5%, with larger gains for vulnerable populations, while an otherwise identical nudge toward the highest rated hospital delivers none of that gain. The study demonstrates how combining infrastructure investment with digital information dissemination could enhance healthcare access while maintaining cost-effectiveness.

Keywords: Universal health coverage; Limited consideration; Hospital quality; Risk adjustment; India
JEL codes: I11, I13, I18, L15, O12

*Mishra: Department of Economics, Rice University. E-mail: ddm5@rice.edu. I thank Rossella Calvi, Maura Coughlin, Jeremy Fox, Arun Gopalakrishnan, Tarun Jain, Debi Prasad Mohapatra, Isabelle Perrigne, Robin Sickles, Xun Tang, and Matthew Thirkettle for their support and guidance. I also thank participants at the IO Sessions in the Southern Economic Association Annual Meeting, Texas Econometrics Camp, and American Society of Health Economists Conference for their comments. All errors are my own.

1 Introduction

Universal health coverage is one of the seventeen Sustainable Development Goals adopted by the United Nations General Assembly in 2015 (Barredo, Agyepong, Liu, and Reddy, 2015), and developing countries have adopted such programs at scale (World Bank, 2018). Enrolling a household is not the same as improving its health. Once care is free at the point of use, the household still has to decide where to go, and the returns to coverage depend on whether it goes somewhere good. Evidence that coverage translates into better outcomes is mixed (Yu, 2015; Mohanan, Hay, and Mor, 2016). Malani, Holtzman, Imai, Kinnan, Miller, Swaminathan, Voena, Woda, and Conti (2021), in a large randomized trial of hospital insurance eligibility in Karnataka, find that utilization rises and then fades, and suggest that households may not know which hospitals participate.

This paper asks whether patients covered by a large public insurance program choose well among the hospitals available to them, and what stops them if they do not. The answer matters for how one values coverage. If patients already select the best hospital they can reach, then expanding capacity is the binding policy lever. If instead they never evaluate most of their options, then information is the binding lever, and it is far cheaper.

Distinguishing the two is difficult because both produce the same observation. A patient who bypasses a nearby high-quality hospital may dislike something about it that the econometrician does not see, or may never have known it existed. Under the standard discrete choice model, where every alternative is assumed to be evaluated, the second possibility is ruled out by assumption and is absorbed into estimated preferences. The parameters then confound taste with awareness, and any counterfactual that changes what people know is uninformative.

I address this with a model of limited consideration. Each patient draws a consideration set, evaluates only the hospitals in it, and chooses the best of those. Consideration is not observed. Identification comes from two-way exclusion restrictions in the sense of Goeree (2008): a variable that shifts consideration but not utility, and variables that shift utility but not consideration. My consideration shifter is the number of prior surgical visits to a hospital from the patient's own village. A neighbor's previous admission plausibly tells a household that a hospital exists and treats people like them, without changing how much that household values the hospital

conditional on knowing about it. Hospital characteristics such as rating, bed capacity and distance play the reverse role.

The Indian context makes the stakes concrete. Some 86 percent of rural and 82 percent of urban households hold no health insurance, and out-of-pocket spending is regressive: households in the bottom income quintile devote 6.5 percent of total expenditure to outpatient care, against 3.5 percent in the top quintile (NSS, 2018; Sarwal and Kumar, 2021). Coverage expansions can compress these gaps, as Buettgens, Blavin, and Pan (2021) document for the Affordable Care Act, but only through take-up and utilization rather than eligibility alone.

My setting is Aarogyasri, the public insurance program in Andhra Pradesh, which covers roughly 20 million households with no premiums, no deductibles and no copayments. Two features make it well suited to this question. First, prices are set administratively and balance billing is prohibited, so the usual confound between price and quality is absent and choice is driven by quality, distance and information alone. Second, I observe the universe of claims rather than a sample, about 300,000 a year between 2014 and 2019, which lets me build a quality measure from outcomes rather than relying on a survey or a rating.

Measuring quality is the first substantive step, and it produces a result I did not expect. I estimate expected survival for each claim with a small feedforward neural network using only patient and case-mix characteristics, deliberately excluding hospital identity, and take the hospital-level average of observed minus expected survival as quality. Expected survival is constructed by five-fold cross-fitting, so no claim contributes to the model that predicts it, and hospital averages are shrunk toward the mean by empirical Bayes in proportion to how precisely each is measured. The measure has a split-half reliability of 0.69 and a standard deviation across hospitals of 1.38 percentage points of survival, against a mean mortality rate of 1.9 percent. It is a risk-adjusted outcome measure, not a causal effect of attending a given hospital, and Section 3 states precisely what it does and does not identify. This follows Chandra and Staiger (2016) in spirit, replacing their linear specification with one that does not impose functional form on the risk adjustment.

The unexpected result is that this measure is uncorrelated with the Google rating of the same hospital. The rank correlation is 0.06 and not significant, and the efficient estimator bounds the association at 0.11 percentage points of survival per rating point, with a confidence interval that excludes anything larger than a quarter

of one standard deviation of quality. The null is robust to the number of ratings a hospital has, to how reliably its quality is estimated, and to dropping the risk adjustment entirely. The publicly available signal and the clinical outcome are close to orthogonal.

They nonetheless both predict where patients go. In the choice model, the rating and risk-adjusted quality enter jointly and each is significant. Quality there is measured on claims through 2017 while the choices are from 2018, so the regressor is predetermined. Patients are therefore acting on information about clinical quality that the published rating does not carry, presumably through referral networks and local reputation. Willingness to travel quantifies the trade-off: patients will travel 16 to 25 kilometers for a one-point higher rating, and 6 to 12 kilometers for a one standard deviation improvement in risk-adjusted survival.

The consideration parameter is large. A hospital previously used by someone in a patient's village is far more likely to be evaluated at all, with an estimated coefficient of 0.356. Consideration sets are most restricted for minority caste and female patients, who are also the groups whose choices respond least to the published rating. Ignoring this, as Abaluck and Gruber (2011) argue in the context of Medicare Part D, misstates both preferences and the value of default or recommendation policies.

I run two counterfactuals. The first places one hospital in every patient's consideration set with certainty, which is what a recommendation delivered through an existing government health application would do. Directing patients to the highest rated hospital raises the expected rating of the hospital they reach by 0.84 percent and leaves risk-adjusted survival slightly worse. Directing them to the highest quality hospital raises expected risk-adjusted survival by 0.105 percentage points, a relative reduction of about 5 percent, and helps minority caste patients most. Both interventions cost the same. They differ only in what the platform discloses. The second counterfactual asks which hospitals the government should induce to offer coronary angioplasty, and solves that allocation problem by exhaustive search over candidate facilities.

This paper contributes to three literatures. The first is the empirical literature on limited consideration. Methodologically, work in this area divides into approaches that difference out unobserved choice sets and approaches that integrate over them following Manski and Lerman (1977); I take the latter route (Crawford et al., 2021). Recent theoretical work has clarified what choice data alone can reveal about con-

sideration (Manzini and Mariotti, 2014; Cattaneo et al., 2020), though it has mostly been applied where the researcher can manipulate the choice set experimentally (Aguiar et al., 2023). Identification has since been extended to settings with minimal variation in excluded regressors (Barseghyan et al., 2021; Abaluck and Adams-Prassl, 2021); those results are more general than what I use, and I adopt the simpler exclusion-restriction approach because my setting supplies a credible consideration shifter. Empirically, limited consideration rationalizes patterns that full-consideration models struggle with, including unstable risk preferences in Medicare Part D (Heiss et al., 2013), asymmetric weighting of premiums and out-of-pocket costs (Abaluck and Gruber, 2011, 2016), dominated choices in auto insurance (Barseghyan et al., 2021), heuristic shortcuts on health insurance exchanges (Ericson and Starc, 2012), gains from product standardization (Ericson and Starc, 2016), and the finding that beneficiaries facing 46 Medicare Part D plans consider no more than five (Coughlin, 2019). Recommendation design under limited attention has been studied in retail (Kawaguchi et al., 2021). I bring this apparatus to a developing-country public insurance program, where the welfare stakes are mortality rather than premiums.

The second is the literature on information and choice in health care. Hibbard and Peters (2003) document that more information does not reliably produce better decisions. Gaynor, Propper, and Seiler (2016) use a reform of the English National Health Service to show that relaxing choice constraints raised welfare, and emphasize that what mattered was information about options rather than their number. Debnath and Jain (2020) identify a social-network channel that raises awareness of an Indian program, which is the channel my consideration shifter is designed to capture. My contribution is to show that the quality signal patients can actually observe is uninformative about survival, so the standard policy of disclosing ratings does not by itself improve outcomes.

The third is the literature measuring provider quality from administrative data, following Chandra and Staiger (2016) and Chandra, Cutler, and Song (2011), and the broader discrete-choice tradition in industrial organization (Petrin, 2002). I contribute a cross-fitted, machine-learned risk adjuster with an explicit reliability correction, and show that the resulting measure behaves sensibly in a demand model: it predicts choices better than either the published rating or a conventional linear fixed-effect measure.

The remainder of this paper is organized as follows. Section 2 describes the institutional setting, and Section 3 the data and the construction of the quality measure. Section 4 presents the model, identification and estimation. Section 5 reports the estimates. Sections 6 and 7 present the two counterfactuals, and Section 8 concludes.

2 Institutional Background

India has historically shouldered a substantial portion of its health expenditure through out-of-pocket payments borne by patients (World Bank, 2024). The average out-of-pocket expenditure (OOPE) per capita in India is strikingly high, reaching US $299.3 for inpatient services and $390.7 for outpatient services (Ambade et al., 2022). Annual per capita income is $2,250, so a single inpatient episode can absorb more than a tenth of it. Public insurance for low-income households, along the lines of Medicaid in the United States, is the standard policy response.

In 2007 Andhra Pradesh became the first Indian state to run its own government-sponsored health insurance program, the Rajiv Aarogyasri Community Health Insurance. This program specifically targets households below the poverty line. However, owing to the state's relatively high poverty line, a significant portion of the population effectively falls under its coverage (Debnath and Jain, 2020).

Aarogyasri ("*absence of disease*") also covers some households above the poverty line, namely holders of an Annapurna card, an Anthyodaya Anna Yojana card, or a Journalist card. Enrollment is automatic for any eligible cardholder and costs the household nothing. Coverage is for inpatient care, concentrated in tertiary services.

There are no co-payments for covered procedures and providers may not balance-bill. Patients sign a statement at discharge confirming they paid nothing out of pocket, transport included. Financing comes entirely from general state revenue. This matters for what follows: because the price a patient faces is zero at every hospital, the choice among them cannot be driven by price, and the usual confound between price and quality is absent.

Public and private hospitals apply for empanelment and not all are accepted. A hospital with fewer than 50 beds is disqualified. An empanelled hospital can be removed if it fails the performance standards in its contract, so the set of available hospitals changes over time.

Initial coverage was about $3,000 per household below the poverty line, paid

directly to the hospital. Take-up was high, and patients travelled long distances for treatment (Debnath and Jain, 2020).

Andhra Pradesh was reorganized in 2014, with 10 of its 23 districts forming the new state of Telangana. The scheme continued in both states under separate names. I study the residual Andhra Pradesh for 2014 to 2019.

The program targets households without regular access to care and covers roughly 20 million families. Eligible families are provided an eligibility card based on their income, and only in-patient coverage is provided for 938 listed surgeries. The program coverage period is one year at a time. There are no premiums and no deductibles. Patients choose which hospital to seek treatment, and are currently provided an annual amount of around $7000 that can be used towards treatment.

Empanelment required roughly 50 beds, specified infrastructure, and a minimum number of duty doctors. The hospital should earmark a space of 50 sqft. for a dedicated kiosk related to the scheme and have process flows for biometric authentication of cardholders. They must also conduct health camps from time to time. Hospitals need to fill in the details of the surgery and submit them to the government to get reimbursed for their costs.

A significant milestone was reached in 2017 with the launch of "Ayushman Bharat" ("*blessed with a long life*"), a pan-India government insurance program. Aarogyasri predates it by a decade, which is why I study Aarogyasri: its operational processes and its data are settled. The success of this model is evidenced by its substantial utilization, with approximately 3 million procedures conducted in its first 8 years of operation (Debnath and Jain, 2020). The program's digital infrastructure, managed by a dedicated trust, supports broad and consistent data collection, making it a suitable subject for empirical analysis.

3 Data

I use administrative claims covering the universe of program users between 2014 and 2019, roughly 300,000 claims a year, and supplement them with hospital characteristics collected through the Google Maps API.

Administrative Claims Data: The claims record patient demographics (age, gender, caste, and village of residence) alongside the surgery claimed, admission and

discharge dates, the treating hospital and its location, the amount claimed, the amount reimbursed, and an indicator for in-hospital death.

The dataset allows for longitudinal analyses, with de-identified patient and household IDs enabling the construction of individual health histories and household health indices. However, in this work I focus on a single year and diagnosis.

Three maps describe where the program operates and where care is supplied. Figure 1 shows the study state, Andhra Pradesh, within India. It sits on the south-eastern coast, covers roughly 160,000 square kilometers, and had a population of about 49 million at the 2011 census, so it is comparable in area to Georgia and in population to Spain. The program runs statewide.

Figure 2 plots every hospital empanelled in the program, with district shading giving the count per district. Supply is concentrated. Hospitals cluster in a handful of urban districts along the coast and in the south, while several interior districts have very few. Figure 3 overlays the villages that patients travel from. Demand is dispersed in a way that supply is not: patients originate across the whole state, including from districts with almost no empanelled capacity.

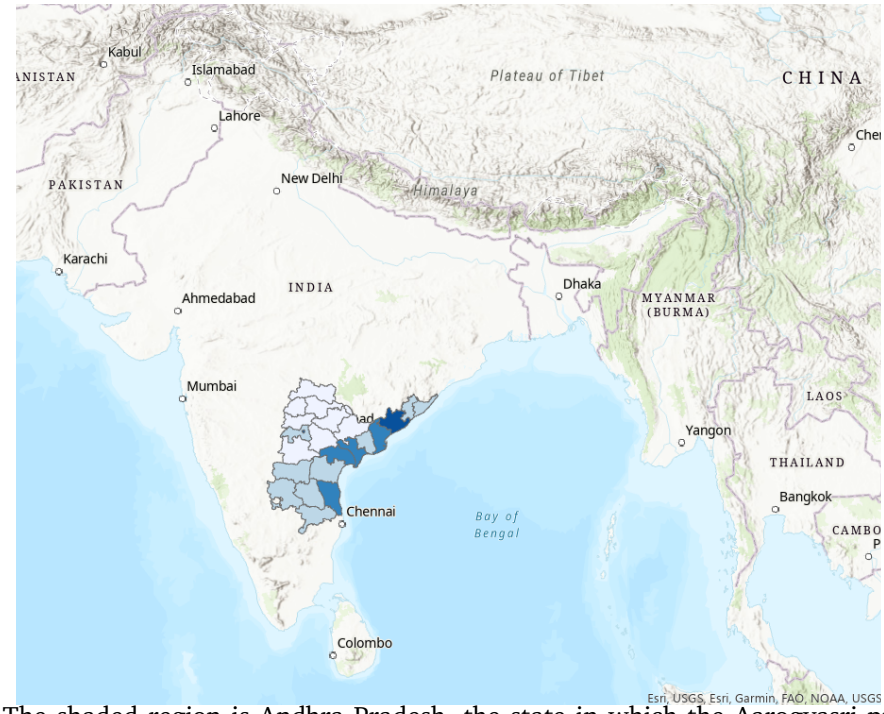
This mismatch does two things for the analysis. It generates the long travel distances documented below, with an average of about 300 kilometers to the hospitals in a patient's choice set. And because each village sits at a different point relative to that uneven supply, the set of hospitals within any given distance differs sharply from one patient to the next. That variation is what identifies how patients trade distance against quality.

The hospital-related information is equally detailed, covering bed capacity, address, and mortality rate from the previous year. Additionally, the dataset includes information on specialties declared by empanelled hospitals, as well as a list of hospitals removed from the program for suboptimal care quality.

Focus on Coronary Angioplasty The question concerns deliberate choice, so the procedure has to be one where patients exercise it. A fracture sends a patient to whichever hospital is reachable. A coronary angioplasty is typically planned, which leaves room for a household to weigh alternatives. I therefore estimate the choice model on all claims for coronary angioplasty in 2018. Restricting to one procedure and one year is for tractability and the framework extends to others.

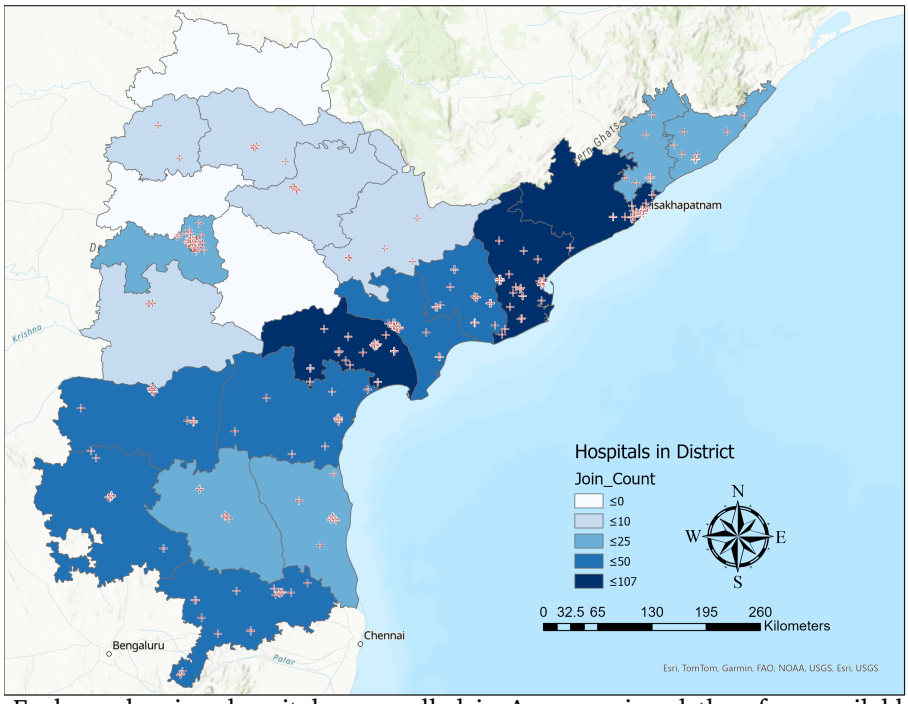
My focus on coronary angioplasty is motivated by both its clinical significance

Figure 1. Location of the program: Andhra Pradesh within India



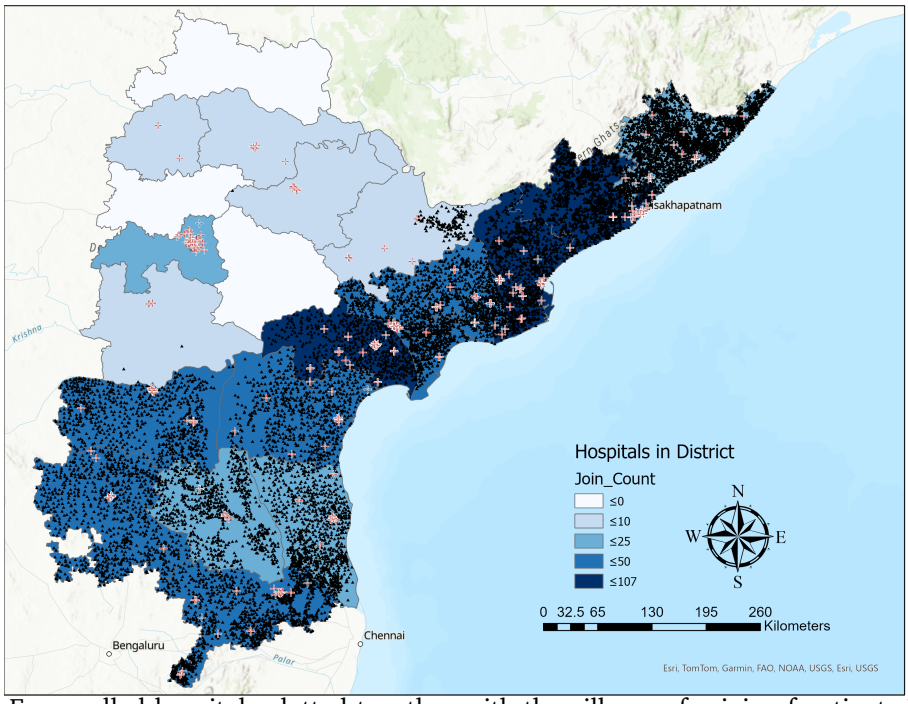
Note: The shaded region is Andhra Pradesh, the state in which the Aarogyasri program operates and from which all claims in this paper are drawn. *Source*: Indian Census.

Figure 2. Location of hospitals empanelled in the program



Note: Each marker is a hospital empanelled in Aarogyasri and therefore available to a covered patient at no cost. District shading gives the number of empanelled hospitals in the district. Capacity is concentrated in a small number of districts. *Source*: Aarogyasri trust, Government of Andhra Pradesh.

Figure 3. Empanelled hospitals and the villages patients travel from



Note: Empanelled hospitals plotted together with the villages of origin of patients in the claims data. Patient origins are spread across the state while hospitals are clustered, which is the source of both the long travel distances and the across-patient variation in choice sets used for identification. *Source*: Aarogyasri trust, Government of Andhra Pradesh.

and India's unique cardiovascular disease (CVD) burden. Coronary angioplasty is a minimally invasive procedure that restores blood flow to the heart by opening blocked coronary arteries. This procedure represents a critical treatment option for patients with coronary artery disease or those experiencing acute cardiac events. The significance of studying angioplasty access in India is underscored by the country's substantial CVD burden. As documented by Kalra et al. (2023), cardiovascular diseases are responsible for approximately one-third of global deaths as of 2017, with India bearing the highest CVD burden among less developed nations, accounting for 12% of all cases globally.

The Indian CVD epidemic differs from the global pattern in three ways. Onset comes earlier than the global average, case fatality is higher, and more deaths are premature. CVD's contribution to total mortality in India has increased from 15.2% in 1990 to 25% in 2017. Perhaps most alarmingly, the country has witnessed more than a two-fold increase in both ischemic heart disease and stroke prevalence between 1990 and 2016. Angioplasty is therefore a procedure where the choice of hospital plausibly matters for survival, and where the volume of cases is large enough to estimate that choice precisely.

Google Maps API Data: I supplement the claims with data collected through the Google Maps API. I geocode each hospital address in the claims and record its Google rating and the number of ratings behind it. I geocode patient villages the same way, and compute the distance between each village and each hospital from the resulting coordinates. These are straight-line distances. Road distance would be preferable and is available from the same API, and I intend to substitute it.

I use Google reviews as a measure of publicly observable hospital reputation. The program's users are predominantly rural, and the people who write Google reviews are unlikely to be drawn from the same population, so the rating measures what is visible to a prospective patient rather than what the average patient experiences. I treat it accordingly, and construct a separate measure of clinical quality from the claims data below.

Are Google Ratings reliable? My use of Google ratings as a quality metric is further supported by recent research on rating platform differences. While Li and Hecht (2021) identify systematic differences between Google and Yelp ratings in the

restaurant context, finding Google ratings to be approximately 0.7 stars higher on average, their work underscores the widespread use and influence of Google ratings in local search decisions. In healthcare, where alternative rating platforms are less prevalent, Google ratings take on particular significance as a primary source of publicly available quality information. The demonstrated correlation between Google ratings and established healthcare quality metrics (Davlyatov et al., 2023) provides additional validation for my methodological approach, even as I acknowledge potential limitations highlighted by Ellenbogen et al. (2022).

I also record, for each hospital, the number of pharmacies and other health facilities within a five-minute drive, using the Geoapify API. Omarali et al. (2023) show that such proximity measures capture dimensions of service accessibility that hospital characteristics alone miss. Complementary services of this kind may affect both where patients go and how they fare.

Estimating Quality from claims data I now construct risk-adjusted survival metrics as an additional quality measure. Following the methodological framework of Chandra and Staiger (2016), these adjustments help account for patient heterogeneity and provider-specific effects in healthcare delivery. This approach enables me to better isolate hospital quality from confounding factors such as patient selection and case-mix variation.

The idea is to predict how likely a patient was to survive given who they were and what was done to them, and to attribute the gap between that prediction and the outcome to the hospital. I pool claims across all surgery codes, exclude pediatric cases, and standardize gender categories. This leaves 1,004,525 claims across 549 hospitals, with a mortality rate of 1.94%.[1] The surgery dates run from 2013 to 2018; the years under which the source files are labeled are claim years rather than surgery years.

Rather than impose a linear index, I estimate expected survival with a small feed-forward neural network. Let S_{ihpmt} denote survival for individual i treated at hospital

[1] One filter warrants care. The claims record a mortality indicator and, separately, a discharge date. From 2017 onward hospitals largely stopped entering a discharge date when the patient died, and the share of deaths with a blank discharge date rises from near zero in 2016 to 45% in 2017 and 99% in 2018. Screening on a length of stay computed from the two dates therefore drops deaths, and drops them mostly in the later years. Length of stay enters nowhere in what follows, so I do not screen on it. Mortality is then stable across the panel, at 2.68% in 2015, 2.07% in 2016, 1.93% in 2017 and 1.78% in 2018.

h for procedure p in mandal m in year t, and let X_{ipmt} collect age, gender, caste, the total number of visits from the household, mandal-level literacy and employment, distance to the treating hospital and its square, procedure category, and surgery year. The network estimates

$$\widehat{S}_{ipmt} = \Lambda\big(g(X_{ipmt})\big), \tag{1}$$

where g is a two-layer network with 64 and 32 hidden units, rectified linear activations, and dropout of 0.2, and Λ is the logistic link. High-cardinality categorical inputs, namely procedure category and mandal, enter through learned embeddings; the remaining categorical inputs are one-hot encoded. The network is fit by binary cross-entropy with class weights, and the resulting probabilities are recalibrated so that predicted and observed survival agree in the aggregate.

The identity of the treating hospital is deliberately excluded from X_{ipmt}. The network is a risk adjuster, not a predictive model of survival: if hospital indicators entered, hospital quality would be absorbed into the fit and the residuals would carry no information about it.

Expected survival is constructed by five-fold cross-fitting. The sample is partitioned at random into five folds, and each claim's $\widehat{S}_{ipmt}$ comes from a network trained on the other four. Without this, residuals would absorb the network's own overfitting and differences across hospitals would be understated. The held-out area under the receiver operating characteristic curve is 0.831 and is stable across folds, ranging from 0.830 to 0.834.

Hospital quality is then the average gap between what happened and what the risk adjuster expected,

$$q_h = \frac{1}{n_h} \sum_{i \in h} \big(S_{ihpmt} - \widehat{S}_{ipmt}\big), \tag{2}$$

where n_h is the number of claims treated at hospital h. Because n_h ranges from a handful of claims to more than twenty thousand, the raw residual mean is dominated by sampling noise at small hospitals. I therefore report an empirical Bayes version that shrinks each hospital toward zero in proportion to the precision with which its quality is measured,

$$q_h^{EB} = \frac{\tau^2}{\tau^2 + \upsilon_h} q_h, \qquad \upsilon_h = \frac{\bar{p}(1-\bar{p})}{n_h}, \tag{3}$$

where υ_h is the sampling variance of the hospital's mean residual and τ^2 is the variance of true quality across hospitals. I estimate τ^2 by splitting each hospital's claims at random into halves and taking the covariance of the two half-sample means, which identifies the shared component without relying on any distributional assumption about the sampling variance. The estimate implies a standard deviation of true quality of 1.38 percentage points of survival, against a mean survival rate of 98.1%. The split-half correlation is 0.53, which implies a reliability of 0.69 for the full measure, so the shrunken measure is the one I use throughout.

I use the shrunken measure q_h^{EB} as my quality measure. As a robustness check I also compute the hospital fixed effects from a linear probability model for survival with the same controls plus fixed effects for mandal, procedure category, caste, and year, which is the specification used by Chandra and Staiger (2016). The two measures correlate at 0.79 across hospitals meeting my reliability restriction, and the substantive conclusions below are unchanged when the linear measure is substituted.

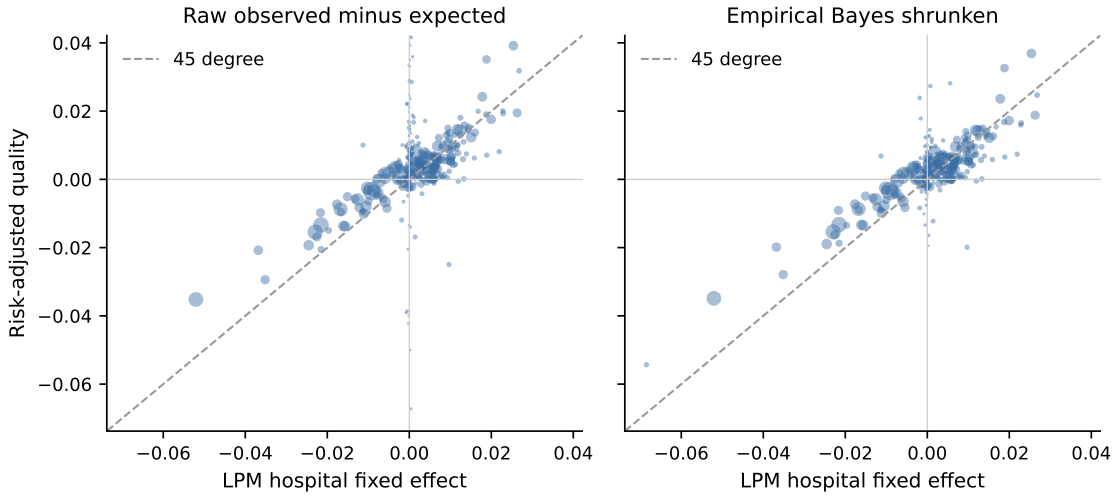
What q_h^{EB} identifies deserves a precise statement. It is the component of survival at hospital h that is not predicted by the observed characteristics in X_{ipmt}. It is a causal measure of hospital quality only under the assumption that assignment of patients to hospitals is unconfounded conditional on X_{ipmt}, which is strong and which I do not claim. Patients choose hospitals, and Section 5 estimates a model in which they do so partly on the basis of quality. If sicker patients sort toward better hospitals on dimensions I do not observe, q_h^{EB} understates the quality of those hospitals; if they sort toward worse ones, it overstates it. The risk adjustment conditions on age, sex, caste, procedure category, mandal, year, and distance, which captures the observable dimensions of case mix, but severity within a procedure category is unobserved in claims data. This is the standard limitation of observed-minus-expected measures and it applies equally to the fixed-effect measure it replaces (Chandra and Staiger, 2016).

I therefore treat q_h^{EB} throughout as a risk-adjusted outcome measure, not as a causal effect of attending hospital h. That is the right object for the questions asked here. When I ask whether the published rating carries information about outcomes,

what matters is the outcome a patient can expect, not the treatment effect of the hospital net of selection. When the measure enters the choice model, it enters as a hospital attribute that patients may or may not respond to.

Figure 4 plots the two measures against one another. The left panel uses the raw residual and the right the shrunken version. Shrinkage matters: small hospitals with extreme raw residuals are pulled toward the mean, and the correlation with the linear measure rises from 0.35 to 0.78. The right panel lies flatter than the 45 degree line by construction, since shrinkage compresses the measure toward zero while leaving the ordering across hospitals intact.

Figure 4. Risk-adjusted quality against the linear probability model fixed effect



Note: Each point is a hospital and marker area is proportional to the square root of its claim count. The dashed line is the 45 degree reference. Left panel, the raw observed-minus-expected residual. Right panel, the empirical Bayes shrunken measure used throughout the paper.

I next ask whether Google ratings track this measure. I regress hospital quality on the Google rating and the number of ratings:

$$q_h^{EB} = \beta_0 + \beta_1 GoogleRating_h + \beta_2 NumRatings_h + \sum_t \beta_t Year_t + \varepsilon_h \tag{4}$$

I restrict this analysis to hospitals with more than 100 ratings to ensure reliability. I also estimate a simpler model without controlling for the number of ratings and surgery year:

$$q_h^{EB} = \beta_0 + \beta_1 GoogleRating_h + \varepsilon_h \tag{5}$$

To further account for potential changes in hospital quality over time, I repeat

the analysis at the hospital-year level, replacing q_h^{EB} with q_{ht}^{EB} constructed from the same residuals grouped by hospital and year.

I present results of this analysis in Table 1. At the hospital level, the coefficient on the Google rating is small and statistically indistinguishable from zero in every specification, with t statistics between 0.58 and 0.94. The rank correlation between the Google rating and risk-adjusted quality is 0.06 and is not significant. The efficient estimator, which weights each hospital by the precision of its quality estimate, puts the association at 0.11 percentage points of survival per rating point with a 95% interval of $[-0.12, 0.33]$, so an effect as large as a quarter of one standard deviation of quality can be ruled out. The same null holds for the linear probability model fixed effects, for the unshrunken residual, and for the unadjusted survival rate, so it is not an artifact of the risk adjustment or of the shrinkage.

These are correlations between two hospital-level measures and nothing more. Neither is randomly assigned and I make no causal claim in either direction. The economically relevant statement is descriptive and is enough for the argument: conditional on nothing, a patient who sorts hospitals by their published rating obtains no information about their risk-adjusted survival.

Three explanations would overturn that reading, and I can address two. Attenuation from noise in the rating would bias the coefficient toward zero, but the ratings are precise: the implied reliability of the rating is 0.85 even under a generous assumption about the dispersion of individual reviews, so attenuation cannot account for a coefficient this close to zero. Noise in the quality measure biases the standard error, not the coefficient, since the measurement error is in the dependent variable. The explanation I cannot rule out is that ratings and quality are related within subgroups in offsetting directions, though I find no such pattern by hospital size, ownership, or rating volume.

These results indicate that Google ratings do not track risk-adjusted survival in this setting. They may still carry information about dimensions patients value and I do not observe, such as waiting times, staff conduct, or non-clinical facilities. Two limitations qualify the result. I collected the ratings in 2024, years after the claims, so time may attenuate a contemporaneous relationship. That objection has force only if hospital quality moves enough over a few years that a 2024 rating, even one that perfectly measured 2024 quality, would say nothing about quality in 2013 to 2017. Two pieces of evidence say it does not.

First, quality is persistent. Estimating the measure separately on 2013-2015 and on 2016-2018 and correlating the two gives 0.63 among the 111 hospitals whose quality is reliably measured in both windows. The early window is short and its measure noisy, so the raw correlation across all 324 hospitals present in both, 0.43, understates persistence; correcting for the estimated reliabilities puts it near 0.9. A rating that tracked quality at any date would track it at all of them.

Second, the association with the rating does not strengthen as the claims approach the date the ratings were collected. Regressing quality on the rating window by window gives 0.178 (standard error 0.160) for 2013-2015 and 0.074 (standard error 0.102) for 2016-2018. Both are indistinguishable from zero, and the point estimate is smaller in the window closer to 2024, which is the opposite of what the timing explanation predicts. Historical ratings would settle the matter and are not retrievable; this is the best available substitute. And the measure captures in-hospital survival alone, a narrow slice of quality at a mortality rate of 1.9%. What survives both caveats is that the signal patients can see and the outcome they care about are different objects. That gap is the information problem this paper studies.

Summary Statistics: I restrict my data to coronary angioplasties only and restrict it to individuals who have undergone only one such surgery in the year 2018.[2] For each individual claim, I first assume that they have a full choice set that implies that each individual considers all empanelled hospitals in the state that offer coronary angioplasty.

For each hospital in the program that offers coronary angioplasty, I create location and rating variables using the application programming interface (API) provided by Google Maps. I also geocode the village of the individuals so that I can generate distance between the village and the hospital. Apart from that I also use hospital specific information such as the bed capacity, and whether or not they are private or public. For each individual, I allow their preferences to be different across gender and caste. Lastly, for each individual, I generate distance between their village and the hospital as well as the number of surgical visits for the hospital in the last years. Furthermore, even though I have data on claim amount, since the prices are set by the government, I assume that the claim amount has minimal impact on hospital choice as there will be no variation. I also note that for almost all households, the

[2]There are less than 0.5% of individuals who underwent more than one procedure in the year.

budget constraint is not violated and hence prices do not impose restrictions on the choice process.

Table 2 presents an overview of summary statistics encompassing all the claims filed for coronary angioplasty recorded in the program for the year 2018. The average patient is around 55 years old but there is significant variation in the ages of individuals, since the youngest patient is a teenager and oldest is around 90. Most patients are male and around a quarter of the patients do not belong to any minority caste denomination.

Google ratings and rating counts both vary widely across hospitals. I assign a rating of 0 to unrated hospitals. The average hospital has around 200 beds but I also have some hospitals have only 13 beds. The largest hospital has around 1500 beds. Most hospitals offering angioplasty are private, which reflects the program's design: empanelment was intended to draw private capacity into serving covered households.

I also see a large variation in distance between villages and hospitals. However, the actual distance travelled might be less since most patients will not go to the farthest hospital. An average of 300 kilometers (180 miles) is nonetheless a long way to travel for care. I also note that there is some variation in the number of surgical visits to hospitals from the village of the individual in the previous year but on an average, it is around 1.

I then construct summary statistics for the hospital-village level variables conditioning on the chosen hospitals only. I find that there is evidence that patients went to closer hospitals as the average individual went to around 70 kilometers ($\sim$ 40 miles) and the maximum was around 800 kilometers ($\sim$ 500 miles) which are less than the distance for the choice set. I also find that chosen hospitals are more likely to have prior surgical visits as well as more surgical visits on an average.

If patients considered only the nearest hospitals, I should observe them bypassing better-rated options. I ask how often a higher-rated hospital was available (as measured by Google Ratings)[3] compared to the choice made by the patients. Table 3 provides the results of this analysis. I note that in around 52% of the cases, there was a higher rated hospital in the entire state. This is to be expected as the patients might be making a distance quality trade off in their choices. Intriguingly,

[3]to account for the measurement error in the ratings, a hospital is considered better than the chosen hospital if the rating is higher by 0.5 rating points

the analysis also reveals instances where patients opt for lower-ranked hospitals despite the availability of higher-quality options nearby. I attribute this behavior to the complexities of limited information and try to explore this in my model.

4 Empirical Strategy

4.1 Full Consideration

Before the notation, the logic in words. A patient ends up at some hospital. I observe that choice, the characteristics of every hospital she could have used, and how far each one is. In the standard model I would assume she compared all of them, and pick the preference parameters that make her observed choice most likely. The difficulty is that a hospital she never heard of and a hospital she considered and rejected look identical in the data: in both cases she did not go. Assuming full consideration forces the model to explain non-use as dislike, so estimated preferences absorb whatever awareness the patient lacked.

The limited consideration model separates the two by adding a stage. Each hospital enters the patient's consideration set with some probability, and she chooses the best hospital among those that entered. Both stages are estimated jointly. What makes that possible is a variable that moves the first stage without moving the second. If prior surgeries at a hospital by other people in the patient's village raise the chance she knows about it, but do not change how much she values it once she does, then variation in that variable traces out the consideration stage separately from preferences. This is the exclusion restriction, and I defend it at length below.

4.1.1 Model

I first set out a standard discrete choice model of hospital demand, in which every patient evaluates every hospital. I then relax that assumption.

Consider a patient i in village v seeking treatment for a specific disease d. I omit the subscript for d as I am focusing on a single disease. I assume that the utility of patient i from choosing hospital h to be given by the following equation:

$$u_{ivh} = x_h'\beta + d_{vh}\delta_0 + d_{vh}^2\delta_1 + x_h'z_i\kappa + d_{vh}z_i\lambda_0 + d_{vh}^2z_i\lambda_1 + \epsilon_{ivh}. \tag{6}$$

Here, x_h has hospital specific characteristics specifically, the average of the Google rating, the number of ratings, hospital bed capacity, an indicator for whether the hospital is private or government run, and the risk-adjusted quality measure q_h^{EB} constructed in Section 3. One detail of timing matters. The choice data are patients treated in 2018, and the quality measure reported in Section 3 is estimated on claims through 2018, which includes those same patients: a patient who chose hospital h and survived raises the measured quality of h. Entering that measure as a regressor for her own choice would be simultaneous. For the choice model I therefore re-estimate quality on claims through 2017 only, so the regressor is predetermined. The two versions correlate at 0.93, and the pre-2018 measure has a split-half reliability of 0.71. Predetermination removes the mechanical channel and rules out reverse causality from 2018 choices to earlier outcomes. It does not deliver exogeneity, since a persistent unobserved hospital attribute would affect survival before 2018 and desirability in 2018 alike. Apart from that I also use variation in distance including a quadratic term as well, $d_{ıh}$ and $d_{ıh}^2$.

I allow the preference for hospital characteristics to vary with demographic characteristics of the individuals, specifically gender and caste. I do this by interacting hospital characteristics and distance with demographics. For the sake of brevity, let

$$u_{ıvh} = \zeta_{ıvh}(\theta) + \epsilon_{ıvh}, \tag{7}$$

where $\theta \equiv (\beta, \kappa, \delta, \lambda)$.

Assuming full consideration, the patient chooses the hospital h, that provides them maximum utility. Denoting Y_i as the hospital chosen by individual i, I have $Y_i \in \arg\max u_{ıvh}$. This implies that probability of hospital h being chosen by patient i is given by

$$p_{hi} = \int \mathbf{1}(u_{ıvh} \geq u_{ıvh'}) dF_\epsilon. \tag{8}$$

I assume the demand unobservable $\epsilon_{ıvh}$ is mean-independent of the observed determinants of utility. The content of this assumption is that whatever a patient likes about a hospital beyond its measured characteristics is uncorrelated with those characteristics. It would fail if, for example, hospitals with better unobserved amenities also systematically located closer to the villages I observe patients travelling from. Taking $\epsilon_{ıvh}$ to be independently and identically Type 1 Extreme Value across patients

and hospitals gives

$$p_{hi}(\theta) = \frac{\exp(\zeta_{ivh}(\theta))}{\sum_{h' \in C_i^*} \exp(\zeta_{ivh'}(\theta))} \tag{9}$$

4.1.2 Identification

Identification is standard (Train, 2009). The preference parameters are pinned down by variation in hospital characteristics across the alternatives a given patient faces, and the demographic interactions by variation in who faces which alternatives. Table 2 shows that both sources of variation are substantial: ratings, bed capacity, ownership and distance all vary widely across the 68 hospitals in the choice set, and patients differ in gender and caste. What this design cannot do is separate a low choice probability that reflects distaste from one that reflects unawareness, which is the motivation for the next subsection.

4.1.3 Estimation

The model is estimated by standard maximum likelihood estimation. One caveat applies to inference throughout. The attribute q_h^{EB} is not observed but estimated in a first stage with its own sampling error, and the standard errors reported below treat it as data. They are therefore understated for the coefficients on quality and on its interactions. The empirical Bayes shrinkage attenuates the point estimate toward zero for the same reason, so the quality coefficient is if anything conservative. A bootstrap over both stages is the remedy and is in progress; the estimates below are point estimates with analytic standard errors. The likelihood as a function of the parameters is given by:

$$L(\theta) = \sum_i \sum_{h \in C_i^*} \mathbf{1}(Y_i = h) p_{hi}(\theta) \tag{10}$$

where $p_{hi}(\theta)$ is defined above. Hence $\hat{\theta}_{fc} \in \arg\max L(\theta)$ represents the estimate for preference parameters assuming that the patients consider all hospitals that offer angioplasty when they make a choice.

4.2 Limited Consideration

4.2.1 Model

Recall that in the full consideration approach, the probability of hospital h being chosen by patient i is given by:

$$p_{hi} = \int \mathbf{1}(u_{ivh} \geq u_{ivh'}) dF_{\epsilon}. \tag{11}$$

Table 3 shows that patients frequently travel past a higher-rated hospital to reach the one they use. With 68 hospitals offering coronary angioplasty in the state, the assumption that each patient evaluates all of them is implausible on its face, and the bypassing pattern is what one would expect if they do not.

I use the limited consideration model of Goeree (2008) to explore patterns in the data further. Let C_i^* represent the full choice set, encompassing all available hospitals, and C_i denote the subset of hospitals that patient i actually considers in making their choice. I further assume that the choice set formation is governed by a parametric process with parameters γ to be estimated, $\Pr(CS_i = c|\gamma)$.

Given Y_i denotes the hospital choice of patient i, the probability that patient i chooses hospital h is given by:

$$\Pr(Y_i = h|\theta, \gamma) = \sum_{c_i \subseteq C_i^*} \Pr(Y_i = h|CS_i = c_i, \theta) \cdot \Pr(CS_i = c_i|\gamma). \tag{12}$$

Here, $\Pr(Y_i = h|CS_i = C_i, \theta)$ represents the probability that h is chosen given that the choice set of the individual has already been formed, and $\Pr(CS_i = c|\gamma)$ represents the probability that a particular choice set is faced by the patient.

4.2.2 Technical Assumptions

For identification and estimation of the model, I make the following technical assumptions:

Assumption 1: The demand unobservables (ϵ_{ivh}) are identically distributed and independent across individuals and hospitals, following a Type 1 Extreme Value dis-

tribution. This implies:

$$\Pr(Y_i = h|CS_i = c, \theta) = \frac{\exp(\zeta_{ivh}(\theta))}{\sum_{r\in c}\exp(\zeta_{ivr}(\theta))} \tag{13}$$

Assumption 2: The consideration of each hospital is independent of the consideration of other hospitals. This allows me to express the choice set formation process as:

$$\Pr(CS_i = c|\gamma) = \prod_{l\in c}\phi_{ivl}(\gamma)\prod_{k\notin c}(1-\phi_{ivk}(\gamma)) \tag{14}$$

Assumption 3: The consideration probability follows a logistic form:

$$\phi_{ivlt}(\gamma) = \frac{\exp(W_{vl}\gamma)}{1+\exp(W_{vl}\gamma)} \tag{15}$$

where W_{vl} represents the number of surgical visits from village v to hospital h in the prior year.

Assumption 4: The number of surgical visits in the prior year affects only consideration and not utility (exclusion restriction). I explain this assumption in detail in the section below.

4.2.3 Identification

My aim is to estimate the parameters that govern preferences θ as well as the parameters that govern the choice set formation process γ. Using standard identification arguments from Goeree (2008), a necessary requirement for my model to be identified is that at least one variable that enters patient utility is excluded from the consideration process. Barseghyan et al. (2021) show that this requirement can be weakened considerably: under a single crossing property, identification obtains with minimal variation in one excluded regressor. I do not need that generality here, because my setting supplies a consideration shifter whose exclusion I can defend on institutional grounds. I use the stronger restriction and explain below why I believe it holds.

I use the number of *any surgical visits* across *all surgical categories* from village v to hospital h in the years 2014-2017, denoted W_{vh}, as a shifter of consideration. The

identifying assumption is that

$$\mathbb{E}\left[\epsilon_{ivh} \mid W_{vh}, x_h, d_{vh}\right] = 0, \tag{16}$$

so that prior surgical volume from the patient's village enters the consideration probability $\phi_{ivh}(\gamma)$ but is mean-independent of the unobserved taste ϵ_{ivh} once hospital characteristics and distance are held fixed.

The threat is worth stating before the defence. Hospitals that many people from a village have used may be hospitals that people from that village like, for reasons I do not observe. If so, W_{vh} belongs in utility, the exclusion fails, and what the model attributes to consideration is partly preference. The estimate of γ would be biased away from zero and the counterfactual gains from information overstated. Three features of the setting bear on how serious this is.

First, the shifter is based on *any surgical visit* and not just visits related to cardiac diseases. This would have been especially concerning had it been the case that hospitals in the program specialized in similar surgeries. However, there are significant differences in surgical category mix in hospitals that provide coronary angioplasty services.[4]

For example, consider hospital A that caters exclusively to cardiac patients and hospital B that caters to all types of surgeries. Then hospital A might only have prior visits that are related to cardiac diseases, but hospital B would have visits for multiple surgery codes. Even if hospital B provided less utility to the patient, the patient might have been more likely to hear about the hospital and hence hospital B is likely to be in the consideration set.[5]

Second, coronary angioplasty is planned, but many of the surgeries entering the shifter are not. A fracture or an accident sends a patient to whichever hospital is available, with little exercise of choice. Visits of that kind still make a hospital known in the village while carrying no information about how much a later angioplasty patient would value it.

Two further points bound what the exclusion buys. First, W_{vh} is measured at

[4] In my sample, among hospitals offering coronary angioplasty, I observe substantial heterogeneity in the range of surgical services provided. The median hospital in this group offers more than 50 distinct types of surgical procedures. At the extremes, I find hospitals providing as few as 5 different surgical services, while others offer in excess of 100 distinct procedures.

[5] In fact, this is one of the reasons I believe that individuals may not be choosing closer higher-rated hospitals, as seen in my discussion about summary statistics.

the village-hospital level and is mechanically related to distance: nearby hospitals get used more. Distance and its square enter utility directly, so γ is identified from variation in prior volume *holding distance fixed*, which is the comparison between two hospitals equally far from a village of which one has been used more. Second, the assumption is untestable. I can show it is consistent with the data, and the limited consideration model fits better than full consideration, but a better fit is not a test of an exclusion restriction. Experimental variation in information would be the clean test and none is available here.

I also assume that certain utility terms such as the Google Ratings, proximal amenities, or risk-adjusted survival in the current year do not strongly impact the consideration process. A plausible reason why this is the case is the fact that there are around 70 hospitals in the choice set and the patient is unlikely to know about quality metrics until they start considering the hospital for a planned visit. This assumption might be violated in the case where the consideration itself is dependent on the hospital characteristics. The program's beneficiaries are unlikely to be well informed about published quality metrics: surveying this same program, Rao et al. (2014) find that roughly 80 percent of respondents learned of the program's existence from their neighbors. If households learn that the program exists through their neighbors, it is reasonable that they learn which hospitals participate the same way.

4.2.4 Estimation

Following Manski and Lerman (1977), I integrate over unobserved choice sets rather than differencing them out, and estimate θ and γ jointly by maximum likelihood. The likelihood is

$$L(\theta) = \sum_i \sum_{h \in C_i^*} \mathbf{1}(Y_i = h) \Pr(Y_i = h | \theta, \gamma). \tag{17}$$

Absent misspecification this can be estimated by simulated maximum likelihood, but three obstacles stand in the way (Crawford et al., 2021). I must choose a functional form for $\Pr(cs_i = c \mid \gamma)$, decide which subsets to evaluate, and confront the dimension of the problem: with J hospitals there are 2^{J-1} possible consideration sets, which for $J = 68$ is far beyond enumeration. Assumptions 2 and 3 address the first, and the simulation procedure below addresses the other two.

4.2.5 Tractable Estimation

To facilitate a more tractable estimation of the model parameters, I use a simulated estimator following Goeree (2008). For each individual i, one would approximate $\Pr(Y_i = h|\theta, \gamma)$ by drawing R choice sets $\left\{c_i^r \mid r = 1, \ldots, R\right\}$ according to the probability $\Pr\left[cs = c_{it}^r \mid \gamma\right]$ and then by averaging out across the resulting conditional choice probabilities. In each simulation, a hospital belongs to the choice set of the individual with the probability governed by γ. The details of the choice set are provided explicitly in Algorithm 1. Subsequently, the estimated choice probability can be expressed as:

$$\widehat{\Pr}\left[Y_i = h \mid \theta, \gamma\right] = R^{-1} \sum_{r=1}^{R} \Pr\left[Y_{it} = h \mid CS_i = c_i^r, \theta\right]. \tag{18}$$

Estimating the log-likelihood through direct simulation is computationally challenging due to the necessity of re-simulation for every parameter value, especially given the substantial number of patients in my context. To address this, I use an importance sampling procedure following Goeree (2008) to approximate $\Pr(Y_i = h|\theta, \gamma)$ as

$$\widehat{\Pr}\left[Y_{it} = j_t \mid \theta, \gamma\right] = R^{-1} \sum_{r=1}^{R} \frac{p\left(c_{it}^r\right)}{g\left(c_{it}\right)} \times \Pr\left[Y_{it} = j_r CS_i = c_{it}^r, \theta\right] \tag{19}$$

where $g(.)$ comes from an initial guess of γ. Hence, I do not need to re-simulate choice sets during parameter search.

In summary, I implement the following algorithm:

Algorithm 1 Importance Sampling procedure from Goeree (2008)

1: Set some initial value for γ and call it γ_0.
2: Given γ_0, compute $\phi^0_{ivl} = \phi_{ivl}(\gamma_0) = \frac{\exp(W_{vl}\gamma_0)}{1+\exp(W_{vl}\gamma_0)}$ for each i and alternative l.
3: For each i draw R choice sets $\left\{c^r_i \mid r = 1, \ldots, R\right\}$ from $\phi^0_{ivl}, l \in J$ where J is the full choice set which in my case is all hospitals that someone from the same district has chosen.
4: Each choice set c^r_i can be drawn as follows:
5: Draw an independent uniform $u^r_{ivl} \in [0, 1]$ for each $l \in J$.
6: Alternative $l \in J$ belongs to c^r_i if and only if $u^r_{ivl} \leq \phi^0_{ivl}$.
7: For each i and t, compute the probability of having drawn each of the R choice sets $\left\{c^r_i \mid r = 1, \ldots, R\right\}$ as $g\left(c^r_i\right) = \prod_{l \in c^r_i} \phi^0_{ivl} \prod_{k \notin c^r_i} \left(1 - \phi^0_{ivk}\right)$.
8: Then, given $\left\{c^r_i \mid r = 1, \ldots, R\right\}$ and their "drawing" probabilities $g\left(c^r_i\right)$ for each i, one can proceed to the estimation of θ and γ by simulated maximum likelihood by maximizing $\sum_i \sum_{h \in C^*_i} \mathbf{1}(Y_i = h) \Pr(Y_i = \hat{h} | \theta, \gamma)$

5 Results and Discussion

Patients trade quality against distance, and the terms of that trade differ across demographic groups (Table 4).

The coefficients are utility weights and are not interpretable in levels, so I read them two ways. The first is relative to distance: dividing a characteristic's coefficient by the distance coefficient converts it into the kilometers a patient would travel for one more unit of that characteristic, which is the willingness-to-travel calculation reported below. The second is per standard deviation, which puts characteristics measured on different scales on a common footing. The rating coefficient of 0.630 and the quality coefficient of 20.43 differ in magnitude only because a rating point and a percentage point of survival are different units. Per standard deviation of each measure they are 0.42 and 0.22, so the published rating and measured clinical quality carry comparable weight in the choice.

General caste male patients show clear preferences for higher-rated hospitals (β = 0.630 in the specification including risk-adjusted quality), while exhibiting expected aversion to distance (β = −3.285). They also respond to risk-adjusted quality itself

(β = 20.43, or 0.22 per standard deviation of the measure). Both attributes enter jointly and are close to orthogonal across hospitals.

The natural reading is that patients act on information about clinical quality that the published rating does not carry. That reading is an inference, not an estimate, and it rests on the quality measure being predetermined. Because q_h^{EB} is built from claims through 2017 and the choices are from 2018, the coefficient cannot reflect the mechanical channel by which a patient's own survival enters the quality of the hospital she chose, nor reverse causality from 2018 choices to earlier outcomes. What it can still reflect is a persistent hospital attribute that raises both survival and desirability and that I do not observe, such as physician skill or equipment. Under that alternative, patients are responding to the attribute rather than to quality as such. Both readings imply that something correlated with survival, and orthogonal to the rating, moves demand. The counterfactual in Section 6 requires only that weaker statement.

Minority caste patients weight bed capacity more heavily (β = 0.057 on the interaction) and the Google rating less. They also weight risk-adjusted quality more (β = 5.591, against a base of 20.43). One reading is that they rely on channels other than the published rating, such as referral networks or prior use by neighbors, which carry information the rating does not. That is an interpretation of a heterogeneous coefficient and I cannot distinguish it from the alternative that minority caste patients face a different set of hospitals within any given distance. Female patients favor hospitals with more nearby pharmacies (β = 0.024) and place less weight on the rating.

Allowing for limited consideration changes what the preference parameters mean. The consideration parameter is γ = 0.356 in the specification with quality, so a hospital previously used by someone in the patient's village is far more likely to be evaluated at all. Preferences estimated under full consideration therefore confound taste with awareness.

My willingness-to-travel estimates quantify these trade-offs, showing patients will travel between 16 and 25 kilometers for a one-point increase in hospital rating (Table 5) and between 6 and 12 kilometers for a one standard deviation increase in risk-adjusted quality, which is 1.06 percentage points of survival (Table 6). These are ratios of estimated coefficients and inherit the identifying assumptions of the model; the confidence intervals are delta-method approximations and do not account for the

fact that q_h^{EB} is itself estimated. The wider confidence intervals for minority caste patients in the limited consideration model suggest more heterogeneous preferences in this group, potentially reflecting varied access to information or transportation resources.

These estimates carry two implications. Consideration runs through prior use in the village, so outreach through local networks reaches the margin that binds, and that margin binds hardest for minority caste and female patients. The published rating does not track survival, so raising the salience of well-rated hospitals is a different intervention from raising the salience of hospitals with better outcomes. Section 6 separates them.

I treat consideration as static within a year. Since consideration here is generated by prior use in the village, it should evolve as the program accumulates users, and the panel structure of the claims data would allow that to be estimated directly.

6 Counterfactual Information Nudges

It is worth being precise about what this experiment does. It does not build a hospital, change a price, or alter any hospital's characteristics. It changes one thing: which hospital the patient is certain to have evaluated. Preferences and the rest of the consideration set stay at their estimated values.

The gain is therefore the value of information alone, computed under the model. It is not an experimental estimate and it inherits every assumption above: the exclusion restriction, the logit error, and independence of consideration across hospitals. It is an upper bound on what a real recommendation would deliver, since a real one would not be attended to with certainty.

Two things the exercise holds fixed deserve attention. The first is capacity. The nudge is concentrated: patients in a region share the same best hospital, so 58 percent of patients are pointed at a single facility. If all of them moved, that hospital would need to treat roughly eight times the caseload of the busiest hospital in the data. That is not what the model predicts, because the nudge only guarantees the hospital is evaluated and patients still weigh distance and everything else. Predicted volumes reallocate by 6.8 percent of patients, and the largest gainer roughly doubles, from 301 to 653 cases a year, which is about one additional patient a day and stays below the 820 cases the busiest hospital already handles. The exercise therefore

stays within the range of caseloads this system operates at. A larger intervention, or one in a region with a single dominant provider, would run into capacity and the realised gain would fall short of the number computed here. Modelling that would require a supply side, which this framework does not have.

The second is hospital behavior. If disclosing quality caused hospitals to compete on it, the long-run effect would differ from what this exercise computes, and the sign of that difference is not obvious.

I use the estimates to simulate targeted information provision. The experiment places one hospital in every patient's consideration set with certainty, which is what a recommendation delivered through an existing health app would do.

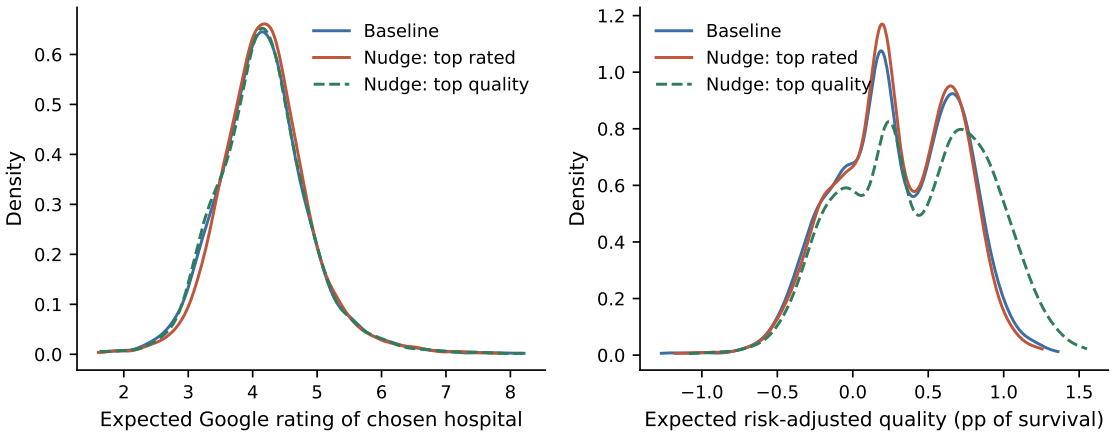
Specifically, for each patient, I modify their consideration set such that $\phi(W_i^{\text{top 1}}) = 1$ for a designated hospital in their region, which enters the consideration set with certainty in both the target and the proposal density. I then compute the expected quality of the hospital the patient ends up at by weighting each hospital's characteristic by its selection probability. I run this twice, once designating the highest Google-rated hospital and once the highest risk-adjusted quality hospital, and report the outcome in both rating points and percentage points of survival.

The results depend on what the nudge points at. Directing patients to the highest rated hospital raises the expected Google rating of the hospital they reach by 0.84%, with minority caste patients gaining most at 0.88% and female patients 0.66%. It does not improve outcomes: expected risk-adjusted quality falls by 0.014 percentage points of survival. This follows directly from the finding in Section 3 that the rating and risk-adjusted survival are uncorrelated, and it is the central caution of this paper for policy.

Directing patients instead to the highest quality hospital in their region raises expected risk-adjusted survival by 0.105 percentage points, against a baseline mortality rate of 2.0%, a relative reduction of roughly 5%. Gains are largest for minority caste patients at 0.110 percentage points and 0.108 for female patients, and the expected rating of the chosen hospital does not move. Figures 5, 6 and 7 show both nudges against the baseline for each group. These improvements are noteworthy given the low implementation costs of digital information provision compared to traditional outreach methods like health camps, but they require the platform to publish a quality measure rather than a rating.

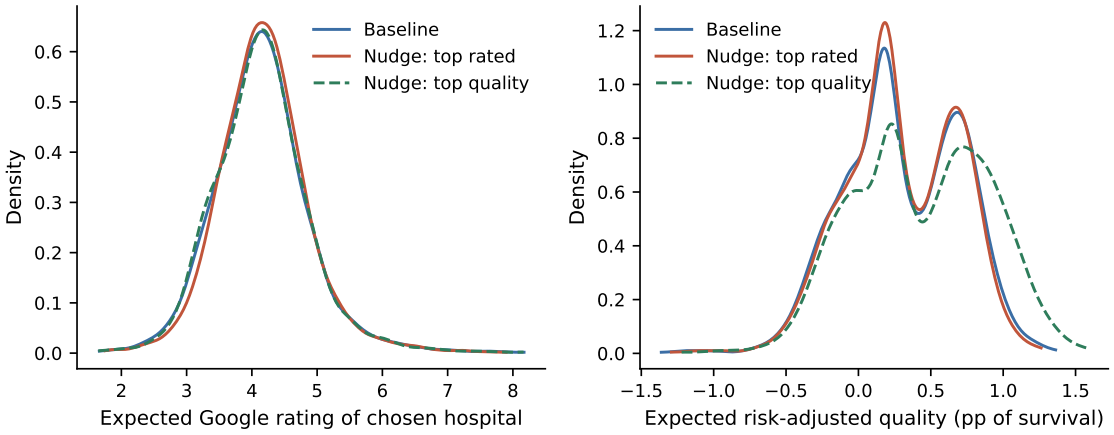
This intervention could be implemented by enhancing existing government health-

Figure 5. Counterfactual 1: all patients



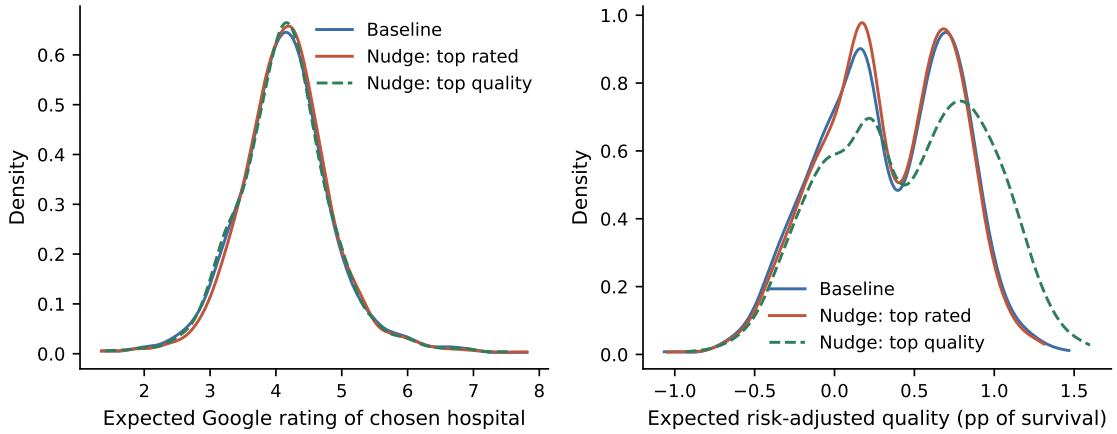
Note: Distribution across patients of the expected characteristic of the hospital reached, under the estimated model (Baseline) and under each nudge. Left panel, Google rating. Right panel, risk-adjusted quality in percentage points of survival. The rating nudge shifts the left panel and leaves the right panel where it was.

Figure 6. Counterfactual 1: minority caste patients



Note: As in Figure 5, restricted to minority caste patients.

Figure 7. Counterfactual 1: female patients



Note: As in Figure 5, restricted to female patients.

care apps to include personalized hospital recommendations based on quality metrics and patient location. The current government health portal already contains basic hospital information, augmenting this with targeted quality information and geographic filtering could help patients make more informed choices while keeping implementation costs minimal. I can also look at welfare increases due to addition of the top n hospitals in the choice set with certainty which is a trivial addition to the method.

7 Counterfactual Hospital Empanelment

I next ask which hospitals the government should induce to offer coronary angioplasty. I use the full consideration estimates, which assume patients evaluate any newly empanelled hospital, so the resulting gains are an upper bound.

I restrict potential candidates to hospitals that have not provided coronary angioplasty in the past five years and have at least 150 beds, yielding 35 eligible facilities. For each individual i, I calculate welfare using the parameter estimates from the full consideration model as:

$$\text{Welfare}_i = \log(1 + \sum_{j \in J} \exp(\mu_{ij}))$$

where J represents all hospitals (existing and potential new). The optimization

problem is formulated as:

$$\begin{aligned} \max_{x} \quad & \sum_{i \in I} \log\left(1 + \sum_{j \in J} \exp(\mu_{ij}) + \sum_{h \in H} x_h \exp(\mu_{ih})\right) \\ \text{subject to} \quad & \sum_{h \in H} x_h = k \\ & x_h \in \{0, 1\} \quad \forall h \in H \end{aligned}$$

where x_h indicates hospital selection and k is the target number of new hospitals. I solve this using both exhaustive search and mixed integer nonlinear programming, finding identical optimal allocations. By using the full consideration estimates, I assume perfect information about the new facilities, which provides an upper bound on potential welfare gains.

The selected hospitals, compared to non-selected candidates, tend to have higher Google ratings (4.2 vs 3.7 average) and slightly better geographic positioning (308 vs 314 km average distance to patients statewide), and are not systematically higher on risk-adjusted quality. Adding these five hospitals raises welfare by 7.0 percent relative to the existing choice set, computed under the full consideration estimates and therefore an upper bound. The optimal set mixes public and private facilities across a range of bed capacities, so quality ratings alone do not determine which hospitals are worth adding. Three of the five sit below the median on risk-adjusted quality. That is not a contradiction: the objective rewards a hospital that is close to patients who currently have no nearby option, and in this allocation geography dominates quality. Table 7 reports the comparison. Targeted expansion of specialized services can therefore widen access while accounting for quality and geographic distribution.

The Government of India has announced plans to upgrade selected district hospitals to medical colleges and expand specialized services. The exercise above gives one way to choose among candidates that accounts for both travel distance and the composition of who is currently underserved.

8 Conclusion

Using claims for coronary angioplasty under India's largest state health insurance program, I show that patients do not evaluate most of the hospitals available to them. Patients value quality and proximity, but restricted consideration leads many to a hospital that is worse on both than an option they never assessed.

The introduction of limited consideration into my empirical framework helps explain previously puzzling choice patterns. The consideration parameter is $\gamma = 0.356$ with a standard error of 0.012, so prior use in a patient's village shapes which hospitals enter the choice set. This effect varies across demographic groups, with minority caste and female patients showing distinct preference patterns that suggest more restricted access to information.

A second finding concerns how quality is measured. I construct a risk-adjusted survival measure from the claims data using a cross-fitted neural network, and find that it is uncorrelated with the Google rating across hospitals. Both nonetheless predict where patients go, and both enter the choice model at conventional significance. Patients therefore appear to act on information about clinical quality that the publicly available rating does not convey. For policy, this means that an intervention which directs patients toward highly rated hospitals is not the same intervention as one which directs them toward hospitals with better survival, and the two should be evaluated separately.

My counterfactual analyses offer two concrete policy recommendations. Extending coronary angioplasty to five well-chosen hospitals raises welfare by 7.0 percent. I select them by optimizing over geographic distribution and quality jointly, which is the same problem the government faces in its current program of upgrading district hospitals through public-private partnerships.

Second, simple information interventions through mobile technology could yield meaningful improvements, but only if they point at the right thing. Ensuring patients consider the top-rated hospital raises the rating of the hospital they choose by 0.84% and leaves risk-adjusted survival slightly worse. Ensuring they consider the highest quality hospital raises expected risk-adjusted survival by 0.105 percentage points, a relative reduction of roughly 5%, with the largest gains for minority caste patients. These improvements could be achieved at relatively low cost by enhancing existing government healthcare apps, provided the information supplied is a quality measure

rather than a star rating.

Three extensions seem worthwhile. The framework applies to a single procedure and could be estimated across diagnosis codes, which would show whether consideration binds more tightly where the procedure is less common. Consideration is static here and could be made dynamic using the panel. And I take hospital quality as given; if hospitals compete on a rating that does not track survival, the supply-side response to disclosure is itself a question.

Coverage is not the binding constraint in this program. What patients know about their options is, and the public signal available to them does not measure the outcome that matters. Expanding capacity and disclosing quality are complements, but only if what is disclosed is quality.

Tables

Table 1. Quality Estimation Results

	(1) hospital	(2) hospital	(3) hospital-year	(4) hospital-year
Hospital Ratings	0.0715	0.0845	0.0287	0.0431
	(0.77)	(0.95)	(0.57)	(0.90)
nratings	0.000002		0.000001	
	(1.11)		(0.96)	
2014.surgery_year			-0.1262	
			(-1.60)	
2015.surgery_year			0.1027	
			(1.85)	
2016.surgery_year			0.1787***	
			(3.65)	
2017.surgery_year			0.1892***	
			(3.46)	
2018.surgery_year			0.1400*	
			(2.23)	
Constant	-0.0384	-0.0853	-0.0654	0.0154
	(-0.09)	(-0.22)	(-0.33)	(0.07)
Surgery Year FE	No	No	Yes	No
Observations	378	378	1,197	1,197
R-squared	0.0038	0.0031	0.0101	0.0009

t statistics in parentheses, Total ratings more than 100

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is the empirical Bayes shrunken risk-adjusted quality measure, in percentage points of survival. Sample is 1,004,525 claims, surgery years 2013-2018, all surgery codes, pediatric cases excluded. Hospitals with more than 100 Google ratings and a nonzero rating. Columns (1) and (2) use one observation per hospital and columns (3) and (4) one per hospital-year, as equations (4) and (5) are written.

Table 2. Summary Statistics: Coronary Angioplasty

	Mean	SD	Min	Max
Patients				
Age of the patient	56.57	10.90	19.00	93.00
Patient is Female	0.30	0.46	0.00	1.00
Patients belong to General Caste	0.29	0.45	0.00	1.00
Hospitals				
Google Rating	4.04	0.89	0.00	5.00
Number of Google Ratings (thousands)	6.03	13.66	0.00	88.64
Bed Capacity of Hospital	154.68	238.70	13.00	1489.00
Hospital is Private	0.94	0.24	0.00	1.00
Hospital - Village				
Distance (in 100s of kms) between hospital and village	2.77	1.76	0.00	9.73
No of surgical visits from the same village in the last year	0.67	5.30	0.00	296.00
Hospital - Village conditional on choice				
Distance (in 100s of kms) between hospital and village	0.67	0.77	0.00	7.84
No of surgical visits from the same village in the last year	7.86	16.97	0.00	296.00
No of Individuals				10,878
No of Hospitals				68
Observations				739,704

Data for the year 2018. There are 10878 cases of Medical Management of Coronary Angioplasty and there are 68 hospitals in the full choice set.

Table 3. Choice and Rating: Coronary Angioplasty

	Mean	S.D.	Min	Max
Conditional on choice				
Distance	0.67	0.77	0.00	7.84
Google Rating of Hospital	4.17	0.74	0.00	5.00
Bed Capacity of Hospital	133.96	198.36	13.00	1489.00
Hospital within 20km of closest hospital	0.63	0.48	0.00	1.00
Is there a better hospital (rating_diff > 0.5)				
In Entire State	0.52	0.50	0.00	1.00
Closer than chosen hospital	0.37	0.48	0.00	1.00

Note: Data for the year 2018. There are 10878 cases of Medical Management of Coronary Angioplasty and there are 68 hospitals in the full choice set.

Table 4. Main Estimates

Variable	Standard Logit	Limited Logit	Logit with Quality	Limited Logit with Quality
Google Rating	0.596	0.642	0.630	0.676
	(0.027)	(0.029)	(0.028)	(0.030)
No of Ratings	0.003	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Bed Capacity	0.024	-0.011	0.037	0.011
	(0.012)	(0.013)	(0.013)	(0.014)
Private Hospital	0.124	0.537	-0.035	0.363
	(0.090)	(0.096)	(0.095)	(0.102)
Distance	-3.274	-2.806	-3.285	-2.846
	(0.191)	(0.207)	(0.194)	(0.211)
Distance ^ 2	0.483	0.297	0.472	0.291
	(0.107)	(0.114)	(0.108)	(0.117)
Hospitals in 5 min drive	0.008	0.007	0.008	0.008
	(0.001)	(0.001)	(0.001)	(0.001)
Pharmacy in 5 min drive	-0.012	-0.005	-0.016	-0.007
	(0.004)	(0.005)	(0.005)	(0.005)
Risk-adjusted Quality			20.425	23.537
			(2.383)	(2.650)
Quality unobserved			-0.687	-0.556
			(0.098)	(0.105)
×_Minority Caste				
Google Rating	-0.052	-0.046	-0.042	-0.025
	(0.030)	(0.032)	(0.031)	(0.033)
No of Ratings	0.006	0.005	0.008	0.006
	(0.002)	(0.002)	(0.002)	(0.002)
Bed Capacity	0.046	0.049	0.063	0.066
	(0.013)	(0.014)	(0.014)	(0.015)
Private Hospital	-0.131	-0.099	-0.240	-0.225
	(0.099)	(0.105)	(0.105)	(0.113)
Distance	0.217	0.127	0.274	0.176
	(0.215)	(0.234)	(0.219)	(0.238)
Distance ^ 2	-0.189	-0.161	-0.220	-0.191
	(0.121)	(0.130)	(0.123)	(0.132)
Hospitals in 5 min drive	0.002	0.002	0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Pharmacy in 5 min drive	0.018	0.018	0.017	0.017
	(0.005)	(0.005)	(0.005)	(0.006)
Risk-adjusted Quality			5.591	6.854
			(2.667)	(2.977)
Quality unobserved			-0.065	-0.097
			(0.115)	(0.123)
×_Female				
Google Rating	-0.071	-0.075	-0.072	-0.079
	(0.029)	(0.031)	(0.030)	(0.032)
No of Ratings	-0.004	-0.006	-0.004	-0.005
	(0.002)	(0.002)	(0.002)	(0.002)
Bed Capacity	0.030	0.031	0.028	0.029
	(0.011)	(0.012)	(0.012)	(0.013)
Private Hospital	0.141	0.126	0.159	0.149
	(0.098)	(0.105)	(0.104)	(0.113)
Distance	-0.088	-0.036	-0.120	-0.043
	(0.216)	(0.234)	(0.220)	(0.240)
Distance ^ 2	0.058	0.033	0.070	0.035
	(0.122)	(0.131)	(0.124)	(0.135)
Hospitals in 5 min drive	0.000	-0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Pharmacy in 5 min drive	0.025	0.027	0.023	0.025
	(0.005)	(0.005)	(0.005)	(0.005)
Risk-adjusted Quality			0.188	0.468
			(2.617)	(2.950)
Quality unobserved			-0.258	-0.228
			(0.126)	(0.134)
Consideration Parameters				
No of previous surgical visits		0.370		0.356
		(0.012)		(0.012)
Constant		-0.198		-0.189
		(0.014)		(0.015)

Note: Data for the year 2018. There are 10878 cases of Medical Management of Coronary Angioplasty and there are 68 hospitals in the full choice set. Standard errors in parentheses. Risk-adjusted Quality is the cross-fitted, empirical-Bayes shrunken observed-minus-expected survival measure of Section 3, in survival-probability units, centred across the choice set and estimated on claims through 2017 so that it is predetermined relative to the 2018 choices. Quality unobserved flags the seven choice-set hospitals with no pre-2018 claims, whose quality is set to the mean. All four columns are estimated on a common sample. Standard errors treat the quality measure as data and are therefore understated for its coefficients.

Table 5. Distance traversed (in kms) for quality increase from rating 3 to 4

	Standard Logit	Limited Logit	Logit with Quality	Limited Logit with Quality
General (Male)	18.734 (15.873, 21.596)	23.475 (19.277, 27.673)	19.725 (16.717, 22.732)	24.356 (20.018, 28.694)
Minority Caste	18.117 (13.706, 22.527)	22.517 (16.284, 28.751)	19.850 (15.021, 24.679)	24.601 (17.817, 31.385)
Female	16.050 (12.282, 19.818)	20.456 (14.854, 26.057)	16.815 (12.913, 20.717)	21.151 (15.386, 26.915)

Note: 95% confidence intervals in parentheses, by the delta method. Distance solves $\beta_q + \delta_0 x + \delta_1 x^2 = 0$ for the smaller root, converted from hundreds of kilometres to kilometres.

Table 6. Distance traversed (in kms) for a one standard deviation increase in risk-adjusted quality

	Logit with Quality	Limited Logit with Quality
General (Male)	6.634 (4.913, 8.355)	8.819 (6.459, 11.178)
Minority Caste	9.201 (6.131, 12.271)	12.082 (7.853, 16.310)
Female	6.464 (3.999, 8.929)	8.868 (5.366, 12.370)

Note: 95% confidence intervals in parentheses, by the delta method. Distance solves $\beta_q + \delta_0 x + \delta_1 x^2 = 0$ for the smaller root, converted from hundreds of kilometres to kilometres. One standard deviation of risk-adjusted quality across the choice set is 1.06 percentage points of survival.

Table 7. Comparison between Chosen and Not Chosen Hospitals

Hospital	Rating	nratings	Bed Cap.	Private	Pharm. 5km	Hosp. 5km	Quality	Avg. Dist.
HS456	4.90	4.09	3.50	No	0	63	0.40	293
HS480	4.37	2.42	5.00	No	4	50	-1.86	440
HS53	3.60	0.18	1.66	No	0	100	-0.65	367
HS751	4.92	0.41	1.63	Yes	3	100	-0.41	183
HS92	3.42	0.25	11.55	No	2	66	0.76	254
Not Chosen	3.73	1.45	4.01	–	5	44	-0.11	314

Note: Rating refers to Google rating (out of 5), nratings is number of ratings in thousands, Bed Cap. is bed capacity in hundreds, Private indicates hospital ownership, Pharm. 5km and Hosp. 5km indicate number of pharmacies and hospitals within 5km radius respectively, Quality is the risk-adjusted quality measure in percentage points of survival, and Avg. Dist. is the average distance to patients in kilometers. Not Chosen represents average values for the remaining 30 candidate hospitals. The optimal set of 5 is found by exhaustive search over all combinations of the 35 candidates, using the full consideration estimates of column (3) of Table 4.

References

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

Abaluck, J. and J. Gruber (2011). Choice inconsistencies among the elderly: evidence from plan choice in the medicare part d program. *American Economic Review 101*(4), 1180–1210.

Abaluck, J. and J. Gruber (2016). Evolving choice inconsistencies in choice of prescription drug insurance. *American Economic Review 106*(8), 2145–84.

Aguiar, V. H., M. J. Boccardi, N. Kashaev, and J. Kim (2023). Random utility and limited consideration. *Quantitative Economics 14*(1), 71–116.

Ambade, M., R. Sarwal, N. Mor, R. Kim, and S. Subramanian (2022). Components of out-of-pocket expenditure and their relative contribution to economic burden of diseases in india. *JAMA Network Open 5*(5), e2210040–e2210040.

Barredo, L., I. Agyepong, G. Liu, and S. Reddy (2015). Ensure healthy lives and promote well-being for all at all ages. *UN Chronicle 51*(4), 9–10.

Barseghyan, L., M. Coughlin, F. Molinari, and J. C. Teitelbaum (2021). Heterogeneous choice sets and preferences. *Econometrica 89*(5), 2015–2048.

Barseghyan, L., F. Molinari, and M. Thirkettle (2021). Discrete choice under risk with limited consideration. *American Economic Review 111*(6), 1972–2006.

Buettgens, M., F. Blavin, and C. Pan (2021). The affordable care act reduced income inequality in the us: Study examines the aca and income inequality. *Health Affairs 40*(1), 121–129.

Cattaneo, M. D., X. Ma, Y. Masatlioglu, and E. Suleymanov (2020). A random attention model. *Journal of Political Economy 128*(7), 2796–2836.

Chandra, A., D. Cutler, and Z. Song (2011). Who ordered that? the economics of treatment choices in medical care. *Handbook of health economics 2*, 397–432.

Chandra, A. and D. Staiger (2016). Sources of inefficiency in healthcare and education. *American Economic Review 106*(5), 383–387.

Coughlin, M. (2019). Insurance choice with non-monetary plan attributes: Limited consideration in medicare part d. Technical report.

Crawford, G. S., R. Griffith, and A. Iaria (2021). A survey of preference estimation with unobserved choice set heterogeneity. *Journal of Econometrics 222*(1), 4–43.

Davlyatov, G., M. He, G. Orewa, H. Qu, and R. Weech-Maldonado (2023). Are hospice google ratings correlated with patient experience scores? evidence from a national hospice study. *American Journal of Hospice and Palliative Medicine® 40*(12), 1365–1370.

Debnath, S. and T. Jain (2020). Social connections and tertiary health-care utilization. *Health economics 29*(4), 464–474.

Ellenbogen, M. I., P. M. Ellenbogen, N. Rim, and D. J. Brotman (2022). Characterizing the relationship between hospital google star ratings, hospital consumer assessment of healthcare providers and systems (hcahps) scores, and quality. *Journal of Patient Experience 9*, 23743735221092604.

Ericson, K. M. and A. Starc (2012). Heuristics and heterogeneity in health insurance exchanges: Evidence from the massachusetts connector. *American Economic Review 102*(3), 493–497.

Ericson, K. M. M. and A. Starc (2016). How product standardization affects choice: Evidence from the massachusetts health insurance exchange. *Journal of Health Economics 50*, 71–85.

Gaynor, M., C. Propper, and S. Seiler (2016). Free to choose? reform, choice, and consideration sets in the english national health service. *American Economic Review 106*(11), 3521–57.

Goeree, M. S. (2008). Limited information and advertising in the us personal computer industry. *Econometrica 76*(5), 1017–1074.

Heiss, F., A. Leive, D. McFadden, and J. Winter (2013). Plan selection in medicare part d: Evidence from administrative data. *Journal of Health Economics 32*(6), 1325–1344.

Hibbard, J. H. and E. Peters (2003). Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice. *Annual review of public health 24*(1), 413–433.

Kalra, A., A. P. Jose, P. Prabhakaran, A. Kumar, A. Agrawal, A. Roy, B. Bhargava, N. Tandon, and D. Prabhakaran (2023). The burgeoning cardiovascular disease epidemic in indians–perspectives on contextual factors and potential solutions. *The Lancet Regional Health-Southeast Asia*.

Kawaguchi, K., K. Uetake, and Y. Watanabe (2021). Designing context-based marketing: Product recommendations under time pressure. *Management Science 67*(9), 5642–5659.

Li, H. and B. Hecht (2021). 3 stars on yelp, 4 stars on google maps: a cross-platform examination of restaurant ratings. *Proceedings of the ACM on Human-Computer Interaction 4*(CSCW3), 1–25.

Malani, A., P. Holtzman, K. Imai, C. Kinnan, M. Miller, S. Swaminathan, A. Voena, B. Woda, and G. Conti (2021). Effect of health insurance in india: a randomized controlled trial. Technical report, National Bureau of Economic Research.

Manski, C. F. and S. R. Lerman (1977). The estimation of choice probabilities from choice based samples. *Econometrica: Journal of the Econometric Society*, 1977–1988.

Manzini, P. and M. Mariotti (2014). Stochastic choice and consideration sets. *Econometrica 82*(3), 1153–1176.

Mohanan, M., K. Hay, and N. Mor (2016). Quality of health care in india: challenges, priorities, and the road ahead. *Health Affairs 35*(10), 1753–1758.

NSS (2018). Nss 71st round-key indicators of social consumption in india: Health.

Omarali, A., H. L. Liu, Y. Xu, N. Farkhondehpay, and Z. J. Tian (2023). The influence of a house's proximity to amenities on educational attainment and demographics in the us.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of political Economy 110*(4), 705–729.

Rao, M., S. Bergkvist, P. V. Singh, K. Anuradha, and S. Amit (2014). Rajiv aarogyasri community health insurance scheme evaluation survey.

Sarwal, R. and A. Kumar (2021). The long road to universal health coverage.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

World Bank (2018). Universal health coverage study series (unico) - world bank.

World Bank (2024). Out-of-pocket expenditure (% of current health expenditure) - india.

Yu, H. (2015). Universal health insurance coverage for 1.3 billion people: what accounts for china's success? *Health policy 119*(9), 1145–1152.