Revealed Preferences for Hospitals*

Amitabh Chandra† Pragya Kakani‡ Alex Philip§ Adam Sacarny¶

September 11, 2026

Abstract

Measuring provider quality is a central challenge in health economics. We leverage the choices of patients to develop a revealed preference measure of hospital quality. We then validate this approach by comparing it to the revealed preferences of expert patients who are doctors or their spouses. Revealed preferences are mostly unexplained by traditional quality measures and highly correlated between non-experts and experts, suggesting that they contain valuable new information about hospital quality. Finally, we identify disparities in access to more-preferred hospitals by socioeconomic factors. Our work establishes a novel and broadly applicable approach for ascertaining provider quality in health care.

*We thank Brad Shapiro, Josh Gottlieb, Michael McWilliams, Pauline Mourot, Andrew Olenski, and conference participants at the 2025 American Society of Health Economics Annual Meeting, the Whistler Health Economics Conference, and departmental seminars at UCSB, UCLA, USC, and the National University of Singapore for their helpful comments and suggestions. We also gratefully acknowledge funding from the Peterson Center on Healthcare and the National Institute on Aging (P01AG032952) for this work. Kakani also gratefully acknowledges funding support from the National Institute on Aging (K01AG092905). This research was approved by the Institutional Review Boards of NBER, Columbia University, and Weill Cornell Medical College.

†Harvard Business School, Harvard Kennedy School, and NBER, achandra@hbs.edu

‡Weill Cornell Medical College, pka4006@med.cornell.edu

§University of California, Berkeley, alex.philip@berkeley.edu

¶Columbia University Mailman School of Public Health and NBER, ajs2102@cumc.columbia.edu

1 Introduction

In health care, measuring provider quality is a central challenge dating to Arrow (1963), who proposed that "uncertainty as to the quality of the product is perhaps more intense here than in any other important commodity." This challenge remains a topic of close study (Kolstad, 2013; Chandra et al., 2016; Doyle, Graves and Gruber, 2019). Rather than reflecting true differences in quality, existing measures may capture unobserved patient characteristics and strategic upcoding of diagnoses (Geweke, Gowrisankaran and Town, 2003; Dafny, 2005). They also may fail to capture provider attributes that patients and their agents value but are not tracked in administrative data, like amenities and care coordination. These limitations reduce the value of existing quality metrics to market participants, including patients, clinicians, and policymakers.

We draw on consumer theory to provide a new and more comprehensive lens on health care quality by estimating patients' revealed preferences for providers. Our approach relies on the insight that when patients bypass one provider to receive treatment at another, it suggests that they – or agents acting on their behalf such as referring physicians or family members – value the latter provider more than the former. This approach, therefore, incorporates all factors considered by the patient and their agents when choosing a provider. It can be applied across many contexts, including those for which quality is difficult to measure by traditional means. Moreover, this approach is difficult to game via patient selection or strategic reporting.

Using this insight, we develop a revealed preference measure of hospital quality. We focus on patients hospitalized for emergency conditions like heart attacks and insured by Medicare; these patients have access to nearly all hospitals at essentially the same out-of-pocket cost. Therefore, their choices can be thought of as reflecting preferences over hospitals, with travel time acting as a cost rather than price. We then estimate hospital preferences using choice models identified by within-patient variation in travel costs. The resulting

measures aggregate across the many dimensions of care that patients value, including those that are difficult to observe.

We find that revealed preferences are only weakly associated with traditional quality rankings, like those produced by Medicare and published in U.S. News and World Report. Revealed preferences are also poorly predicted by publicly reported quality measures like mortality and readmission rates, use of productivity-raising best practices, and patient-reported hospital experience. Together, these measures explain only about one-fourth of variation in revealed preferences. This finding suggests that much of what patients value is not reflected in these metrics. As a result, a revealed preference approach adds information relative to existing quality indicators.

A natural objection is that patients' behavior might reflect choice errors (Abaluck and Gruber, 2011; Handel, 2013; Bhargava, Loewenstein and Sydnor, 2017). For instance, patients might incorrectly treat hospital advertising or brands as signals of quality. To address the concern that rankings are driven by misinformation or mistakes, we next study patients who are experts in the choice at hand: those who are themselves physicians and their spouses, using data linkages developed previously (Kakani, Matecna and Chandra, 2025). These patients are uniquely able to observe hospital quality through professional experience and networks, reducing the risk of choice errors.

We find that the revealed preferences of experts are strikingly similar to those of the broader patient population, with the two measures being highly correlated. The revealed preferences of both groups also have similar conditional correlations with existing quality measures. These patterns suggest that even the revealed preferences of non-expert patients contain meaningful information about hospital quality.

Finally, we study who benefits from access to higher revealed preference hospitals. We find that racial and ethnic minority, low-income, and rural patients are significantly less likely to receive care at high revealed preference hospitals. These results partly reflect that these groups tend to live farther from these hospitals. However, differences persist within small

geographic areas – for instance, Black patients tend to use lower revealed preference hospitals than white patients in the same ZIP code. These patterns mirror disparities documented in outcomes-based quality measures (Chandra, Kakani and Sacarny, 2024). The results suggest that disparities in access to high-quality care extend beyond what is captured by existing quality measures.

We demonstrate the robustness of our results to multiple alternative approaches. For instance, we show very similar results using only observations from non-winter months, when hospital capacity constraints are less binding (Akintoye et al., 2017; Janke, Burke and Haimovich, 2025). To address the possibility that results may differ for conditions where patients have more scope for choice, we also repeat our analyses for hip and knee replacements, the most common elective surgeries among older adults. This analysis reveals similar key findings, including that experts and non-experts have similar revealed preferences across hospitals. Finally, we show our results are nearly identical when we relax constraints on patients' hospital choice sets through a national gravity model (as in Dingel et al., 2023) rather than using fixed geographic hospital markets.

Our analysis contributes to several economic literatures. First, we extend work on health care provider quality measurement by developing a novel, preferences-based measure of quality. A large literature documents substantial variation in performance across providers (e.g. Jha et al., 2005; Chandra et al., 2016; Levine, Landon and Linder, 2019). Recently, studies have adapted a value-added approach from the education literature (see e.g. Koedel, Mihaly and Rockoff, 2015, for a review) to estimate provider quality, particularly for nursing homes (Olenski and Sacher, 2024; Einav, Finkelstein and Mahoney, 2025). Research has validated some observational quality measures by showing that they are similar to quasi-experimental estimates (Doyle, Graves and Gruber, 2019; Hull, 2020; Chandra, Dalton and Staiger, 2023). However, existing measures are still limited to a narrow set of diagnoses and outcomes that are observable. Moreover, for many measures, variations may reflect selection and coding rather than true quality differences (Joynt Maddox et al., 2019; Bastani, Goh and Bayati,

2019; Gupta, 2021; Olenski and Sacher, 2024).

Our approach complements existing quality measurement work by introducing a holistic measure that captures patients' overall valuation of hospitals and is not limited to select clinical measures, each of unclear value to patients and their agents. By showing that traditional measures leave three-fourths of the variation in revealed preferences unexplained, our work suggests that widely used quality measures should be interpreted as partial indicators of a broader, multidimensional concept of quality that patients value. Our approach also lends itself to extensions to other contexts, like outpatient care, where quality measures are even more limited.

Second, we draw on approaches to rank firms and organizations using consumers' revealed preferences and extend these techniques to health care. This literature has developed revealed preference methods to infer latent quality in non-health care settings like higher education and labor markets (Avery et al., 2013; Sorkin, 2018; Rothstein and Yoon, 2024). We adapt this framework to the hospital sector, and in contrast to prior work, we validate revealed preferences of a general population based on their similarity to those of experts. Our approach is similar to Romley and Goldman (2011), which focused on pneumonia patients in Los Angeles to estimate the cost of improving revealed quality.

Third, we contribute to the literature using the empirical choices of experts as consumers to understand the degree to which consumer choices reflect a lack of information. Outside health care, prior work has considered the role of information asymmetry in real estate markets by examining the outcomes of expert consumers, such as realtors (Levitt and Syverson, 2008; Rutherford, Springer and Yavas, 2005). In health care, studies have compared the utilization and health outcomes of expert patients, such as physicians, nurses, and pharmacists, to others. Several studies have found that non-expert patients are more likely to use unnecessary or low-value care than experts (Bronnenberg et al., 2015; Chen, Persson and Polyakova, 2022; Kakani, Matecna and Chandra, 2025; Johnson and Rehavi, 2016). Others have found that the two groups behave similarly (Frakes, Gruber and Jena, 2021; Finkelstein

et al., 2022).

In contrast to prior work on clinical experts as patients, we are the first, to our knowledge, to focus on experts' choices of providers. Our finding that experts and non-experts exhibit similar hospital valuations suggests that patients – perhaps via their agents, including clinical experts – may make better-informed choices over providers than is commonly appreciated. Indeed, prior work has suggested that hospital markets exhibit signposts of competition found in other industries like manufacturing (Chandra et al., 2016). Our findings also contrast with work demonstrating persistent choice errors by patients in other health care markets, such as health insurance (Abaluck and Gruber, 2011; Handel, 2013; Bhargava, Loewenstein and Sydnor, 2017). This may reflect that patient choices over hospitals are informed by medical professionals including physicians and first responders (Doyle, Graves and Gruber, 2019).

Finally, we contribute to the large literature documenting disparities in the quality of providers used by patients across socioeconomic and demographic groups, including race. This work has shown that Black (Bach et al., 2004; Jha et al., 2007; Chandra, Kakani and Sacarny, 2024) and lower-income Medicaid patients (Jha, Orav and Epstein, 2011) are disproportionately treated by lower quality providers. A parallel literature highlights geographic inequities, with rural patients receiving worse quality of care from their local providers (Joynt et al., 2011; Loccoh et al., 2022). We confirm these disparities persist when hospital quality is defined through patients' revealed preferences rather than traditional approaches. This alignment with prior evidence shows that disparities are not an artifact of deficiencies in existing metrics, and that they remain when quality is measured holistically.

2 Background and Approach

2.1 Setting

We use data from the Medicare program, which provides near-universal coverage for older adults (age 65 and up) and certain disabled people in the US. We focus on Original (fee-

for-service) Medicare, where the program pays providers directly. Original Medicare is an ideal setting because its patients have access to nearly all hospitals in the US. Moreover, patients face effectively equivalent out-of-pocket costs regardless of which hospital they choose.[1] These features remove provider networks and out-of-pocket costs as drivers of patients' choices and emphasize the potential for competition over quality instead.

Our main analyses study hospitalizations for five emergency medical conditions that have been the focus of a large health economics literature: acute myocardial infarction (AMI, also called heart attack), heart failure, hip fracture, pneumonia, or stroke. These diagnoses account for nearly one in ten hospitalizations in the US and about 15% of hospitalizations in Medicare (Liang, Moore and Soni, 2020). Though these conditions require immediate hospital treatment, patients still exhibit scope for hospital choice. For instance, we find only about half of patients receive treatment at their nearest hospital, consistent with prior work (Chandra et al., 2016).

2.2 Analytic Approach

We then develop a model to isolate patients' revealed preferences for hospitals. The goal of this model is to operationalize the insight that when patients bypass a close-by hospital for one that is farther away, they are revealing that they prefer the latter hospital over the former. In other words, after accounting for travel costs, econometricians can use patients' choices to infer their implicit valuations of the hospitals in their option sets.

This implies a utility function that takes the following form:

$$U_{ph} = \alpha_h + f\left(traveltime_{ph}; \boldsymbol{\rho}\right) + \varepsilon_{ph}, \tag{1}$$

where U_{ph} is the utility for patient p of a potential hospital choice h, α_h is the incremental

[1] Most Medicare beneficiaries have supplemental plans or Medicaid coverage, curtailing or eliminating their out-of-pocket costs (Cubanski et al., 2018). For those without extra coverage, the Medicare cost-sharing structure for inpatient care is mainly a deductible that is well below the cost of a hospitalization. Thus patients' out-of-pocket costs effectively do not vary across hospitals.

utility of the hospital, $traveltime_{ph}$ is the travel time from the patient to the hospital (entering as a function parameterized by the vector $\boldsymbol{\rho}$), and ε_{ph} is the unobserved component of utility. Patients choose the hospital yielding the greatest utility, i.e.:

$$C_{ph} = U_{ph} > U_{ph'} \; \forall \; h' \in H_m, \tag{2}$$

where C_{ph} indicates their choice and H_m is the set of hospitals in the patient's market m.

We impose additional structure on the utility function and choice set to facilitate estimation. Specifically, we assume ε_{ph} follows a Type 1 extreme value distribution. We also define $f\,()$ as quadratic, so that the marginal cost of travel may vary as times rise, and allow $\boldsymbol{\rho}$ to vary across the five conditions comprising the patient sample. We also restrict the choice set to hospitals in the same geographic market as the patient's residence. The resulting choice probabilities take the form:

$$\mathbb{E}(C_{ph}) = \frac{\exp\left[\alpha_h + f\left(traveltime_{ph}; \boldsymbol{\rho}\right)\right]}{\sum_{h' \in \mathrm{H}_m} \exp\left[\alpha_{h'} + f\left(traveltime_{ph}; \boldsymbol{\rho}\right)\right]}. \tag{3}$$

This is the conditional expectation function of the conditional logit model, which we can then use to estimate the parameters α_h and $\boldsymbol{\rho}$. We estimate the model market-by-market, yielding hospital effect estimates $\hat{\alpha}_h$ and travel cost estimates $\hat{\boldsymbol{\rho}}_m$.

The $\hat{\alpha}_h$ can be interpreted as the estimated utility gained from selecting hospital h compared to an arbitrarily chosen reference hospital. The level and scale of these effects are arbitrary because any affine transformation of all utilities leaves the choices given by equation 2 unchanged. Because we partition the sample by market, each hospital effect is also only identified relative to others within a market. Therefore, to ease interpretation, we transform the effect estimates into a market-specific z-score $\hat{Z}_h$. We refer to this object as our revealed preference measure.

We then regress estimated revealed preferences on hospital covariates, including traditional quality measures, to understand what predicts patients' hospital valuations and

whether the covariates can account for variation in these valuations. Namely, we estimate:

$$\hat{Z}_h = \gamma_m + \mathbf{X}_h\boldsymbol{\beta} + \varphi_h, \tag{4}$$

where γ_m are market effects, since revealed preferences are only identified among hospitals in the same market, and $\mathbf{X}_h$ are hospital covariates. The key objects are the coefficient estimates $\hat{\boldsymbol{\beta}}$ and the regression R^2, which we adjust for measurement error.[2]

3 Data

3.1 Patients

Our primary data is a 100% sample of Medicare hospitalizations for the five emergency medical conditions in 2010-2019. The sample consists of index events, or the patient's first hospitalization for the condition in the past year. It is limited to short-term acute care hospitals, the type of facility that typically treats these patients.[3] We follow hospitals longitudinally across Medicare identifier changes via Dartmouth Institute hospital tracking files.

We require patients be fully enrolled in Original Medicare Parts A (inpatient care) and B (outpatient care) for 12 months preceding and one month following index hospitalizations. These restrictions ensure we can observe patients' index hospital choice and their full suite of care before and after that hospitalization. This includes patients' historical health conditions (defined as indicators for 87 condition categories) from diagnoses from other hospitalizations in the year preceding their index event and 30-day mortality and hospital readmission.

[2]We adjust for measurement error in two places. First, when constructing the $\hat{Z}_h$, we divide the $\hat{\alpha}_h$ by an estimate of the standard deviation that accounts for measurement error. Second, we correct the regression R^2 to account for the presence of measurement error in $\hat{Z}_h$. See Appendix A.

[3]We limit to hospitals designated as short-term or critical access hospitals by Medicare. We remove facilities that are not relevant for emergency care or not reliably open to Original Medicare patients including cancer, Tribal, Federal, intellectual disabilities, developmental, or Kaiser facilities per American Hospital Association (AHA) survey data.

3.2 Hospital Quality, Rankings, and Characteristics

We collect traditional measures of hospital quality included in public reporting efforts, pay-for-performance schemes, and the clinical and health economics literature. We describe them briefly here and in detail in the appendices. They cover three domains, the first of which is patients' health outcomes: mortality and readmission. Using the same sample of patients for whom we observe hospital choices, we estimate hospital-level rates of both outcomes at 30-days, risk-adjusted for patient demographics and historical health conditions (Chandra et al., 2016; Chandra, Kakani and Sacarny, 2024). Our approach is similar to that used by the Centers for Medicare and Medicaid Services (CMS), which reports hospital-level mortality and readmission rates and uses them in incentive schemes like the Hospital Readmissions Reduction Program (Gupta, 2021).

The second quality domain covers hospital processes of care thought to raise productivity, also called timely and effective care. These include best practices such as aspirin for AMI or timely antibiotics for pneumonia. They have been the subject of extensive clinical (Jencks, Huff and Cuerdon, 2003; Jha et al., 2005) and economic study (Skinner and Staiger, 2007, 2015); CMS also incentivizes them through the Hospital Value Based Payment Program (HVBP). We draw on the measures reported by CMS in 2014, the midpoint of our analysis period. We include all measures that are commonly reported (i.e. for at least 1,000 hospitals) for conditions in our analyses, using surgery measures to stand in for hip fracture care; while hip fracture processes are not directly assessed, surgery is typically required. There are 19 such measures (Table A1). We standardize each measure by expressing it as a z-score, average the measures together, and z-score the result.

The final domain is patient satisfaction with their hospital experience. Hospitals conduct a standard survey of patient experience called the Hospital Consumer Assessment of Provider & Systems (HCAHPS). Measures include patients' rating of hospital cleanliness and the quality of clinician communication. Like the process measures, CMS incentivizes

these indicators through HVBP. We transform the 11 patient satisfaction measures reported in 2014 into a single z-score using the same procedure we used for the process measures (see Table A2 for details).

We next collect data on two key hospital rankings available to patients. One is hospitals' overall star ratings from Medicare Care Compare (scored from 1 to 5) in 2015, the first year they were reported. These ratings aggregate over several quality measures, including those in the three domains described above (Centers for Medicare & Medicaid Services, 2016). The other is whether the hospital was ranked by US News and World Report as a top 50 facility in any specialty relevant to study conditions (cardiology, neurology, pulmonology, and orthopedics) in 2014. This ranking is based on hospital outcomes, mostly mortality; structural features including volume, staffing, and technology; reputation from physician surveys; and patient safety (Olmsted et al., 2014).

Finally, we collect hospital characteristics from the AHA hospital survey, using survey responses from 2014 (or 2015, if the 2014 value is missing). These include hospital bed size, ownership type (for-profit, non-profit, or government), medical school affiliation, and whether it has two relevant advanced technologies: cardiac catheterization and robotic surgery.

3.3 Hospital Choice Set

The conditional logit models require not only patients' realized hospital choices, but also the alternative hospitals available to them in their market. We define markets using Dartmouth Hospital Referral Regions (HRRs), which divide the US into 306 areas such that patients living in each area tend to stay there for hospital care. Consistent with that definition, and for computational tractability, we exclude the 13% of hospitalizations occurring outside patients' HRR of residence. This restriction also helps to remove patients traveling during their hospitalization. To confirm that defining markets by HRRs and excluding patients crossing HRRs do not drive our findings, we test robustness to a national gravity model – this yields nearly identical results (see section 4.5).

Then, for each market-year, we define the choice set as all hospitals with at least one event. To enter the choice set, we also require that a hospital have at least 0.1% market share in that year, and that its median market share across years during which the hospital operates is at least 0.1%. This restriction removes hospitals that are unlikely to be bona fide options for patients with emergency conditions (e.g. an orthopedic hospital that entered our data because its patient incidentally experienced a stroke). Thus, our sample consists of patients treated in their market of residence with their choices defined as other open and relevant hospitals in the same market.

The choice model also requires drive times between patients and hospitals in their markets. We use Project OSRM routing software and OpenStreetMap data to find the drive time from the centroid of beneficiaries' ZIP code of residence to the hospital's geolocation in AHA data.[4] We exclude patients for whom drive times could not be calculated. Appendix Table A3 provides details on sample construction steps.

3.4 Linkage to Occupational Data

To assess the revealed preferences of experts, we leverage a novel linkage to identify physicians and their spouses in the Medicare data. Kakani, Matecna and Chandra (2025) describes the linkage in detail. In brief, we start with a directory of physicians who accepted Medicare between 1985 and 2007. We then match this directory to a consumer marketing data source from the firm Infutor, also known as Verisk, which includes names, addresses, ages, and Social Security Numbers for over 240 million individuals in the US with over 30 years of history. These data are built from voting and death records, home ownership and credit card data, and other public information and have been used in prior work (Diamond, McQuade and Qian, 2019; Bernstein et al., 2022; Diamond and Diamond, 2024).

Reassuringly, 82% of Medicare enrollees match to Infutor data, consistent with this data

[4]We assume patients are traveling from their home to their hospital. While not universally true, over 70% of out-of-hospital cardiac arrests and 69-79% of strokes are thought to occur at home (Kelly-Hayes et al., 1995; Tsao, Connie W., Aday and et al. on behalf of the American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee, 2023).

capturing the bulk of older adults in the US. In this data, we identify an individual's spouse as another individual in the data within 15 years of age, the opposite sex, and sharing the same last name and at least one address. With this approach, we find spouses for 63% of Medicare enrollees in Infutor data.

3.5 Summary Statistics

Table 1 illustrates summary statistics separately overall and for the largest markets. We focus on the largest markets in our analyses of experts because they have sufficient volume of experts to reliably estimate their revealed preferences.[5]

In Panel A, which shows patient characteristics, a few patterns are notable. Only about half of patients choose their closest hospital, with the typical patient using a hospital located approximately 10 minutes farther than their closest option. This observation suggests that patients are willing to travel for certain hospitals, behavior that our choice model will exploit to uncover revealed preferences. In large markets, patients have more choices and are slightly less likely to use the closest one. Compared to the overall sample, experts are older but have fewer chronic conditions and better care outcomes, consistent with the socioeconomic gradient of health in the US (Chetty et al., 2016).

Panel B shows characteristics of the hospitals in our analyses. As expected, hospitals in large markets tend to have more beds and patient episodes. The typical hospital is in a market with 21 facilities in total (34 facilities in large markets).

[5]Specifically, we start with the 50 markets with the greatest patient volume. In 3 of those markets, our procedure to adjust experts' revealed preferences for measurement error fails, leaving 47 markets in our large market sample.

4 Results

4.1 Revealed Preferences vs. Other Ranking Systems

We begin by showing the association between revealed preferences and traditional quality ranking schemes.[6] Figure 1, Panel A illustrates that revealed preferences appear only weakly related to star ratings from Medicare Care Compare. While more stars are meant to indicate higher quality, the revealed preference distributions are visually similar at each star level. Column 1 of Appendix Table A5 confirms this result quantitatively, showing that only the 5-star group has detectably higher revealed preferences on average. Compared to 1-star hospitals, 5-star hospitals have two-fifths of a standard deviation higher revealed preferences.

The association is clearer for US News and World Report rankings. Figure 1, Panel B shows that top 50-ranked hospitals have higher revealed preferences on average, and Column 2 of Appendix Table A5 estimates the difference as about 1 standard deviation. Still, there is meaningful overlap in the distributions, and the bulk of variation remains unexplained. As Column 3 shows, the two ranking systems together only explain 5% of the variation in revealed preferences.

4.2 Revealed Preferences vs. Traditional Quality Measures

We then characterize the relationship between revealed preferences and traditional quality measures. Table 2 illustrates that association by estimating equation 4 with varying hospital covariates.

In column 1, we report univariate regressions of revealed preferences on each quality measure. Each measure is a highly statistically significant predictor of revealed preference. Quantitatively, a 1 percentage point higher mortality rate is associated with a quarter of a standard deviation lower revealed preference; the estimate for readmission is similar at

[6] In Appendix Table A4 we also provide illustrative revealed preference rankings in two large markets.

about one third of a standard deviation. Evidence-based care is positively associated with revealed preference, with a standard deviation greater compliance with these processes associated with two-fifths of a standard deviation higher revealed preference. Finally, we see a counterintuitive result for patient experience: better experience is associated with lower revealed preference.[7]

Column 2 reports a multivariate regression of revealed preference on the quality measures. The coefficients are qualitatively similar to those of column 1 and each quality measure remains a significant predictor. Together, these measures explain 27% of the variation in revealed preference. Column 3 adds hospital structural features to the regression. The coefficients on the quality measures attenuate by roughly one-half to two-thirds, though all remain highly statistically significant. Conditionally, government-run hospitals are less preferred, while medical school affiliated facilities and facilities with advanced technology are more preferred. With these features in the regression, we can explain nearly three-fifths of the variation in revealed preference.

4.3 Revealed Preferences of Experts

The previous section showed that existing quality measures are largely unable to explain patients' revealed preferences for hospitals. This finding could be because revealed preferences reflect choice errors due to limited information. Such errors would plausibly be less likely among expert patients. Thus, we compare the revealed preferences of expert and non-expert patients. To facilitate estimation given the small number of experts in the sample, we focus on the largest markets by patient volume. These account for 40% of all patients (Table 1).

The first column of Table 3 reports results for non-experts, who make up 99.7% of patients in these markets. As expected, the estimates are similar to the national findings already reported, though the traditional quality measures only explain about one-fifth of

[7]When we study hip and knee replacement patients, a non-emergent population, we find that patient experience and revealed preference are positively correlated (section 4.5). One possible explanation for these discordant findings is that patients allocate to hospitals based on experience when they have more time to decide or the condition is not life-threatening.

the variation here. The second column changes the outcome to the revealed preferences of experts. Compared to non-experts, expert patients put similar weight on observable quality measures. The coefficients on the measures are similar, and all remain significant predictors of revealed preferences except patient experience. Again, the measures explain about one-fifth of the variation.[8]

Column 3 compares the two groups directly by regressing the revealed preferences of experts on those of non-experts. We find that the two groups' preferences are tightly linked. A hospital with a one standard deviation higher revealed preference measure among non-experts is expected to have a 0.9 standard deviation higher revealed preference measure among experts. Non-experts' revealed preferences can also explain the bulk of variation in the revealed preferences of experts: the R^2 of this regression is 81%.

The final column asks whether experts put weight on traditional quality metrics after conditioning on the revealed preferences of non-experts. Compared to the unconditional approach (column 2), the coefficients on the quality measures attenuate greatly. Only the coefficient on mortality remains significant, showing that experts put some extra weight on this attribute. We run the same analyses with hospital structural features and find similar results, though experts also put additional weight on medical school affiliation (Appendix Table A6).

4.4 Access to High Revealed Preference Hospitals

Finally, we evaluate differences across patient populations in their access to high revealed preference hospitals. We do so by conducting patient-level regressions in the set of patients analyzed in the choice models. We regress hospital revealed preference measures on patient characteristics. Since we only identify revealed preferences within a market, we include market fixed effects and interpret results as within-market differences in access.

[8]Given the small samples of experts, the correction for measurement error is crucial here, raising the R^2 from 12.1% to 19.4%. It plays comparatively little role when analyzing non-experts because their revealed preferences are estimated precisely.

To illustrate the roles of geography and choice in access, the first two columns of Table 4 report univariate regressions in which the outcome is the revealed preference of the patient's closest hospital (column 1) or the hospital they actually used (column 2). These show racial disparities, with a Black-white gap in the average revealed preference of the nearest hospital that grows when looking at the hospital actually used. The larger gap in column 2 indicates that it comes at least partly from patients' hospital choice, not just where patients live. Similarly, Hispanic patients, Medicaid eligible patients, and those in lower-income areas both live near and use less-preferred hospitals.

While rural patients' nearest hospitals have 1 standard deviation lower revealed preference measures than urban patients on average, the gap for the hospitals they actually use is only about half of that. The attenuation reflects the role of choice in compressing a rural-urban gap that would prevail if patients mechanically went to their nearest hospital. For rural patients, nearby facilities are often less preferred, and some bypass them for more-preferred hospitals that are farther away.

In column 3, we replace market fixed effects with ZIP code fixed effects so that the regressions compare patients living in the same small areas. We observe that racial and socioeconomic disparities persist within a finer geographic level. They are therefore not simply the result, for instance, of white patients living closer to high revealed preference hospitals within markets. Column 4 reports one multivariate regression and shows that disparities are similar whether unconditional or conditional on the full set of covariates.

4.5 Robustness

We confirm the robustness of our findings to alternative sample decisions and modeling approaches. We describe the alternatives briefly here and in full detail in Appendix C. Appendix Tables A7-A11 report robustness checks of the findings in all main tables.

First, we analyze hip and knee replacements, which are among the most commonly performed hospital procedures in the Medicare population. In contrast to the emergency

patients we have focused on, these patients typically schedule treatment in advance and face little time pressure when making their hospital choice. We find that results are similar in this cohort, though patient experience is positively associated with revealed preference here and the explanatory power of traditional quality metrics is lower.

Second, we consider an alternative parameterization of revealed preference that expresses it as a marginal rate of substitution (MRS) between hospital choice and travel time in minutes, a willingness to travel metric. Negative values indicate greater revealed preference, i.e. an MRS of -10 means that the hospital gives utility equal to a 10 minute shorter trip. Unlike our baseline approach, the variance of this metric not trivially equal to one in each market, and Appendix Figure A1 plots its distribution.[9] While the scaling of MRS is different from the z-score, the explanatory power of traditional quality measures under this approach is quite similar to baseline. The correlation between the MRSs of experts and non-experts is somewhat weaker, though R^2 is still 65%.

Third, we consider the possibility that patients' hospital choices may be limited by capacity constraints. If this occurs, a desirable hospital's revealed preference measure could be biased downward because it turns away patients who otherwise would have chosen it. Such bias could obscure the explanatory power of traditional quality metrics. Reassuringly, hospitals generally have excess capacity, with 95% having average occupancy rates under 85% (MedPAC, 2024). Still, to investigate the role of capacity constraints, we leverage the fact that occupancy is typically highest during the winter when respiratory infections and cardiovascular events are more common (Akintoye et al., 2017; Janke, Burke and Haimovich, 2025). We therefore evaluate robustness to excluding admissions occurring during the winter season, finding very similar results.

Finally, we consider a more flexible market definition than the fixed geographic approach of HRRs. The approach addresses the concern that revealed preferences for some hospitals

[9]In the unusually geographically large market of Alaska, all 10 hospitals had an absolute value of MRS over 200 minutes. The standard deviation of MRS excluding Alaska is just 31 minutes. These observations would exert disproportionate weight in the regressions, so we remove them from MRS analyses.

might be biased because we omit patients traveling to them from outside their home HRR. Specifically, rather than limiting to patients who stay in their HRR of residence for treatment, we include any hospital within 50 miles in their choice sets. We then run one national choice model. This approach exploits an equivalence between conditional logistic regression and a Poisson gravity model (as in Dingel et al., 2023) for tractability. It yields nearly identical results to baseline.

Overall, we find that the results are highly robust across the sensitivity analyses. Quality measures still only explain a minority of variation in hospital revealed preferences (Appendix Table A7). Revealed preferences for experts are also highly similar to those estimated on a general population across specifications (Appendix Tables A8 and A9). Finally, the socioeconomic and geographic disparities in use of high revealed preference hospitals persist across specifications (Appendix Tables A10 and A11).

5 Discussion and Conclusion

This paper develops and validates a revealed preference approach to estimating hospital quality, using patients' choices to measure the value they place on each hospital. Patients' revealed preferences for hospitals are not well correlated with existing ranking systems. While existing measures of quality are associated with revealed preferences, they explain only about one-fourth of the variation in them. Clinical experts have revealed preferences that are highly correlated with those of non-experts, and the two groups put similar weights on observable quality. Previously established racial, socioeconomic, and geographic disparities in access to high-quality hospitals persist when assessing quality using this more holistic approach.

Our findings show that revealed preference rankings capture otherwise unobservable information regarding the quality of hospitals, even in markets where quality measures currently exist. The method also permits assessing quality in areas without reliable measures where choice is still observed, as is commonly the case in administrative data. For instance, it

is possible to construct revealed preference measures for cancer care or outpatient surgery. Moreover, these findings suggest that revealed preference rankings can capture information about the valuations of clinical experts even if constructed among non-experts. Health care data typically does not identify which patients are themselves clinicians, so this finding further emphasizes the utility and tractability of this method.

Our approach may be useful for policymakers seeking to improve the functioning of health care markets. First, revealed preference measures of quality could provide guidance to patients seeking care – indeed, we present them for the Boston and New York City markets in Appendix Table A4. Second, they are difficult or impossible to game, addressing concerns about selection and upcoding in traditional approaches. Third, they can be calculated and disseminated for conditions where quality is hard to measure.

Taken together, this study highlights that market allocation processes can capture meaningful quality signals in the health care sector. It also raises several questions for future work. These include investigation of what accounts for the unexplained component of patients' preferences, whether and why revealed preferences differ across patient groups, and relatedly, the mechanisms underlying disparities in access to high-quality providers.

References

Abaluck, Jason, and Jonathan Gruber. 2011. "Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program." *American Economic Review*, 101(4): 1180–1210.

Akintoye, Emmanuel, Alexandros Briasoulis, Alexander Egbe, Oluwole Adegbala, Samson Alliu, Muhammad Sheikh, Manmohan Singh, Abdelrahman Ahmed, Sagar Mallikethi-Reddy, and Diane Levine. 2017. "Seasonal variation in hospitalization outcomes in patients admitted for heart failure in the United States." *Clinical Cardiology*, 40(11): 1105–1111.

Arrow, Kenneth. 1963. "Uncertainty and the welfare economics of medical care." *American Economic Review*, LIII(5).

Avery, Christopher N., Mark E. Glickman, Caroline M. Hoxby, and Andrew Metrick. 2013. "A Revealed Preference Ranking of U.S. Colleges and Universities*." *The Quarterly Journal of Economics*, 128(1): 425–467.

Bach, Peter B., Hoangmai H. Pham, Deborah Schrag, Ramsey C. Tate, and J. Lee Hargraves. 2004. "Primary Care Physicians Who Treat Blacks and Whites." *New England Journal of Medicine*, 351(6): 575–584. _eprint: https://www.nejm.org/doi/pdf/10.1056/NEJMsa040609.

Bastani, Hamsa, Joel Goh, and Mohsen Bayati. 2019. "Evidence of Upcoding in Pay-for-Performance Programs." *Management Science*, 65(3): 1042–1060.

Bernstein, Shai, Rebecca Diamond, Abhisit Jiranaphawiboon, Timothy McQuade, and Beatriz Pousada. 2022. "The Contribution of High-Skilled Immigrants to Innovation in the United States." National Bureau of Economic Research w30797, Cambridge, MA.

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor. 2017. "Choose to Lose: Health Plan Choices from a Menu with Dominated Options." *The Quarterly Journal of Economics*, 132(3): 1319–1372.

Bronnenberg, Bart J., Jean-Pierre Dubé, Matthew Gentzkow, and Jesse M. Shapiro. 2015. "Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium *." *The Quarterly Journal of Economics*, 130(4): 1669–1726.

Centers for Medicare & Medicaid Services. 2016. "First Release of the Overall Hospital Quality Star Rating on Hospital Compare." tex.howpublished: Fact Sheet, U.S. Centers for Medicare & Medicaid Services.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson. 2016. "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector." *American Economic Review*, 106(8): 2110–2144.

Chandra, Amitabh, Maurice Dalton, and Douglas Staiger. 2023. "Are Hospital Quality Indicators Causal?" National Bureau of Economic Research w31789, Cambridge, MA.

Chandra, Amitabh, Pragya Kakani, and Adam Sacarny. 2024. "Hospital Allocation and Racial Disparities in Health Care." *The Review of Economics and Statistics*, 106(4): 924–937.

Chen, Yiqun, Petra Persson, and Maria Polyakova. 2022. "The Roots of Health Inequality and the Value of Intrafamily Expertise." *American Economic Journal: Applied Economics*, 14(3): 185–223.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. "The Association Between Income and Life Expectancy in the United States, 2001-2014." *JAMA*, 315(16): 1750.

Cubanski, Juliette, Anthony Damico, Tricia Neuman, and Gretchen Jacobson. 2018. "Sources of supplemental coverage among medicare beneficiaries in 2016." tex.howpublished: Data Note, KFF (Kaiser Family Foundation).

Dafny, Leemore S. 2005. "How Do Hospitals Respond to Price Changes?" *American Economic Review*, 95(5): 1525–1547.

Diamond, Rebecca, and William Diamond. 2024. "Racial Differences in the Total Rate of Return on Owner-Occupied Housing." National Bureau of Economic Research w32916, Cambridge, MA.

Diamond, Rebecca, Tim McQuade, and Franklin Qian. 2019. "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco." *American Economic Review*, 109(9): 3365–3394.

Dingel, Jonathan, Joshua Gottlieb, Maya Lozinski, and Pauline Mourot. 2023. "Market Size and Trade in Medical Services." National Bureau of Economic Research w31030, Cambridge, MA.

Doyle, Joseph, John Graves, and Jonathan Gruber. 2019. "Evaluating Measures of Hospital Quality: Evidence from Ambulance Referral Patterns." *The Review of Economics and Statistics*, 101(5): 841–852.

Einav, Liran, Amy Finkelstein, and Neale Mahoney. 2025. "Producing Health: Measuring Value Added of Nursing Homes." *Econometrica*, 93(4): 1225–1264.

Finkelstein, Amy, Petra Persson, Maria Polyakova, and Jesse M. Shapiro. 2022. "A Taste of Their Own Medicine: Guideline Adherence and Access to Expertise." *American Economic Review: Insights*, 4(4): 507–526.

Frakes, Michael, Jonathan Gruber, and Anupam Jena. 2021. "Is great information good enough? Evidence from physicians as patients." *Journal of Health Economics*, 75: 102406.

Geweke, John, Gautam Gowrisankaran, and Robert J. Town. 2003. "Bayesian Inference for Hospital Quality in a Selection Model." *Econometrica*, 71(4): 1215–1238.

Gupta, Atul. 2021. "Impacts of Performance Pay for Hospitals: The Readmissions Reduction Program." *American Economic Review*, 111(4): 1241–1283.

Handel, Benjamin R. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review*, 103(7): 2643–2682.

Hull, Peter. 2020. "Estimating Hospital Quality with Quasi-experimental Data."

Janke, Alexander T., Laura G. Burke, and Adrian Haimovich. 2025. "Hospital 'Boarding' Of Patients In The Emergency Department Increasingly Common, 2017–24." *Health Affairs*, 44(6): 739–744.

Jencks, Stephen F., Edwin D. Huff, and Timothy Cuerdon. 2003. "Change in the Quality of Care Delivered to Medicare Beneficiaries, 1998-1999 to 2000-2001." *JAMA*, 289(3): 305.

Jha, Ashish K., E. John Orav, and Arnold M. Epstein. 2011. "Low-Quality, High-Cost Hospitals, Mainly In South, Care For Sharply Higher Shares Of Elderly Black, Hispanic, And Medicaid Patients." *Health Affairs*, 30(10): 1904–1911.

Jha, Ashish K., E. John Orav, Zhonghe Li, and Arnold M. Epstein. 2007. "Concentration and Quality of Hospitals That Care for Elderly Black Patients." *Archives of Internal Medicine*, 167(11): 1177–1182.

Jha, Ashish K., Zhonghe Li, E. John Orav, and Arnold M. Epstein. 2005. "Care in U.S. Hospitals — The Hospital Quality Alliance Program." *New England Journal of Medicine*, 353(3): 265–274.

Johnson, Erin M., and M. Marit Rehavi. 2016. "Physicians Treating Physicians: Information and Incentives in Childbirth." *American Economic Journal: Economic Policy*, 8(1): 115–141.

Joynt, Karen E., Yael Harris, E. John Orav, and Ashish K. Jha. 2011. "Quality of Care and Patient Outcomes in Critical Access Rural Hospitals." *JAMA*, 306(1): 45–52.

Joynt Maddox, Karen E., Mat Reidhead, Jianhui Hu, Amy J. H. Kind, Alan M. Zaslavsky, Elna M. Nagasako, and David R. Nerenz. 2019. "Adjusting for social risk factors impacts performance and penalties in the hospital readmissions reduction program." *Health Services Research*, 54(2): 327–336. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-6773.13133.

Kakani, Pragya, Simone Matecna, and Amitabh Chandra. 2025. "Expert Patients' Use of Avoidable Health Care."

Kelly-Hayes, Margaret, Philip A. Wolf, Carlos S. Kase, Frederick N. Brand, Jacqueline M. McGuirk, and Ralph B. D'Agostino. 1995. "Temporal Patterns of Stroke Onset." *Stroke*, 26(8): 1343–1347.

Koedel, Cory, Kata Mihaly, and Jonah E. Rockoff. 2015. "Value-added modeling: A review." *Economics of Education Review*, 47: 180–195.

Kolstad, Jonathan T. 2013. "Information and Quality When Motivation is Intrinsic: Evidence from Surgeon Report Cards." *American Economic Review*, 103(7): 2875–2910.

Levine, David M., Bruce E. Landon, and Jeffrey A. Linder. 2019. "Quality and Experience of Outpatient Care in the United States for Adults With or Without Primary Care." *JAMA Internal Medicine*, 179(3): 363–372.

Levitt, Steven D., and Chad Syverson. 2008. "Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions." *The Review of Economics and Statistics*, 90(4): 599–611.

Liang, Lan, Brian Moore, and Anita Soni. 2020. "National Inpatient Hospital Costs: The Most Expensive Conditions by Payer, 2017."

Loccoh, Eméfah C., Karen E. Joynt Maddox, Yun Wang, Dhruv S. Kazi, Robert W. Yeh, and Rishi K. Wadhera. 2022. "Rural-Urban Disparities in Outcomes of Myocardial Infarction, Heart Failure, and Stroke in the U.S." *Journal of the American College of Cardiology*, 79(3): 267–279.

MedPAC. 2024. "Health Care Spending and the Medicare Program." MedPAC.

Olenski, Andrew, and Szymon Sacher. 2024. "Estimating Nursing Home Quality with Selection." *Review of Economics and Statistics*, 1–31.

Olmsted, Murrey G., Emily Geisen, Joe Murphy, Denise Bell, Melissa Morley, and Marshica Stanley. 2014. "Methodology: U.S. News & World Report Best Hospitals 2014-15."

Romley, John A., and Dana P. Goldman. 2011. "How Costly is Hospital Quality? A Revealed-Preference Approach." *The Journal of Industrial Economics*, 59(4): 578–608. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-6451.2011.00468.x.

Rothstein, Jesse, and Albert Yoon. 2024. "Rankings without U.S. News: A revealed preference approach to evaluating law schools." *Journal of Empirical Legal Studies*, 21(2): 279–336. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jels.12380.

Rutherford, R. C., T. M. Springer, and A. Yavas. 2005. "Conflicts between principals and agents: evidence from residential brokerage." *Journal of Financial Economics*, 76(3): 627–665.

Skinner, Jonathan, and Douglas Staiger. 2007. "Technology adoption from hybrid corn to beta-blockers." In *Hard-to-measure goods and services: Essays in honor of zvi griliches*. 545–570. University of Chicago Press.

Skinner, Jonathan, and Douglas Staiger. 2015. "Technology Diffusion and Productivity Growth in Health Care." *Review of Economics and Statistics*, 97(5): 951–964.

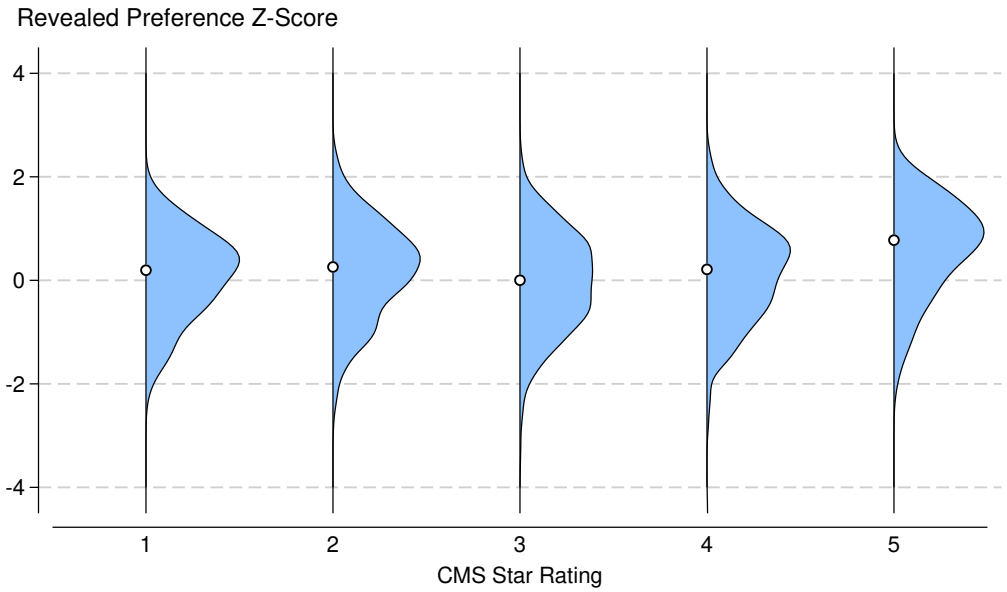
Sorkin, Isaac. 2018. "Ranking Firms Using Revealed Preference*." *The Quarterly Journal of Economics*, 133(3): 1331–1393.

Tsao, Connie W., Aaron W. Aday, and Zaid I. Almarzooq et al. on behalf of the American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. 2023. "Heart Disease and Stroke Statistics—2023 Update: A Report From the American Heart Association." *Circulation*, 147(8): e93–e621.

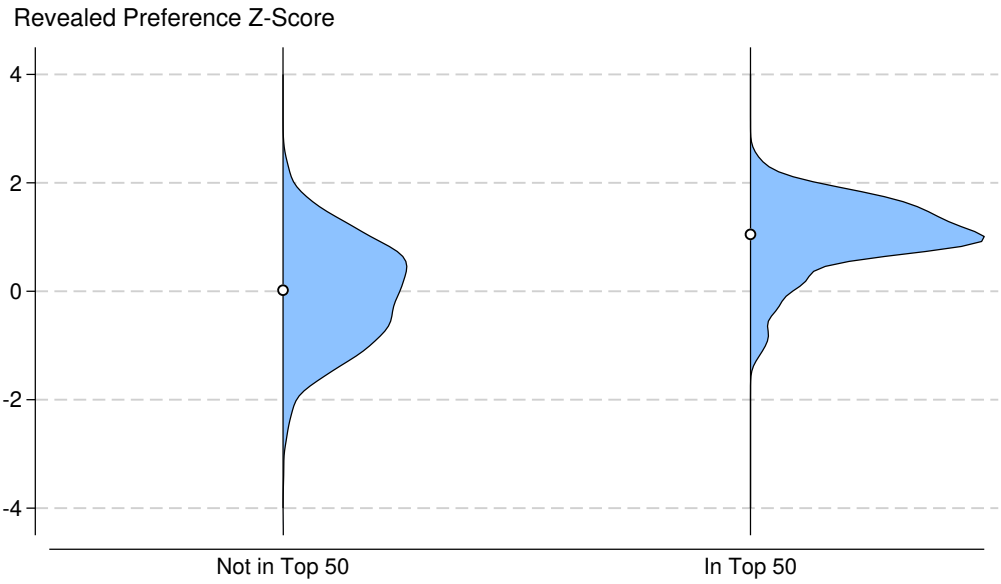
Figures and Tables

Figure 1: Distributions of revealed preferences by traditional quality rankings

A. Medicare Care Compare star rating



B. U.S. News & World Report ranking



Notes: This figure illustrates distributions of revealed preference z-scores across hospitals by star rating from Medicare Care Compare (Panel A) and by whether the hospital was rated by U.S. News & World Report in the top 50 (within relevant specialties (cardiology, neurology, pulmonology, and orthopedics; Panel B). White circles indicate medians.

Table 1: Summary statistics on patients and hospitals

A. Patient characteristics

	All markets		Large markets	
Patients:	All	Experts	All	Experts
Age	80.85	81.97	80.92	82.10
Female	0.568	0.451	0.569	0.453
Race-ethnicity				
White	0.847	0.881	0.838	0.878
Black	0.079	0.014	0.087	0.016
Hispanic	0.045	0.024	0.042	0.025
Other/Unknown	0.029	0.081	0.033	0.081
Medicaid eligible	0.165	0.018	0.163	0.019
Rural	0.042	0.014	0.022	0.008
Area avg. income ($1,000s)	79.66	109.3	85.51	117.9
No. of comorbidities	2.059	1.643	2.134	1.719
30-day outcomes				
Mortality	0.121	0.109	0.118	0.105
Readmission	0.165	0.147	0.170	0.152
No. of hospital choices	22.99	23.36	34.72	34.23
Chose closest hospital	0.531	0.497	0.479	0.435
Minutes to				
Chosen hospital	21.87	17.58	21.23	16.83
Closest hospital	12.78	10.48	12.00	9.833
Expert	0.003	1.000	0.003	1.000
Patients	9,309,821	24,507	3,746,103	11,883

B. Hospital characteristics

	All markets	Large markets
Ownership		
Non-profit	0.638	0.694
For-profit	0.176	0.187
Government	0.186	0.119
Beds	183.0	270.0
Patients in analysis	2,542.3	3,772.5
30-day risk-adjusted outcomes		
Mortality	0.120	0.113
Readmission	0.165	0.167
Medicare star rating (1-5)	3.056	2.985
US News top 50	0.024	0.051
Medical school	0.300	0.404
Robotic surgery	0.332	0.508
Cardiac catheterization	0.418	0.635
Hospitals in market	21.10	33.78
Markets	305	47
Hospitals	3,662	993

Notes: This table presents characteristics of patients (Panel A) and hospitals (Panel B) in the hospitals analyzed in study regressions. Each cell is an average. Both panels distinguish between all markets and the 47 large markets that are the focus of the analysis of experts. Panel A also distinguishes between all patients and experts.

Table 2: Relationship between hospital characteristics and revealed preferences

Specification:	(1) Univariate	(2) Multivariate	(3) Multivariate
Quality measures			
Mortality, 30 days in percentage points	-0.229 (0.0117)	-0.211 (0.0109)	-0.0912 (0.00875)
Readmission, 30 days in percentage points	-0.315 (0.0193)	-0.301 (0.0174)	-0.116 (0.0142)
Evidence-based care z-score	0.417 (0.0236)	0.340 (0.0222)	0.167 (0.0165)
Patient experience z-score	-0.142 (0.0255)	-0.250 (0.0227)	-0.0959 (0.0174)
Hospital structural features			
Ownership			
Non-profit			ref.
For-profit			-0.0318 (0.0355)
Government			-0.220 (0.0325)
Medical school			0.333 (0.0279)
Robotic surgery			0.484 (0.0302)
Cardiac catheterization			0.787 (0.0319)
R^2, unadjusted		0.273	0.577
R^2, meas. err. adjusted		0.274	0.578
Hospitals	3,662	3,662	3,662

Notes: This table presents regressions of revealed preference z-scores on hospital quality measures and structural features. Column 1 presents univariate regressions in which each cell is a different regression. Columns 2 and 3 present multivariate regressions in which each column is a regression. All regressions include market fixed effects. Both R^2 are within-market; they are unadjusted and adjusted for regressand measurement error, respectively. Robust standard errors in parentheses.

Table 3: Relationship between traditional quality measures and the revealed preferences of experts and non-experts

Revealed preferences of:	(1) Non-experts	(2) Experts	(3) Experts	(4) Experts
Mortality, 30 days in percentage points	-0.194 (0.0230)	-0.247 (0.0309)		-0.0754 (0.0225)
Readmission, 30 days in percentage points	-0.206 (0.0331)	-0.159 (0.0500)		0.0240 (0.0408)
Evidence-based care z-score	0.641 (0.0940)	0.518 (0.107)		-0.0501 (0.0664)
Patient experience z-score	-0.146 (0.0488)	-0.0866 (0.0622)		0.0429 (0.0466)
Non-expert revealed preference z-score			0.905 (0.0445)	0.886 (0.0468)
R^2, unadjusted	0.188	0.104	0.440	0.446
R^2, meas. err. adjusted	0.189	0.195	0.821	0.834
Hospitals	993	993	993	993

Notes: This table presents regressions of revealed preference z-scores on hospital quality measures, varying whether the revealed preferences are calculated for non-experts (column 1) or experts (columns 2-4). Each column represents one multivariate regression. All regressions include market fixed effects. Both R^2 are within-market; they are unadjusted and adjusted for regressand measurement error, respectively. Robust standard errors in parentheses.

Table 4: Association between patient characteristics and the revealed preferences of their closest and chosen hospitals

Hospital:	(1) Closest	(2) Chosen	(3) Chosen	(4) Chosen
Race-ethnicity				
White	ref.	ref.	ref.	ref.
Black	-0.0098	-0.0476	-0.0373	-0.0304
	(0.0011)	(0.0010)	(0.0009)	(0.0009)
Hispanic	-0.129	-0.113	-0.0223	-0.0034
	(0.0016)	(0.0015)	(0.0011)	(0.0011)
Other	-0.0283	0.0085	-0.0002	0.0122
	(0.0018)	(0.0016)	(0.0012)	(0.0012)
Age (Decades)	0.0334	-0.0295	-0.0509	-0.0498
	(0.0004)	(0.0003)	(0.0002)	(0.0002)
Female	0.0162	-0.0347	-0.0429	-0.0228
	(0.0006)	(0.0005)	(0.0004)	(0.0004)
Medicaid eligible	-0.152	-0.214	-0.112	-0.108
	(0.0008)	(0.0008)	(0.0006)	(0.0006)
No. of comorbidities*	0.0610	-0.0955	-0.0956	-0.0533
	(0.0090)	(0.0083)	(0.0060)	(0.0060)
Area avg income ($100s)	0.0911	0.0683		
	(0.0001)	(0.0001)		
Rural	-1.001	-0.558		
	(0.0015)	(0.0019)		
Geographic Controls Specification	Market Univariate	Market Univariate	ZIP Univariate	ZIP Multivariate
Patients	9,131,832	9,131,832	9,131,659	9,131,659

Notes: This table presents regressions of the revealed preference z-scores of a patient's hospital on patient characteristics using the patient-level sample of hospitalizations. In column 1, the outcome is the revealed preference of the patient's closest hospital, while in columns 2-4 the outcome is the revealed preference of the hospital the patient actually chose. Columns 1-3 present univariate regressions in which each cell is a different regression, column 4 presents a single multivariate regression. Regressions include geographic controls (fixed effects): market-level (columns 1-2) or ZIP code-level (columns 3-4) fixed effects. All models exclude observations with any missing covariates. Analysis of the closest hospital is based on the closet in-sample hospital. Robust standard errors in parentheses. * Coefficients for comorbidities are multiplied by 100 for ease of exposition.

Appendix To:

Revealed Preferences for Hospitals

Amitabh Chandra, Pragya Kakani, Alex Philip, and Adam Sacarny

September 11, 2026

Appendix A Adjustment for measurement error

The conditional logit models emit estimates of hospital effects $\hat{\alpha}_h$ that measure the true effects $\alpha_h^\star$ with error. If we did not account for it, this measurement error would cause two issues. First, it would tend to attenuate the revealed preference z-scores, making it hard to interpret regression coefficients. Second, it would downward-bias R^2 estimates, so the true explanatory power of traditional quality metrics would not be clear. In this appendix, we review how we address these issues.

A.1 Revealed preference z-scores

First, failure to account for measurement error would make it harder to interpret the regression coefficients $\boldsymbol{\beta}$ in the regression given by equation 4 in the main text. These coefficients would express effects in terms of standard deviations of $\hat{\alpha}_h$ inclusive of the noise. Noise varies across markets and samples – for instance, all else the same, the effects will be estimated more precisely, and the variance of the noise will be lower, in markets and samples with more patients. As a result, noise also varies between analyses of experts and non-experts, with much more noise in the former since it contains fewer patients than the latter. A standard deviation of hospital effect estimates among experts is comprised of more noise, and less signal, than a standard deviation of hospital effect estimates in the much larger non-expert sample.

Instead, we would like the $\boldsymbol{\beta}$ to express effects as expected changes in standard deviations of the true $\alpha_h^\star$. This approach would make comparisons across expert and non-expert samples more interpretable. Therefore, to calculate $\hat{Z}_h$, we must derive a corrected standard deviation.

Suppose there is a noisy measure of the hospital's unobserved true revealed preference:

$$\alpha_h = \alpha_h^\star + \nu_h \qquad \alpha_h^\star \sim (\mu_m, \sigma_m^2) \qquad \nu_h \sim (0, \kappa_h^2).$$

In practice, we observe this object normalized to average to 0 in each market: $\hat{\alpha}_h = \alpha_h - \bar{\alpha}_m$. This setup reflects the normalization of the $\hat{\alpha}_h$ from the conditional logit. This implies:

$$\hat{\alpha}_h = [\alpha_h^\star - \bar{\alpha}_m^\star] + \omega_h \quad \omega_h \sim (0, \varsigma_h^2),$$

where ω_h is a noise term. The conditional logit emits an estimate $\hat{\varsigma}_h^2$ of its variance — this is the squared standard error of the hospital effect. The standard errors are calculated by selecting the hospital fixed effects to be used in the second-stage regression and applying the normalization that they average to 0.

Now consider the sample variance estimator S_m^2. This estimator reduces to:

$$\begin{aligned} S_m^2 &= \frac{1}{N_m - 1} \sum_h (\hat{\alpha}_h^2 - \overline{\hat{\alpha}}_m)^2 \\ &= \frac{1}{N_m - 1} \sum_h (\alpha_h^\star - \bar{\alpha}_m^\star)^2 + \frac{1}{N_m - 1} \sum_h \omega_h^2 \end{aligned}$$

Taking expectations shows that this estimator is biased due to the measurement error:

$$\mathbb{E}[S_m^2] = \sigma_m^2 + \frac{1}{N_m - 1} \sum_h \varsigma_h^2$$

We therefore define the bias-corrected sample variance estimator:

$$\tilde{S}_m^2 = S_m^2 - \frac{1}{N_m - 1} \sum_h \hat{\varsigma}_h{}^2,$$

where the second term is the degrees-of-freedom corrected average of the squared standard errors of the hospital effects.

We then define the standardized revealed preference measure as:

$$\hat{Z}_h = \frac{\hat{\alpha}_h}{\tilde{S}_m}.$$

A.2 R^2

The second issue occurs in the analysis of the explanatory power of hospital covariates in the second-stage regression via R^2. Absent adjustment, the R^2 will tend to under-estimate the fraction of underlying variation in revealed preferences that the covariates can explain. Note that because hospitals' revealed preferences are not separately identified across markets, the regressions include market fixed effects and we focus on the within-market R^2.

We now derive the asymptotic bias of the within-market R^2. First, note that we can also think of the adjusted z-score as a noisy estimate of the true (unobserved) z-score:

$$\hat{Z}_h = Z_h + \phi_h \quad \phi_h \sim (0, \iota_h^2)$$

Starting with the regression given by equation 4, we demean by markets and add measurement error:

$$Z_h = \ddot{\mathbf{X}}_h \boldsymbol{\beta} + \ddot{\varphi}_h + \phi_h,$$

where double-dots denote de-meaned variables; we omit them from $\hat{Z}_h$ and φ_h because the true and noisy z-scores are already defined to have a mean of 0 in each market.

Written this way, it is clear that the within-market R^2_W will tend to be downward-biased in the limit. To assess the asymptotic bias, we consider how this object will behave as

$N \to \infty$:

$$\begin{aligned} R^2_W &= 1 - \frac{RSS}{WSS} \\ &= 1 - \frac{\sum_h \left(\hat{Z}_h - \ddot{X}_h\hat{\beta}\right)^2}{\sum_h \hat{Z}_h^2} \\ &\to 1 - \frac{\frac{1}{N}\sum_h \left(\hat{Z}_h - \ddot{X}_h\beta\right)^2}{\frac{1}{N}\sum_h \hat{Z}_h^2} \\ &\to 1 - \frac{\frac{1}{N}\sum_h \ddot{\varphi}_h^2 + \frac{1}{N}\sum_h \phi_h^2}{\frac{1}{N}\sum_h Z_h^2 + \frac{1}{N}\sum_h \phi_h^2} \\ &\to 1 - \frac{\sigma^2_{\ddot{\varphi}} + \sigma^2_\phi}{1 + \sigma^2_\phi}, \end{aligned}$$

where RSS and WSS both include the unwanted term σ^2_ϕ (we assume the averages of $\ddot{\varphi}^2_h \to \sigma^2_{\ddot{\varphi}}$ and $\ddot{\phi}^2_h \to \sigma^2_\phi$). Note, for instance, that WSS/N limits to greater than 1 even though in the ideal (but infeasible) regression, it is the limiting variance of a z-score.

To correct the object for measurement error, we use the squared standard errors of the z-scores, $\hat{\iota}^2_h$, as estimates of the variances of the measurement errors and subtract their sum from RSS and WSS:

$$\tilde{R}^2_W = 1 - \frac{\sum_h \left(\hat{Z}_h - \ddot{X}_h\hat{\beta}\right)^2 - \sum_h \hat{\iota}^2_h}{\sum_h \hat{Z}_h^2 - \sum_h \hat{\iota}^2_h}.$$

Appendix B Hospital quality measures

B.1 Risk-adjusted mortality and readmission

We estimate hospital 30-day risk-adjusted mortality and readmission rates and use them as traditional quality measures in the regressions. Our approach is based on CMS mortality and readmission measures and is adapted from prior work (Chandra et al., 2016; Chandra, Kakani and Sacarny, 2024). Specifically, we start with the full sample of hospitalizations described in Section 3.1. We then limit to hospitals that have 25 or more hospitalizations during our study period so that the risk-adjusted rate can be reliably measured.

We then estimate hospital random effects regressions of the following form:

$$Y_p = \beta X_p + \gamma_h + \delta_t + \xi_c + \epsilon_p,$$

where Y_p is the mortality or readmission outcome for patient episode p.[10] Here, X_p is a vector of patient covariates that includes (i) interactions between age groups of five years (66-69, 70-74, 75-79, 80-84, 85-89, 90+), race (White, Black, Hispanic, Other/Unknown), sex, and (ii) comorbidities drawn from diagnosis codes from hospitalizations that occurred during the previous 12 months.[11]

The γ_h are hospital random effects, while δ_t and ξ_c are fixed effects for year and clinical cohort (acute myocardial infarction, heart failure, hip fracture, pneumonia, or stroke). For each outcome, mortality or readmission, we estimate the regression and extract BLUPs of the hospital random effects, $\hat{\gamma}_h$. We normalize $\hat{\gamma}_h$ so that its patient-weighted mean equals that of Y_p. We define this object as the risk-adjusted mortality or readmission measure.

[10] More specifically, 30-day mortality is defined as death within 30 days of admission and 30-day readmission is defined as a subsequent hospitalization at any short-term acute care or critical access hospital within 30 days of discharge.

[11] Our comorbidities are indicators for the 87 chronic condition categories used in Version 21 (2015 Model Software) in the CMS Risk-Adjustment Model for Medicare Advantage plans (CMS, 2015).

B.2 Patient experience and process of care measures

We collect patient experience and process of care measures from Medicare Care Compare. For process of care measures, we focus on those reported for at least 1,000 hospitals in domains that are directly relevant to our clinical cohorts in year 2014. Specifically, we include process measures for acute myocardial infarction, congestive heart failure, pneumonia, stroke, or surgery. We use the surgery measures because there is no domain for hip fracture but treating this condition typically requires surgical intervention. The measures are listed in Table A1. We normalize each measure into a z-score. Then, for each hospital, we average together all available z-scores. Finally, we take this composite measure and z-score it.

The patient experience measures are drawn from the 2014 CMS Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey. In the HCAHPS survey, patients are asked to evaluate hospitals on the questions listed in Table A2. For most questions, patients could respond "never", "sometimes", "usually", or "always". The public data combines the first two responses and gives shares of patients responding never/sometimes, usually, or always. We assign these responses values of 1, 2, or 3, respectively, and create a share-weighted score.[12] Analogous to our approach with process of care measures, we z-score each question, average together the z-scores, and take the z-score of the composite measure.

[12]For the "overall rating" question, patients could score the hospital on a scale from 0-10; the public data gives shares scoring 1-6, 7-8, and 9-10. The "would recommend" question gives shares saying "would not", "would probably", and "would definitely". As for the other questions, we value these responses at 1, 2, or 3, respectively. The "recovery at home" question is given as yes/no, which we translate to 1 or 0, respectively.

Appendix C Robustness

C.1 Revealed preferences for hip and knee replacement

In our first robustness exercise, we calculate revealed preferences among patients receiving hip and knee replacements. Because hip and knee replacements are non-emergent, they are typically planned and thus may exhibit greater scope for patient choice.

Our sample construction and revealed preference estimation approach parallels that used in the emergent patient sample that is the focus of our main analyses. The only difference is that we define the sample as index events for hip or knee replacement. We limit to the 82% of patients seeking care within their market of residence. While the bulk of patients remain in their market of residence for hip and knee replacement, the share exiting is somewhat greater here than in the main sample, reflecting that this care is more deferrable.

The traditional quality measures for hip and knee replacement differ somewhat from the ones we evaluated for emergent conditions in the main text. Because mortality is rare following these surgeries, CMS does not report a hospital mortality measure for these procedures and we also omit it from our analyses. We follow CMS and estimate a 30-day risk-adjusted readmission rate for hospitals using the approach described in Appendix B.1. For process measures, we use the same approach as the main analyses but calculate the z-score limiting only to surgical domain measures in Table A1. Because the HCAHPS is not condition-specific, we use the same patient experience z-score here as for the main analyses.

C.2 Measuring revealed preference as a marginal rate of substitution

We next test the robustness of our findings to measuring revealed preference as a marginal rate of substitution (MRS) rather than a z-score. We define this measure as the MRS between travel time and hospital choice. It is normalized to average to 0 within a market, and thus

reflects patients' willingness to travel to a particular hospital relative to the average hospital in the market. A more negative MRS indicates a more-preferred hospital. This measure can be interpreted in terms of minutes; for instance, assuming the utility function given by equation 1, choosing a hospital with an MRS of -10 gives the patient utility equal to 10 minutes less travel time, compared to picking the average hospital in the market.

The MRS between travel time and hospital h for patient p is:

$$MRS_{ph} = \frac{\partial U_{ph}/\partial C_h}{\partial U_{ph}/\partial traveltime_{ph}} = \frac{\alpha_h}{f'(traveltime_{ph}; \boldsymbol{\rho})}.$$

Since f is a quadratic function of travel time with coefficients that are cohort-specific (i.e. AMI, heart failure, hip fracture, pneumonia, stroke), let ρ_1 and ρ_2 be the simple averages of the cohort-specific polynomial travel time terms. Then, let $\overline{traveltime}_m$ be the average realized travel time in minutes for all patients in the market. We estimate the hospital's MRS as:

$$\hat{MRS}_h = \frac{\hat{\alpha}_h}{\hat{\rho}_1 + 2 \cdot \hat{\rho}_2 \cdot \overline{traveltime}_m},$$

where the hats denote estimates from the choice model. As when generating the z-scores in the main text, $\hat{\alpha}_h$ is normalized to average to 0 in each market, so the same is true of $\hat{MRS}_h$. Appendix Figure A1 illustrates the distribution of MRS across hospitals nationally.

C.3 Patient-specific markets

We also estimate revealed preferences using an alternative market concept, namely one in which we define the choice set for each patient as hospitals within a 50 mile radius of their ZIP code of residence. This approach avoids the restrictive assumptions of fixed geographic markets. For instance, it is more likely to capture the relevant choice set for patients living near the borders of their HRRs, since these patients likely see hospitals on the other side of the borders as options. However, this approach does not lend itself to market-by-market

estimation of revealed preferences since there is no longer a fixed set of markets partitioning patients and hospitals. Instead, we must take a national approach to revealed preference estimation.

We start with the same patient sample described in Section 3.1. Then, we impose similar choice set restrictions to those used in the main analyses, with some simplifications. For a given patient, we define the choice set as all valid hospitals located within 50 miles of their ZIP code of residence. With this approach, we drop the 9% of patients who travel more than 50 miles to the hospital (analogously, we dropped the 13% of hospitals who exited their HRR of residence in the main analyses). For a hospital to be a valid choice, we require that among patients within 50 miles of the facility, its market share is at least 0.1%. This restriction helps to remove hospitals that are not truly destinations for emergent patients, similar to the 0.1% market share restriction used in the main analysis.

In principle, one could estimate a conditional logistic regression with this data. In practice, doing so would not be tractable even on a well-equipped server. The design matrix of the hospital fixed effects would be very large, equivalent to adding thousands of columns of dummy variables to a dataset that already spans all hospital choices over a decade in Medicare. This size would overwhelm the memory available on most servers. Even with sufficient memory, the model would converge extremely slowly.

Our approach leverages a set of equivalences that allow us to obtain the same point estimates using Poisson regression on aggregated data. We can therefore exploit recent advances that permit rapid, memory-saving estimation of fixed effects Poisson models. The equivalences connect conditional logistic regression, Poisson regression, and aggregated Poisson regression. Consider the following Poisson regression in which data is at the patient-by-choice level:

$$C_{ph} = \exp(\alpha_p + \alpha_h + \rho_1 \cdot distance_{ph} + \rho_2 \cdot distance_{ph}^2) + \varepsilon_{ph},$$

in which the choice indicator is regressed on patient fixed effects, hospital fixed effects, and

a quadratic for the distance from patient to hospital.[13] The likelihood function for this regression is identical, up to a constant, to the likelihood function for a conditional logistic regression of choice on hospital effects and a quadratic in distance. Both therefore return the same maximum likelihood estimates $\hat{\alpha}_h$, $\hat{\rho}_1$, and $\hat{\rho}_2$.[14]

We exploit a further equivalence to make this model tractable. Note that all right-hand side variables in the above equation are defined at the ZIP code-by-hospital level. This is because we measure patient location based on ZIP code of residence; all patients in a ZIP code have the same hospital choices, the same hospital fixed effect design matrix, and the same vector of distances to hospitals. We can therefore aggregate up to the ZIP code-by-hospital level and run the following Poisson regression:

$$N_{zh} = \exp(\alpha_z + \alpha_h + \rho_1 \cdot distance_{zh} + \rho_2 \cdot distance_{zh}^2) + \varepsilon_{zh},$$

where N_{zh} is the number of patients in ZIP code z choosing hospital h. This regression again has the same likelihood up to a constant and yields identical estimates $\hat{\alpha}_h$, $\hat{\rho}_1$, and $\hat{\rho}_2$. We estimate this regression using the R "fepois" function, which converges rapidly.

In the main analyses, the hospital effects are only separately identified within a market. This robustness analysis, which is essentially an AKM model but with ZIP codes and hospitals instead of workers and firms, has related identification conditions. Here, hospital effects are separately identified within mobility groups, i.e. connected subgraphs in the graph in which nodes are hospitals and ZIP codes and edges denote hospitals in the choice sets of ZIP codes.

We extract the hospital effect estimates from the Poisson regression. Because the effects are not separately identified across mobility groups, in each HRR we identify the mobility group with the most hospitals and restrict to these facilities. In the final regression, 99.6%

[13] We use distance in miles, not travel time in minutes, for this model. Given the similarity in the ultimate results, the change does not appear to be consequential.

[14] For a derivation of the general equivalence between multinomial logistic regression and Poisson regression, see Guimarães (2004).

of the included hospitals are in the single largest mobility group, with additional mobility groups each in the Hawaii and Alaska HRRs. We demean the hospital effects within each HRR, yielding effect estimates $\hat{\alpha}_h$ that are comparable to those in the main analysis.

An additional consideration is inference, since the procedure to create z-scores while adjusting for measurement error (Appendix A) requires standard errors for the $\hat{\alpha}_h$. We calculate them with the Bayesian bootstrap. We use the Bayesian, rather than traditional, bootstrap because it ensures that mobility groups in the bootstrap replications will match those in the original run – since all observations get positive weights, any hospital with positive market share in a ZIP code will continue to have positive market share in each bootstrap replication.

Consider the original patient-level sample used to calculate the N_{zt}. If we weight each patient ω_p, we can define a patient count for a bootstrap replicate run as $\tilde{N}_{zt} = \sum_{\{p \,:\, C_{ph}=1\}} \omega_p$. We let the weights be distributed Exp(1), which means $\tilde{N}_{zt}$ is the sum of N_{zt} independent Exponential random variables; thus $\tilde{N}_{zt} \sim \text{Gamma}(N_{zt}, 1)$. We generate the patient counts for each bootstrap replicate run by sampling from the Gamma distribution in this way. We conduct 300 bootstraps and calculate standard errors for each $\hat{\alpha}_h$ as the standard deviation of the hospital's fixed effect across the replicate runs.

Finally, we z-score the $\hat{\alpha}_h$ within each HRR so that these objects are as comparable as possible to the revealed preference z-scores of the main model. We use the measurement error adjustment described in Appendix A when calculating the standard deviation to generate the z-scores. In the robustness regressions, we also adjust the R^2 for measurement error with these procedures.

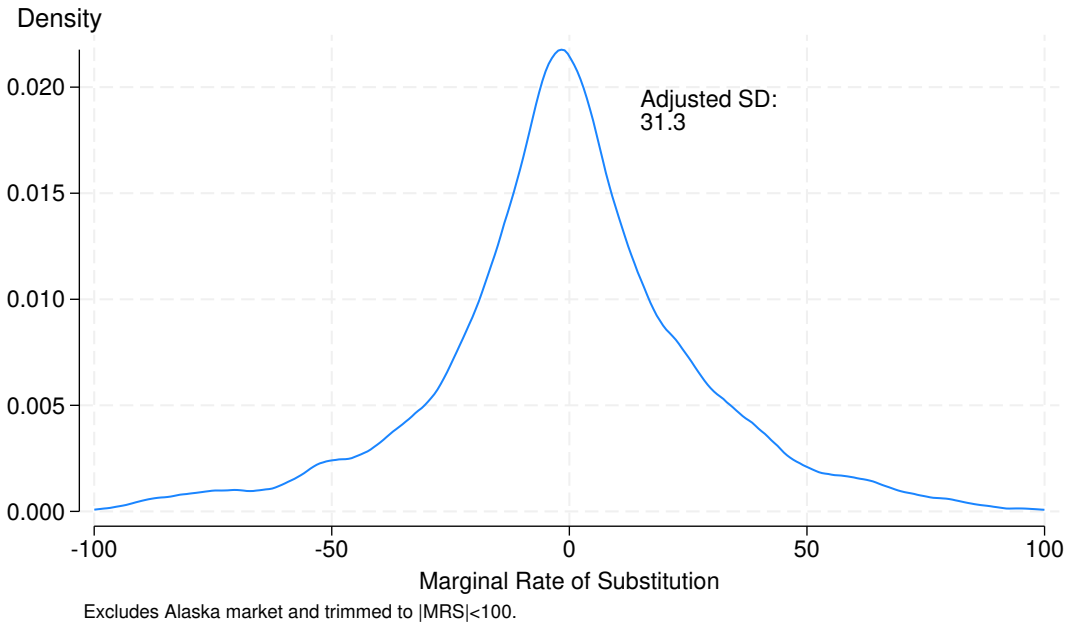
Appendix References

CMS, "Model Software," 2015.

Guimarães, Paulo, "Understanding the Multinomial-Poisson Transformation," *The Stata Journal: Promoting communications on statistics and Stata*, August 2004, *4* (3), 265–273.

Appendix Figures and Tables

Figure A1: Distribution of hospital marginal rates of substitution



Note: Figure illustrates distribution of across hospitals of the marginal rate of substitution between hospital choice and travel time. Marginal rates of substitution are normalized such that the within-market mean is 0. The 1.6% of hospitals with absolute marginal rates of substitution over 100, including all hospitals in the Alaska market, are suppressed from plot.

Table A1: Evidence-based care measures included in measure z-score

Acute Myocardial Infarction
Percutaneous Coronary Intervention at Arrival
Statin Prescription Provided at Discharge
Aspirin Prescription Provided at Discharge
Congestive Heart Failure
LVS Function Evaluated
ACE or ARB Provided for LVSD
Pneumonia
Appropriate initial antibiotic selection for community-acquired pneumonia
Stroke
VTE prophylaxis or documented reason by day of or day after admission
Assessed for rehabilitation services
Thrombolytic therapy within three hours, for patients arriving within two hours
Antithrombotic therapy administered by end of hospital day 2
Statin prescribed at discharge
Educational materials provided during hospital stay*
Anticoagulation prescribed for atrial fibrillation/flutter at discharge
Surgery
Beta-blocker continued perioperatively for prior users
Prophylactic antibiotic within one hour prior to incision
Prophylactic antibiotics consistent with guidelines
Prophylactic antibiotics discontinued within 24 hours after surgery end time
Urinary catheter removed on postoperative day 1 or 2
Appropriate VTE prophylaxis within 24 hours before to 24 hours after surgery

Note: Measures drawn from 2014 Care Compare process-of-care indicators reported for over 1000 hospitals. Each measure is estimated as the share of patients satisfying the criteria among patients for whom the measure is applicable. Abbreviations: LVS = Left Ventricular Systolic; VTE = Venous Thromboembolism; ACE = Angiotensin-Converting Enzyme; ARB = Angiotensin II Receptor Blocker.

Table A2: Patient experience measures included in measure z-score

Cleanliness of hospital environment
Communication with doctors
Communication with nurses
Communication about medicines
Responsiveness of hospital staff
Given information on symptoms to look for after discharge
Pain is well-controlled
Quietness of hospital environment
Patients understood care
Would recommend the hospital
Overall hospital rating

Note: Measures drawn from 2014 HCAHPS responses.

Table A3: Sample construction flow

	Hospitals	Patients
Initial sample	4,644	10,957,763
Require patients treated in market of residence	4,620	9,530,534
Require a valid travel time	4,620	9,526,721
Require hospitals meet choice set restrictions = Choice model starting sample	4,450	9,521,161
Exclude markets where choice model did not converge and hospitals where fixed effect could not be estimated = Choice model estimation sample	4,430	9,498,551
Exclude hospitals with incomplete quality or structural measures = Final regression sample	3,662	9,309,821

Notes: This table illustrates the flow of hospitals and patients in our analyses as we impose inclusion and exclusion criteria. The initial sample is described in Section 3.1 in the main text. The next steps impose criteria required for entry into the choice model starting sample and are described in Section 3.3. We then run the choice model market-by-market on this sample. Rarely, the choice model fails to converge for a market or the hospital fixed effect cannot be estimated. Omitting these markets and hospitals, which represent a tiny fraction of patients, yields the choice model estimation sample. For regression analyses in the main text, we further limit to hospitals with complete data on hospital quality and structural measures. This step excludes disproportionately small facilities – they account for 17% of hospitals but just 2% of patient episodes. The result is the final regression sample analyzed in the text.

Table A4: Revealed preference rankings of hospitals in Boston and Manhattan

A. Boston, Massachusetts

Rank	Hospital	Z-score	Market Share
1	Massachusetts General Hospital	1.2	7%
2	Lahey Hospital	1.2	5
3	Beth Israel Deaconess Medical Center	1.1	5
4	Brigham and Women's Hospital	1.0	4
5	Mount Auburn Hospital	0.8	2
6	North Shore Medical Center	0.8	4
7	Newton-Wellesley Hospital	0.6	3
8	Beverly Hospital	0.6	4
9	Southcoast Hospitals Group	0.6	9
10	St. Elizabeth's Medical Center	0.6	2

B. Manhattan, New York

Rank	Hospital	Z-score	Market Share
1	New York Presbyterian Hospital	1.9	15%
2	The Mount Sinai Hospital	1.6	7
3	NYU Langone Hospitals	1.4	8
4	Mount Sinai Beth Israel	1.4	9
5	Mount Sinai Morningside	1.2	6
6	Lenox Hill Hospital	1.1	4
7	Staten Island University Hospital	0.8	7
8	Maimonides Medical Center	0.4	12
9	New York Presbyterian Brooklyn Methodist	0.4	6
10	NYC Health + Hospitals / Bellevue	0.3	1

Note: This table presents the revealed preference rankings, z-scores, and market shares for the top 10 hospitals in two markets as estimated from all patients. Panel A presents Boston, Massachusetts (HRR = 227) and Panel B presents Manhattan, New York (HRR = 303). Abbreviations: Hospital Referral Region (HRR).

Table A5: Relationship between hospital revealed preferences and other ranking systems

	(1)	(2)	(3)
US News Top 50	1.176		1.159
	(0.0778)		(0.0809)
CMS Star Rating			
1		ref.	ref.
2		0.0498	-0.0239
		(0.100)	(0.0972)
3		-0.231	-0.304
		(0.102)	(0.0985)
4		-0.0608	-0.151
		(0.106)	(0.102)
5		0.525	0.405
		(0.151)	(0.155)
R^2, within	0.0341	0.0212	0.0541
R^2, within, ME adjusted	0.0342	0.0212	0.0542
Hospitals	3358	3358	3358

Note: Table predicts standardized hospital revealed preferences using alternative ranking systems including the US News and World Reports and CMS Star Ratings. The regressions include market fixed effects. US News Top 50 is an indicator for whether the hospital is listed as a top 50 facility in any specialty relevant to study conditions (cardiology, neurology, pulmonology, and orthopedics) in 2014 by US News and World Reports. Abbreviations: Measurement Error (ME), Centers for Medicare and Medicaid Services (CMS).

Table A6: Relationship between hospital characteristics, including structural features, and the revealed preferences of experts and non-experts

	Non-Experts (1)	Experts (2)	Experts (3)	Experts (4)
Quality measures				
Mortality, 30 Days in percentage points	-0.107	-0.166		-0.0733
	(0.0182)	(0.0272)		(0.0218)
Readmission, 30 Days in percentage points	-0.0793	-0.0498		0.0186
	(0.0270)	(0.0470)		(0.0404)
Evidence-Based Care Z-score	0.260	0.162		-0.0626
	(0.0753)	(0.0906)		(0.0664)
Patient Experience Z-score	-0.0653	-0.000173		0.0562
	(0.0399)	(0.0577)		(0.0501)
Hospital structural features				
For-Profit	-0.0770	0.0135		0.0800
	(0.0726)	(0.109)		(0.0909)
Government	-0.166	-0.157		-0.0140
	(0.0821)	(0.179)		(0.165)
Medical School Affiliation	0.325	0.487		0.206
	(0.0531)	(0.0956)		(0.0841)
Robotic Surgery	0.455	0.431		0.0384
	(0.0593)	(0.0930)		(0.0901)
Cardiac Catheterization	0.760	0.597		-0.0586
	(0.0664)	(0.0972)		(0.0707)
Non-Expert Rev Pref Z-score			0.905	0.863
			(0.0445)	(0.0577)
R^2, within	0.471	0.240	0.440	0.451
R^2, within, ME-adjusted	0.472	0.448	0.821	0.843
Hospitals	993	993	993	993

Note: This table presents a version of Table 3 that additionally includes hospital structural features as controls. See notes to Table 3 for more details. All regressions include market fixed effects. Both R^2 are within-market; they are unadjusted and adjusted for regressand measurement error, respectively. Robust standard errors in

Table A7: Robustness of relationship between hospital characteristics and revealed preferences

	Baseline	Hip/Knee Replacement	Marginal Rate of Substitution	Exclude winter months	Patient-specific markets
	(1)	(2)	(3)	(4)	(5)
Mortality, 30 Days in percentage points	-0.211		-6.185	-0.212	-0.216
	(0.0109)		(0.385)	(0.0109)	(0.0110)
Readmission, 30 Days in percentage points	-0.301	-0.225	10.69	-0.300	-0.300
	(0.0174)	(0.0186)	(0.619)	(0.0174)	(0.0172)
Evidence-Based Care Z-score	0.340	0.436	-11.60	0.341	0.343
	(0.0222)	(0.0588)	(0.825)	(0.0223)	(0.0230)
Patient Experience Z-score	-0.250	0.136	10.22	-0.249	-0.243
	(0.0227)	(0.0255)	(0.731)	(0.0227)	(0.0227)
R^2, within	0.273	0.167	0.289	0.272	0.275
R^2, within, ME adjusted	0.274	0.168	0.289	0.273	0.276
Hospitals	3662	2983	3652	3662	3652

Note: This table presents sensitivity analyses of Table 2 column 2 (repeated here in column 1). Column 2 analyzes revealed preferences and quality for hip and knee replacement. In column 3, the outcome is a marginal rate of substitution between hospital choice and travel time rather than a Z-score (more negative values indicate more preferred hospitals). Column 4 excludes observations from winter months when calculating revealed preference Z-scores. Column 5 use a national gravity model to allow patient-specific markets and choice sets (hospitals within 50 miles of the patient's residence) to estimate revealed preference Z-scores. See Appendix C for more details. All regressions include market fixed effects. Robust standard errors in parentheses.

Table A8: Robustness of relationship between traditional quality measures and the revealed preferences of experts and non-experts

	Baseline		Hip and Knee Replacement		Marginal Rate of Substitution	
	(1)	(2)	(3)	(4)	(5)	(6)
Mortality, 30 Days in p.p.	-0.194	-0.247			3.813	3.608
	(0.0230)	(0.0309)			(0.544)	(0.569)
Readmission, 30 Days in p.p.	-0.206	-0.159	-0.170	-0.161	5.923	3.554
	(0.0331)	(0.0500)	(0.0336)	(0.0428)	(0.914)	(0.844)
Evidence-Based Care Z-score	0.641	0.518	0.646	0.389	-17.94	-10.99
	(0.0940)	(0.107)	(0.121)	(0.151)	(3.131)	(2.504)
Patient Experience Z-score	-0.146	-0.0866	0.240	0.314	6.232	2.763
	(0.0488)	(0.0622)	(0.0481)	(0.0725)	(1.138)	(1.116)
Patient sample	Non-Experts	Experts	Non-Experts	Experts	Non-Experts	Experts
R^2, within	0.188	0.104	0.146	0.0815	0.199	0.120
R^2, within, ME-adjusted	0.189	0.194	0.147	0.136	0.199	0.139
Markets	47	47	50	50	47	47
Hospitals	993	993	781	781	993	993
	Excluding winter months		Patient-specific markets			
	(7)	(8)	(9)	(10)		
Mortality, 30 Days in p.p.	-0.194	-0.214	-0.180	-0.222		
	(0.0243)	(0.0300)	(0.0237)	(0.0292)		
Readmission, 30 Days in p.p.	-0.208	-0.223	-0.241	-0.182		
	(0.0338)	(0.0481)	(0.0353)	(0.0430)		
Evidence-Based Care Z-score	0.702	0.446	0.651	0.567		
	(0.0999)	(0.111)	(0.0884)	(0.0799)		
Patient Experience Z-score	-0.149	-0.0725	-0.109	0.0143		
	(0.0515)	(0.0615)	(0.0476)	(0.0580)		
Patient sample	Non-Experts	Experts	Non-Experts	Experts		
R^2, within	0.190	0.105	0.206	0.145		
R^2, within, ME-adjusted	0.191	0.169	0.207	0.214		
Markets	47	47	45	45		
Hospitals	921	921	907	907		

Note: This table presents sensitivity analyses of Table 3 columns 1-2 (repeated here in columns 1-2). Columns 3-4 analyze revealed preferences and quality for hip and knee replacement. In columns 5-6, the outcome is a marginal rate of substitution between hospital choice and travel time rather than a Z-score (more negative values indicate more preferred hospitals). Columns 7-8 exclude observations from winter months when calculating revealed preference Z-scores. Columns 9-10 use a national gravity model to allow patient-specific markets and choice sets (hospitals within 50 miles of the patient's residence) to estimate revealed preference Z-scores. See Appendix C for more details. All regressions include market fixed effects. Robust standard errors in parentheses.

Table A9: Robustness of relationship between revealed preferences of non-experts and experts

	Baseline	Hip/Knee Replacement	Marginal Rate of Substitution	Exclude winter months	Patient-specific markets
	(1)	(2)	(3)	(4)	(5)
Revealed Preference Z-score, Non-Experts	0.905	0.927		0.840	0.868
	(0.0445)	(0.0388)		(0.0478)	(0.0317)
Revealed Preference MRS, Non-Experts			0.676		
			(0.0415)		
R^2, within	0.440	0.514	0.561	0.440	0.513
R^2, within, ME adjusted	0.821	0.861	0.653	0.708	0.759
Markets	47	50	47	47	45
Hospitals	993	781	993	921	921

Note: This table presents sensitivity analyses of Table 2 column 3 (repeated here in column 1). Column 2 analyzes revealed preferences and quality for hip and knee replacement. In column 3, the outcome and regressor are marginal rates of substitution between hospital choice and travel time rather than a Z-score (more negative values indicate more preferred hospitals). Column 4 excludes observations from winter months when calculating revealed preference Z-scores. Column 5 use a national gravity model to allow patient-specific markets and choice sets (hospitals within 50 miles of the patient's residence) to estimate revealed preference Z-scores. See Appendix C for more details. All regressions include market fixed effects. Robust standard errors in parentheses.

Table A10: Robustness of association between patient characteristics and the revealed preferences of their chosen hospitals

	Baseline		Hip and Knee Replacement		Marginal Rate of Substitution	
	(1)	(2)	(3)	(4)	(5)	(6)
Race-Ethnicity						
White	ref.	ref.	ref.	ref.	ref.	ref.
Black	-0.0476	-0.0373	-0.0575	-0.0725	-0.812	0.449
	(0.00102)	(0.000859)	(0.00237)	(0.00223)	(0.0225)	(0.0189)
Hispanic	-0.113	-0.0223	-0.213	-0.117	1.897	0.0366
	(0.00146)	(0.00110)	(0.00343)	(0.00282)	(0.0412)	(0.0286)
Other/Unknown	0.00853	-0.000168	0.00448	-0.0226	-0.865	-0.449
	(0.00159)	(0.00123)	(0.00309)	(0.00260)	(0.0452)	(0.0304)
Age (Decades)	-0.0295	-0.0509	-0.0306	-0.0397	1.099	1.426
	(0.000328)	(0.000235)	(0.000827)	(0.000659)	(0.00944)	(0.00641)
Female	-0.0347	-0.0429	-0.0111	-0.0151	0.839	1.104
	(0.000546)	(0.000389)	(0.00104)	(0.000836)	(0.0158)	(0.0107)
Medicaid eligible	-0.214	-0.112	-0.315	-0.213	4.965	2.557
	(0.000800)	(0.000584)	(0.00277)	(0.00226)	(0.0226)	(0.0156)
No. of Comorbidities*	-0.0955	-0.0956	-1.266	-0.986	-0.352	2.626
	(0.00825)	(0.00595)	(0.0425)	(0.0348)	(0.224)	(0.152)
Area Avg Income ($100s)	0.0683		0.0636		-1.496	
	(0.0000935)		(0.000177)		(0.00220)	
Rural	-0.558		-0.308		19.12	
	(0.00185)		(0.00306)		(0.0716)	
Geographic Controls	Market	ZIP	Market	ZIP	Market	ZIP
Patients (N)	9,131,832	9,131,659	2,397,008	2,396,087	9,120,620	9,120,457

Note: This table presents sensitivity analyses of Table 4 columns 2-3 (repeated here in columns 1-2). Columns 3-4 analyze revealed preferences and quality for hip and knee replacement. In columns 5-6, the outcome is a marginal rate of substitution between hospital choice and travel time rather than a Z-score (more negative values indicate more preferred hospitals). See Appendix C for more details. All regressions include market fixed effects. Robust standard errors in parentheses. *Coefficients for comorbidities are multiplied by 100 for ease of exposition.

Table A11: Robustness of association between patient characteristics and the revealed preferences of their chosen hospitals, continued

	Baseline		Exclude winter months		Patient-specific markets	
	(1)	(2)	(3)	(4)	(5)	(6)
Race-Ethnicity						
White	ref.	ref.	ref.	ref.	ref.	ref.
Black	-0.0476	-0.0373	-0.0453	-0.0367	-0.0645	0.0373
	(0.00102)	(0.000859)	(0.00102)	(0.000859)	(0.00104)	(0.000888)
Hispanic	-0.113	-0.0223	-0.111	-0.0219	-0.147	-0.0260
	(0.00146)	(0.00110)	(0.00146)	(0.00110)	(0.00148)	(0.00114)
Other/Unknown	0.00853	-0.000168	0.00922	-0.000155	0.00879	0.00488
	(0.00159)	(0.00123)	(0.00159)	(0.00123)	(0.00156)	(0.00126)
Age (Decades)	-0.0295	-0.0509	-0.0299	-0.0513	-0.0312	-0.0518
	(0.000328)	(0.000235)	(0.000329)	(0.000235)	(0.000329)	(0.000242)
Female	-0.0347	-0.0429	-0.0348	-0.0431	0.0370	-0.0435
	(0.000546)	(0.000389)	(0.000546)	(0.000389)	(0.000547)	(0.000400)
Medicaid eligible	-0.214	-0.112	-0.214	-0.113	-0.216	-0.115
	(0.000800)	(0.000584)	(0.000800)	(0.000584)	(0.000803)	(0.000601)
No. of Comorbidities*	-0.0955	-0.0956	-0.0928	-0.955	-0.112	-0.0969
	(0.00825)	(0.00595)	(0.00825)	(0.00595)	(0.00829)	(0.00612)
Area Avg Income ($100s)	0.0683		0.0682		0.0631	
	(0.0000935)		(0.0000936)		(0.0000952)	
Rural	-0.558		-0.559		-0.528	
	(0.00185)		(0.00186)		(0.00180)	
Geographic Controls	Market	ZIP	Market	ZIP	Market	ZIP
Patients (N)	9,131,832	9,131,659	9,131,832	9,131,659	9,133,344	9,133,164

Note: This table presents sensitivity analyses of Table 4 columns 2-3 (repeated here in columns 1-2). Columns 3-4 exclude observations from winter months when calculating revealed preference Z-scores. Columns 5-6 use a national gravity model to allow patient-specific markets and choice sets (hospitals within 50 miles of the patient's residence) to estimate revealed preference Z-scores. See Appendix C for more details. All regressions include market fixed effects. Robust standard errors in parentheses. *Coefficients for comorbidities are multiplied by 100 for ease of exposition.