Demand for Quality in the Presence of Information Frictions: Evidence from the Nursing Home Market

Alden Cheng

March 31, 2023

Click Here for Most Recent Version

Abstract

This paper studies consumers’ demand for quality in the nursing home market, where information frictions are a source of concern. Using administrative data on the universe of nursing home residents, I estimate quality of nursing homes in California, and use these estimates as inputs into a structural demand model. I find substantial variation in nursing home quality: one standard deviation higher quality is associated with 2 percent lower risk-adjusted 90-day mortality rate. Yet, despite the high stakes for residents, average demand for quality is very low, even after accounting for unobserved supply-side constraints arising from selective admissions practices by nursing homes. Patterns of demand heterogeneity highlight information frictions as a major reason for this low demand: residents who were younger, highly educated, free from dementia, and who made their choices after the introduction of the star rating system were more responsive to quality. Counterfactual simulations based on estimates of the structural demand model and a competing risks model suggest that eliminating information frictions can reduce deaths by at least 8 to 28 percent, and potentially even more if supply side responses are considered.

I am incredibly grateful to Alberto Abadie, Joseph Doyle, Amy Finkelstein, Jon Gruber, and Ben Handel for their support and advice. In addition, Leila Agha, Josh Angrist, Nikhil Agarwal, MacKenzie Alston, Meta Brown, David Card, David Chan, Ben Chartock, Victor Chernozhukov, Jonathan Cohen, Kenneth Couch, Stefano DellaVigna, Delia Furtado, Tatyana Deryugina, Todd Elder, Ashvin Gandhi, Lorenz Goette, Aaron Goodman, Martin Hackmann, Adam Harris, Daniel Keniston, Kyoo il Kim, Kurt Lavetti, Jetson Leder-Luis, Haizhen Lin, Diana Sverdliv Lisker, Justin Marion, Philip Marx, Martin Mattsson, Nidhiya Menon, David Molitor, Kelsey Moran, Anh Nguyen, Francesc Ortega, Abigail Ostriker, Elena Prager, Charlie Rafkin, Julian Reif, Nina Roussille, Anna Russo, Akshar Saxena, Frank Schilbach, Mark Shepard, Maggie Shi, Ben Shiller, George Siddarth, David Simon, Rahul Singh, Martina Uccioli, Gerry Wedig, Jeremy West, Laura Wu, Jia Xiang, Ariel Zucker, and seminar participants at Gies College of Business, MIT, Nanyang Technological University, and Queens College, CUNY provided valuable feedback. I also thank Mohan Ramanujan at the NBER for his assistance. All mistakes are my own.
1 Introduction

Health expenditures in the US add up to roughly 20 percent of GDP. Yet, healthcare quality in certain sectors leaves much to be desired, with nursing homes being a prime example. An important factor contributing to this underinvestment in quality is information frictions (Arrow 1963; Gaynor 2006; Salop and Stiglitz 1977): when consumers are poorly informed, firms are incentivized to supply suboptimal amounts of quality (Dranove and Satterthwaite 1992).¹

Despite the importance of information frictions about quality of care, limited empirical has studied this directly; this is because of the inherent difficulties in measuring quality (due to issues such as selection bias and misreporting), and the fact that consumer misperceptions about quality of care are often difficult to identify and quantify.² Partly due to these reasons, most of the literature in this space has focused on settings where mistakes are more clear cut, such as insurance demand (Abaluck and Gruber 2012, 2016, 2020; Handel 2013; Handel and Kolstad 2015; Handel, Kolstad, Minten, and Spinniewijn 2021) and treatment decisions (Kolstad 2013; Chan, Gentzkow, and Yu 2022; Mullainathan and Obermeyer 2022).³ However, it is unclear whether lessons learned from these settings apply to consumer choice over healthcare quality, given that the nature of information frictions is very different qualitatively.⁴ Therefore, this paper takes a step in filling this gap in the literature by measuring information frictions about quality of care in nursing homes and quantifying the consequences.

I choose the nursing home setting to study this issue for three reasons. First, this is a setting wherein quality significantly impacts consumer wellbeing. Indeed, understaffing, abuse, and negligence has been well documented, especially at for-profit nursing homes (Comondore et al. 2009), and this has often resulted in severe health consequences, even death, for residents (Washington Post 2020). Second, poor nursing home quality is an issue of great interest to policymakers; for example, President Biden

¹Highlighting the perceived difficulty of assessing nursing home quality, Senator Ron Wyden remarked that it was “easier to shop for washing machines than it is to select a nursing home” during a congressional hearing in 2007 (New York Times 2007).

²It is challenging to identify information frictions about quality of care from choice data due to heterogeneous preferences (Handel and Schwartzstein 2018). In principle, researchers can identify information frictions using surveys that test consumers’ knowledge of certain publicly available quality measures. However, it is often hard to interpret the magnitudes of such survey results without a better understanding of consumers’ preferences for these quality measures, and the fact that these measures are publicly available also raises the possibility that they may be gamed by care providers.

³In the insurance setting, there is a natural dollar metric, which makes it easier to identify mistakes (e.g., choosing dominated plans) and to quantify the costs of these errors. As for treatment decisions, there are certain medical settings wherein the appropriate treatment is relatively uncontroversial.

⁴In the case of health insurance demand, underlying reasons for suboptimal choices often involve consumer misperception about their likely future health expenditures and consumer innumeracy in general, which are both different from consumers’ perceptions about the quality of care providers. Similarly, in the case of treatment decisions, the decision maker is typically an “expert” (having received years of medical training) operating in a repeated choice environment (having faced similar treatment decisions for previous patients), whereas consumers choosing their care provider are typically not specifically trained in this area and are less experienced with making this choice.
proposed a slew of reforms aimed at improving nursing home quality in his 2022 State of the Union speech. Third, nursing homes are an important part of the healthcare sector: 1.3 million Americans live in nursing homes (CDC), and more than half of those aged 57–61 today will spend some time in a nursing home (Hurd, Michaud, and Rohwedder, 2017) — and thus lessons learned about this industry are valuable in themselves.

To study information frictions about nursing home quality, I proceed in three steps. First, I estimate quality of nursing homes in California based on risk-adjusted mortality, a commonly used health outcome in the economics literature (Doyle, Graves, Gruber, and Kleiner 2015; Deryugina and Molitor 2020; Finkelstein, Gentzkow, and Williams 2021; Abaluck, Bravo, Hull, and Starc 2021). I avoid overfitting with the large set of potential controls using a variable selection method motivated by double machine learning (Belloni, Chernozhukov, and Hansen, 2014), and I account for statistical imprecision by estimating quality using empirical Bayes. In addition, to address concerns about selection on unobservables, I use a validation test from the value-added literature, demonstrating that variation in the quality of residents’ chosen nursing homes induced by differential distances to their prior address predicts resident outcomes one-for-one.

In the second step, I use these quality estimates to study residents’ demand for quality, taking into account unobserved supply-side constraints arising from selective admissions practices by nursing homes (Gandhi 2019). I identify the structural demand model using distance and temporary occupancy fluctuations as demand and supply instruments, and I estimate the model using Gibbs sampling with data augmentation to avoid the curse of dimensionality (Agarwal and Somaini 2022).

In the final step, I simulate the consequences of information frictions, the sign of which are a priori unclear due to the theory of the second best. To do this, I combine the structural demand model

---

5The empirical Bayes model reduces the mean-squared error of the quality estimates by shrinking them towards the mean according to their variance.

6More precisely, by “residents’ demand for quality” I refer to the demand for quality of the decision maker, which is not necessarily the resident (e.g., it could be a family member). This distinction is not important for the main findings of this paper, but may be relevant to follow-up questions (e.g., regarding the design of an effective information intervention).

7The number of possible eligible choice sets for each resident increases exponentially with the number of potential nursing homes in her choice set (exceeding $10^{60}$ for some residents), and thus methods such as maximum likelihood are computationally infeasible.

8I use a broad definition of information frictions that encompasses suboptimal choices owing to the costs of acquiring and processing information, as well as those that result from psychological distortions in the process (Handel and Schwartzstein, 2018).

9This is because there are multiple market imperfections: information frictions and selective admissions practices by nursing homes. In particular, reducing information frictions for a given cohort of residents imposes a negative externality on future cohorts by increasing occupancy at high-quality nursing homes, making it more likely that these nursing homes will reject future residents. Handel (2013) provides an example of how eliminating one market imperfection may worsen outcomes when there are multiple imperfections (in the context of inertia and adverse selection).

10More precisely, by the “costs of information frictions”, I refer to the potential deaths averted if residents’ demand for quality (based on their decision utility) resemble the demand for quality previously estimated in other healthcare settings.
(which allows me to simulate admissions while accounting for nursing homes’ selective admissions practices) with a competing risks model (which I use to simulate deaths and discharges).\textsuperscript{11} I also consider the long-run effect of eliminating information frictions (where I allow nursing homes to adjust their quality) using a simple model of quality competition between nursing homes, which I calibrate using the introduction of five-star ratings by the Centers for Medicare and Medicaid Services (CMS) at the end of 2008.

My nursing home quality estimates reveal substantial variation in quality across nursing homes in California (both overall and locally), and the validity of these estimates is supported by the distance-based instrumental variables (IV) strategy. The estimates imply that a resident who goes to a nursing home with one standard deviation higher quality is 2 percent less likely to die within 90 days of admission (all else equal), a 27 percent reduction relative to the baseline mortality rate. Moreover, my quality measure is correlated in expected ways with publicly available nursing home characteristics (such as skilled staffing levels and cited deficiencies), but provides additional predictive power for resident outcomes above and beyond these variables.

Despite the high stakes for residents, I estimate that average demand for quality is an order of magnitude smaller than existing estimates in the literature.\textsuperscript{12} Additionally, residents do not seem to take full advantage of publicly available information about nursing home quality, suggesting that information frictions may be present. Moreover, patterns of demand heterogeneity provide further evidence that information frictions may be responsible for the low demand: older and cognitively impaired residents are less responsive to quality differences, whereas residents who have at least a Bachelor’s degree, and who made their choices after an information intervention by the CMS (specifically, the introduction of the five-star ratings system) are more responsive to quality differences.

My counterfactual simulations suggest that eliminating information frictions may reduce deaths by at least 8–28 percent (or 250–850 deaths annually in California nursing homes) in the short run (i.e., keeping quality fixed), and potentially several times more in the long run (when allowing nursing homes to adjust quality in response to changes in demand). Moreover, I find that eliminating information frictions has favorable distributional consequences: reduction in mortality is concentrated among residents with the greatest baseline information frictions, and there is little evidence of negative

\textsuperscript{11}Simulating resident exits is necessary given that nursing homes’ admission decisions depend on occupancy. In addition, because mortality is the outcome of interest, resident exits must be decomposed deaths and discharges. Moreover, the fact that deaths and discharges are competing risks necessitates the use of a competing risks model.

\textsuperscript{12}The demand estimate implies that residents are willing to tolerate a 28.6 percent higher probability of dying within 90 days in exchange for a one-mile reduction in travel distance.
spillovers for any subgroup of residents due to increased crowding at high-quality nursing homes.

This paper is connected to several strands of literature. As mentioned earlier, my main results add to the literature on behavioral frictions in healthcare (Abaluck and Gruber 2012, 2016, 2020; Handel 2013; Handel and Kolstad 2015; Handel, Kolstad, Minten, and Spinnewijn 2021; Kolstad 2013; Chan, Gentzkow, and Yu 2022; Mullainathan, and Obermeyer 2022) by studying frictions in a choice environment that is very different qualitatively from previous work in this area (which typically focuses on insurance demand and treatment decisions). Relatedly, this paper also adds to a previous body of work on determinants of choice quality more generally. In particular, my finding that disadvantaged residents tend to make worse nursing home choices echoes results in other settings, such as Handel, Kolstad, Minten, and Spinnewijn (2021) in the context of insurance choice in the Netherlands and Walters (2012) in the context of school choice.

Methodologically, the present paper is one of the first to estimate demand while accounting for unobserved choice set constraints due to dynamic supply side incentives. The two main papers that previously estimate such a model are Gandhi (2019) and Agarwal and Somaini (2022), both of which I borrow from.13 Separately, my estimation of nursing home quality is related to a vast literature on value-added models (Chetty, Friedman, and Rockoff 2014; Abaluck, Bravo, Hull, and Starc 2021; Angrist, Hull, Pathak, and Walters 2021),14 and also provides one of the first applications of double machine learning in this area (Belloni, Chernozhukov, and Hansen, 2014; Chernozhukov et al. 2018).15

Finally, this paper contributes to a growing literature on nursing homes in economics. Much of the prior literature has focused on the supply-side, for example studying the effect of private equity ownership on nursing home behavior (Gandhi, Song, and Upadrashta 2020; Gupta, Howell, Yannelis, and Gupta 2021), admission and discharge decisions by nursing homes (Gandhi 2019; Hackmann, Pohl, and Ziebarth 2020), and the effect of reimbursement rates and competition (Hackmann, 2019),

---

13 This is related more broadly to a literature on demand estimation with unobserved choice set constraints arising from various sources. The source of these choice set constraints typically motivates the model and identification strategy, which may be based on consideration set formation (Abaluck and Adams-Prassl 2021; Barseghyan, Molinari, and Thir-kettle 2021; Barseghyan, Coughlin, and Teitelbaum 2021), consumer search (Abaluck and Compiani 2020), or two-sided matching (Menzel 2015; Diamond and Agarwal 2017; He, Sinha, and Sun 2020; Agarwal and Somaini 2022). Since the choice set constraints in this paper are due to nursing home behavior, I use a two-sided matching model for my demand estimation.

14 Much of the earlier literature on value-added estimation has focused on education settings (Chetty, Friedman, and Rockoff 2014; Angrist, Hull, Pathak, and Walters 2017; Angrist, Hull, Pathak, and Walters 2021), although more recently quality estimation has also become increasingly common in health economics (Fletcher, Horwitz, and Bradley 2014; Doyle, Graves, Gruber, and Kleiner 2015; Hull 2018; Finkelstein, Gentzkow, and Williams 2021; Abaluck, Bravo, Hull, and Starc 2021; Cooper, Doyle, Graves, and Gruber 2022; Einav, Finkelstein, and Mahoney 2022).

15 The challenge with using double machine learning in value-added estimation is that the “treatment variable” is typically a high-dimensional vector of choice dummies, whereas double machine learning typically deals with variable selection when there is only a single (or at most a few) treatment variable(s). I address this by using a model reduction strategy, before using the standard post-double-selection method to choose the appropriate set of controls (Belloni, Chernozhukov, and Hansen 2014).
although there is also a smaller number of studies on consumer choice (Gandhi 2019) and quality estimation (Einav, Finkelstein, and Mahoney 2022). The present paper contains elements of all three, but the main focus is on consumer choice.

This paper proceeds as follows. In section 2, I provide background on the nursing home industry and its residents, and describe the data I use for my analysis. I then estimate nursing home quality and validate these estimates in section 3, before using them to study residents’ demand for quality and information frictions in section 4. In section 5, I quantify the costs of information frictions via counterfactual simulations, and I conclude in section 6.

2 Background

2.1 Nursing Home Industry

There are roughly 15,000 nursing homes in the US providing care for about 1.3 million Americans (CDC), and an estimated 56 percent of Americans aged 57–61 are expected to spend at least one night in a nursing home during their lifetimes (Hurd, Michaud, and Roh wedder 2017). Nursing home residents vary widely in their medical conditions and needs, but as a crude approximation, there are two broad categories: short-stay and long-stay. Short-stay residents typically require rehabilitative care following an acute care hospital stay — for example, to recover from knee or hip replacement surgery. These patients are expected to recover sufficiently during their nursing home stay to be discharged, and are typically covered by Medicare. By contrast, long-stay residents often suffer from chronic conditions (e.g., cognitive decline), decreasing the likelihood that they will be discharged in the short term. Many of these residents are covered by Medicaid, which has substantially lower reimbursement rates than Medicare. Finally, short-stay residents account for the majority of nursing home admissions, but at any given point in time, roughly half of residents residing in nursing homes are long-stay.

Despite substantial health expenditures in nursing homes, which totaled approximately $170 billion in 2016 (or roughly 5 percent of all healthcare spending in the US), both CMS data and anecdotal evidence show that poor quality remains endemic. These issues are especially prevalent at for-profit

---

16More precisely, throughout this paper I colloquially refer to skilled nursing facilities (SNFs) that are certified by the CMS as nursing homes.
17Medicare reimburses up to 100 days following a hospital stay that lasts at least three days (with full coverage for the first 20 days and partial coverage for days 21–100). Reimbursement rates are also a function of resident care requirements and cost of inputs.
18Specifically, because long-term care insurance is uncommon in the US (Brown and Finkelstein, 2007, 2008, 2011), long-stay residents typically pay out-of-pocket until they become eligible for Medicaid (if they do not already qualify at the time of admission).
nursing homes (Comondore et al., 2009), which account for the majority of nursing homes. This paper focuses on one potential reason that the profit incentive has not resulted in greater quality investments: information frictions faced by residents about nursing home quality.

2.2 Choice of Quality Measure

Measuring information frictions about nursing home quality requires a definition of quality. In reality, nursing home quality is multidimensional, and reflecting this notion, a variety of quality measures have been used in the past, including skilled staffing levels, deficiency citations during annual inspections by regulators, resident outcomes, and the CMS five-star rating system introduced towards the end of my sample (which aggregates many of these components into a single index). While it may be difficult to account for all dimensions of quality in my analysis, most of these quality measures are likely to be positively correlated. Hence, I use 90-day risk-adjusted survival rate as my primary quality measure, which is easier to measure and interpret for several reasons.

First, death tends to be recorded quite accurately, whereas there is greater scope for intentional or unintentional misreporting for other quality measures. For example, gaming of staffing levels, resident outcomes, deficiency citations, and the five-star ratings by nursing homes has been well documented over the years. Moreover, even unintentional errors in misreporting can be problematic: for instance, if understaffing at low-quality nursing homes results in fewer adverse outcomes being discovered, then quality measures based on these outcomes will be biased upwards for low-quality nursing homes.

Second, death is less likely than other outcome-based measures to be affected by truncation bias. This refers to the scenario wherein a resident suffers an adverse outcome (other than death) but dies More recently, the rise of private equity ownership in the nursing home sector is also believed to have contributed to lower quality (Gandhi, Song, and Upadrashta 2020; Gupta, Howell, Yannelis, and Gupta 2021).

These inspections may also be conducted as a result of a complaint.

I will show evidence of this in my analysis.

Staffing levels were self-reported by nursing homes during the period of this study, and were often unaudited. A comparison with staffing numbers since payroll-based reporting was introduced in 2016 reveals that these earlier self-reported numbers are likely to be substantially inflated (Geng, Stevenson, and Grabowski 2019).

Schizophrenia diagnoses rose sharply in the years following 2012 when the government began publicly releasing information about inappropriate antipsychotic use (New York Times, 2021).

Some nursing homes temporarily increase inputs during the period wherein the annual inspections are expected to occur.

Because the five-star ratings are a function of deficiency citations, staffing levels, and resident outcomes, gaming of these former measures will affect the star ratings as well (New York Times, 2021).
before having another assessment, and thus the adverse outcome is not recorded in the data.\textsuperscript{26,27}

Third, death shortly after admission is an undesirable outcome for most types of residents.\textsuperscript{28} This is important given that I am studying residents from all payer sources, and goals of care vary substantially for different types of residents.\textsuperscript{29}

Finally, mortality is a commonly used health outcome in the prior economics literature (Doyle, Graves, Gruber, and Kleiner 2015; Deryugina and Molitor 2020; Finkelstein, Gentzkow, and Williams 2021; Abaluck, Bravo, Hull, and Starc 2021), which makes it easier to interpret the magnitude of effect sizes found in my study.\textsuperscript{30}

\subsection{Data}

The primary data source for this paper is the Minimum Data Set 2.0 (MDS). All nursing homes that receive federal funding are required to fill out MDS assessment forms at regular intervals (42 CFR §483.20).\textsuperscript{31,32,33} Data collected from the MDS assessments includes information on residents’ demographics, cognitive status, communication and hearing patterns, vision patterns, mood and behavior patterns, psychosocial well-being, physical functioning and structural problems, continence issues, disease diagnoses (including ICD-9 codes), health conditions, oral health, nutrition, dental status, skin

\textsuperscript{26}In theory, one could attempt to address this by estimating a hazard model to account for the censoring. However, even setting aside the difficulty of estimating many fixed effects in a hazard model, the assumption of uninformative censoring that many hazard models rely on is violated in this setting. This is because censoring induced by mortality is likely correlated with a host of other adverse outcomes in the time window shortly before death.

\textsuperscript{27}I do not observe deaths occurring in other settings after a resident has been discharged, so in this sense, 90-day mortality may also suffer from truncation bias. However, even if I did observe deaths after discharge from nursing homes, it is unclear whether nursing homes are at fault for these deaths. Moreover, in Section 3, I show that my quality estimates remain stable when I use mortality over shorter or longer durations than 90 days as the outcome, suggesting that truncation bias is unlikely to play a major role.

\textsuperscript{28}I drop a relatively small number of residents who were either comatose or already on hospice upon admission, given that short-term mortality may not be an appropriate quality measure for these residents.

\textsuperscript{29}For example, the quality measure based on discharge readiness recently developed by Einav, Finkelstein, and Mahoney (2022) is an excellent quality measure for short-stay residents, which is the population they study. However, it is a less appropriate quality measure for long-stay residents, who are by definition unlikely to be discharged in the short term.

\textsuperscript{30}In contrast, it is more difficult to interpret the magnitudes of demand based on various other quality measures. For example, the economic value of an additional stage two pressure sore is not well-established, which makes it difficult to determine whether a given level of demand for a quality measure based on this variable is “too large” or “too small.”

\textsuperscript{31}The set of nursing homes receiving federal funding account for roughly 96 percent of all nursing homes (Grabowski, Gruber, and Angelelli 2008).

\textsuperscript{32}Assessment forms must completed upon admission, at discharge (or death), quarterly in between, and whenever there is a significant change in status.

\textsuperscript{33}MDS forms are typically filled out by a registered nurse (RN), or at least certified by one. Any willful misrepresentation in the MDS forms may result in penalties under the False Claims Act. This is not limited to upcoding and variables that affect reimbursements directly but also other variables related to resident well-being. This is because nursing homes “must provide services to attain or maintain the highest practicable physical, mental, and psychosocial well-being of each resident” (42 CFR §1395i–3) to be certified to receive federal funding. Hence, any misrepresentation pertaining to resident wellbeing may be interpreted as being related to misrepresentation connected to a requirement for federal funding, and thus falls under the False Claims Act. Moreover, several studies on the accuracy of MDS data have found it to be fairly reliable (Shin and Scherer, 2009).
conditions, activity pursuit patterns, medications, special treatments and procedures, and discharge potential. The richness of this data will play an important role in the risk adjustment for quality estimation in the next section.

I supplement the MDS with data on nursing homes from other sources. This includes the Online Survey Certification and Reporting (OSCAR) surveys (which contain information such as nursing homes’ ownership status and staffing levels), data on deficiency citations, five-star ratings for nursing homes, and Medicare cost reports.

These data sets also contain the information needed to compute distances between residents’ prior addresses and nursing homes as well as nursing home occupancy levels over time, which are key components for my empirical strategies in sections 3 and 4. In particular, I compute distances by combining the five-digit zip code of residents’ prior addresses from the MDS with nursing homes’ street addresses from the OSCAR data, which I convert to GPS coordinates using the Google Maps API, and I determine nursing home occupancy levels using admission and discharge dates from the MDS.

I focus on the first stays of residents between the years 2000 and 2010, given that the MDS 2.0 and OSCAR data overlap over this period. In addition, I restrict the sample to nursing homes in California with at least 100 new stays over this period so that quality can be estimated relatively precisely. Because of the computational expense of the structural demand model, I further restrict my sample period to 2008–2010 for the demand estimation in section 4. Finally, I consider the single year of 2009 for my counterfactual simulations in section 5, given that various assumptions that the simulations depend on (e.g., counterfactual nursing home entries and exits) are more suspect over longer time horizons. More details about the data can be found in Appendix B.

Appendix Tables A.1 and A.2 contain a finer breakdown of the different types of data collected on each resident. The full MDS 2 assessment form is available on the CMS website.

The OSCAR data is available from 2000 onwards from LTCFocus.org, which is maintained by Brown University Center of Gerontology and Healthcare Research. LTCFocus is sponsored by the National Institute on Aging (1P01AG027296) through a cooperative agreement with the Brown University School of Public Health. Data on deficiencies, Medicare cost reports, and five-star ratings are available from the CMS website.

I also convert the five-digit zip codes for residents’ prior addresses into GPS coordinates using the U.S. Department of Housing and Urban Development’s ZIP Code Crosswalk Files for Quarter 1 of 2010, and I compute the distances between residents’ and nursing homes’ addresses using the “geodist” module in Stata.

I choose to study nursing homes in a single state mainly for computational reasons, but also because nursing home regulations differ across states. I choose California due to its large population, and because most of its population live relatively far away from state borders (and thus nursing homes in other states are relatively unlikely to be in nursing home residents’ choice sets).

Additional sample restrictions include the exclusion of nursing homes that do not file cost reports because these tend to be specialized facilities (e.g., specializing in subacute care or mental disease), and the exclusion of a relatively small number of residents who were either comatose or on hospice upon admission. A more detailed account of sample restrictions can be found in Appendix B.

I choose this particular time window because it allows me to study the effect of the introduction of CMS five-star ratings at the end of 2008.
2.4 Summary Statistics

Panel A of Table 1 presents summary statistics for the residents (at admission) in my sample. We observe that the average resident age is 77 years old, the majority of residents are white, and only a minority have a Bachelor’s or graduate degree. Most residents are admitted from an acute care hospital, and reflecting their poor state of health in general, almost a quarter of residents have dementia at admission, and 7.5 percent die within 90 days of admission. Appendix Table A.3a contains more detailed summary statistics for nursing home residents.

Panel B summarizes nursing home statistics, showing that nursing homes have 120 beds on average, occupancy rates tend to be quite high (86 percent on average), and most are owned by chains or are for profit. In addition, I have data on the number of deficiencies that nursing homes are cited for as well as self-reported staffing levels. Reflecting the prevailing wisdom that many nursing homes are understaffed, registered nurses (RNs) provide less than 25 minutes of care for each resident per day on average.

Table 1: Summary Statistics for Residents and Nursing Homes in California (2000–2010)

<table>
<thead>
<tr>
<th>Panel A: Residents (N=653,946)</th>
<th>Panel B: Nursing Homes (J=840)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 77.541 (12.991)</td>
<td>Number of Beds 123.989 (51.812)</td>
</tr>
<tr>
<td>Race: White 0.735 (0.441)</td>
<td>Occupancy Rate 86.198 (10.171)</td>
</tr>
<tr>
<td>Bachelor’s/Graduate Degree 0.131 (0.338)</td>
<td>Chain 0.641 (0.480)</td>
</tr>
<tr>
<td>Admitted from Acute Care Hospital 0.889 (0.314)</td>
<td>For-Profit 0.913 (0.282)</td>
</tr>
<tr>
<td>Has Dementia 0.233 (0.423)</td>
<td>Deficiencies 6.573 (8.156)</td>
</tr>
<tr>
<td>Death Within 90 days of Admission 0.075 (0.264)</td>
<td>RN hours per resident day 0.387 (0.354)</td>
</tr>
</tbody>
</table>

Notes: This table contains summary statistics for residents and nursing homes in California between 2000 and 2010. The unit of observation for nursing homes’ summary statistics is a nursing home-year, and each observation is weighted by the number of residents admitted to the nursing home for their first stay during that year.

40The MDS also records the payer source for residents at admission; although, as I discuss in Appendix B, this is less accurate than claims data.
41These summary statistics are based on OSCAR surveys, which are conducted annually. In addition, these statistics are weighted by number of admissions, although the general patterns for statistics that are not weighted by number of admissions are qualitatively similar (see Appendix Table A.4).
42A histogram of nursing home occupancy rates is also shown in Appendix Figure A.1.
3 Estimation of Nursing Home Quality

In this section, I estimate quality of nursing homes in California and validate these estimates using a distance-based IV strategy. I then briefly discuss various properties of my quality measure, such as its relationship with other potential quality measures, and the degree of geographical concentration in quality.

3.1 Framework for Quality Estimation and Validation

I follow a standard additive causal model from the value-added literature (described in greater detail in Appendix Section C). This model yields the following causal equation:

\[ Y_i = \mu_1 + \sum_{j=2}^{J} \beta_j D_{ij} + X_i' \gamma + u_i, \text{ } E[X_i u_i] = 0, \]

where \( Y_i \) is a dummy for whether resident \( i \) survives at least 90 days after admission, \( \beta_j \) is the causal effect of nursing home \( j \) on survival, \( D_{ij} \) is a dummy for whether resident \( i \) chooses nursing home \( j \), \( X_i \) is a vector of resident characteristics, and \( u_i \) is an unobserved health shock. To account for statistical noise in the estimation procedure, I estimate equation (1) using empirical Bayes (see Appendix K for more details about my implementation of empirical Bayes).

The literature on value-added estimation in education has often found that controlling for lagged values of the outcome variable is important (Chetty, Friedman, and Rockoff 2014). While controlling for lagged outcomes is impossible in the present setting with survival being the outcome variable, I control for residents’ baseline mortality risk based on a rich set of more than 500 health and demographic variables for residents at admission (even without interaction terms).

However, using the full set of controls for quality estimation runs the risk of overfitting, especially because some of the health measures correspond to relatively rare medical conditions, and the sample sizes for nursing homes are not huge. Hence, I use a modification of the post-double-selection method.

\[ Y_i = \mu_1 + \sum_{j=2}^{J} \beta_j D_{ij} + X_i' \gamma + u_i, \text{ } E[X_i u_i] = 0, \]

43 There is also a somewhat philosophical debate over whether the OLS or empirical Bayes estimates are preferable. In particular, if we are interested in learning about the quality of a specific nursing home (say nursing home \( j \)), then minimizing the conditional mean-squared error seems more natural, and we would prefer the OLS estimate. However, in the present setting, I am less interested in the quality of any particular nursing home, and more interested in obtaining reasonable estimates of quality for all nursing homes on average. In this case, minimizing the unconditional mean-squared error makes more sense because it measures how well the quality estimates predict outcomes on average, and so I use the empirical Bayes estimates instead. Moreover, given that my demand estimation results show that average demand for quality is very low, my use of the shrunken quality estimates is conservative, in the sense that using unshrunken quality estimates will likely result in even smaller demand estimates.

44 Alternatively, value-added estimation in education sometimes also uses test score gains as the dependent variable.

45 An extreme example of this would be if some of the controls (that one does not need to control for to obtain consistent estimates of quality) end up being perfectly collinear with some of the nursing home choice dummies, in which case we would be unable to estimate quality for these nursing homes.
in Belloni, Chernozhukov, and Hansen (2014) to select an appropriate set of controls for my quality estimation (see Appendix J for details).

Issues of overfitting and statistical noise aside, and despite the rich set of controls, one may still be concerned about selection on unobservables, which could be due to resident sorting or selective admissions by nursing homes (based on unobserved health \(u_i\)). To check whether my quality estimates \(\{\alpha_j\}\) are likely to be affected by selection bias, I use a validation test from the value-added literature based on the following idea: if the quality estimates are valid, then exogenous variation in quality should predict resident outcomes one-for-one. So, to obtain exogenous variation in the quality of residents’ chosen nursing home, I leverage residents’ preferences for closer nursing homes.

To elaborate, I derive the following structural equation from equation (1):

\[
Y_i = \mu_1 + \lambda \alpha_i \sim t(i) + X'_i \gamma + \tilde{u}_i, \tag{2}
\]

where \(\lambda\) is known as the forecast coefficient, \(\alpha_i \sim t(i) \equiv \sum_{j=1}^{J} \alpha_j \sim t(i) D_{ij}\) is the leave-year-out estimate of quality for the nursing home that resident \(i\) chooses, and \(\tilde{u}_i \equiv u_i + \eta_i\) is a composite structural error term, with \(\eta_i \equiv \sum_{j=1}^{J} \eta_j D_{ij}\) (see Appendix Section C for the derivation of the structural equation and the definition of the “forecast residual” \(\eta_i\)). I estimate equation (2) using IV, instrumenting \(\alpha_i \sim t(i)\) with the leave-year-out quality estimate of nursing homes close to resident \(i\)’s prior address, \(Z_i\), and test whether \(\lambda = 1\). If \(\lambda\) is indeed equal to one, then this is consistent with the quality estimates \(\alpha_j\) being asymptotically unbiased on average.

At a high level, the post-double-selection method uses lasso to choose the appropriate subset of controls for the estimation of treatment effects, in settings with many controls. Because omitted variables bias is determined by the product of the omitted variable’s covariance with the dependent variable, and with the treatment variable, one runs two lasso regressions of the dependent variable and the treatment variable on the controls. To make this procedure more robust, one takes the union of the two sets of controls selected by lasso in these lasso regressions, and estimates the treatment effect controlling for these variables.

A modification of the standard post-double-selection method is required for my setting, given that the standard method applies to the case with a single (or at most a few) treatment variable(s), whereas treatment in my setting corresponds to \(J = 1 > 800\) nursing home choice dummies. To address this, I use a model reduction strategy motivated by a correlated effects approach, which then allows me to apply the standard post-double-selection method.

It turns out that the variable selection procedure does not matter for my quality estimates. In particular, a regression of my main quality estimates (controlling only for the variables selected by the variable selection method mentioned above) on quality estimates with the full set of controls has an R-squared of 0.99.

I use leave-year-out quality estimates to avoid a mechanical relationship between resident outcomes and the quality estimate of residents’ chosen nursing home. Specifically, the outcome \(Y_i\) is used in the estimation of nursing home quality \(\alpha_j\), so using the quality estimates \(\alpha_i \equiv \sum_{j=1}^{J} \alpha_j D_{ij}\) as the endogenous variable will result in a mechanical relationship. Due to the leave-year-out definition of quality, we can also interpret equation (2) as an out-of-sample test of the predictivity of my quality estimates.

I define the quality \(\alpha_j\) of the first nursing home \((j = 1)\) to be zero since the notation of this section treats \(j = 1\) as the “omitted category.”

Strictly speaking, empirical Bayes estimates are not unbiased in finite samples, given that empirical Bayes minimizes the mean-squared error of the quality estimates according to the bias-variance tradeoff (which typically does not involve setting the bias to zero). So, I consider asymptotics wherein the number of residents in each nursing home tends to infinity, in which case the empirical Bayes and OLS estimates will eventually coincide.
For $Z_i$ to be a valid instrument several assumptions must be satisfied.

**IV Assumption 1 (First Stage).** The instrument must be relevant, i.e., we must have $\delta_Z \neq 0$ in the regression equation for the first stage:

$$\alpha_{i}^{\sim t(i)} = \delta_0 + Z_i'\delta_Z + X_i'\delta_X + e_i, \quad \mathbb{E}[(Z_i',X_i')'e_i] = 0. \quad (3)$$

**IV Assumption 2 (Exclusion Restriction).** The instrument must be uncorrelated with unobserved health shocks after accounting for resident characteristics, i.e.,

$$\text{Cov}(\tilde{Z}_i, u_i) = 0, \quad (4)$$

where $\tilde{Z}_i$ denotes the instrument $Z_i$ residualized of resident characteristics $X_i$.

**IV Assumption 3 (Fallback Condition).** The instrument must be uncorrelated with the forecast residual after accounting for resident characteristics, i.e.,

$$\text{Cov}(\tilde{Z}_i, \eta_i) = 0.$$

The first stage assumption can be easily tested by estimating equation (3) using OLS. As for the exclusion restriction, one way to interpret this condition is that the quality of nearby nursing homes can only affect a resident’s outcome through the quality of the nursing home she ultimately chooses (after controlling for her characteristics). This may be violated, for instance, if residents choose where they live based on nursing home quality and preferences varied by unobserved health status $u_i$, or if high-quality nursing homes choose to locate where unobservably healthier residents tend to live.

To account for sorting along these lines, I focus only on local variation in distance to nursing homes, by including county fixed effects in all IV specifications. Precise sorting by residents at the local level is relatively unlikely in this setting, given that housing location decisions are often made decades before residents require nursing home care, and rates of migration among the elderly are low (US Census Bureau 2003). Finally, I also conduct a balance test to check whether the instrument is correlated with observable determinants of health.

I do not focus much on the fallback condition (Abaluck, Bravo, Hull, and Starc 2021), since it is difficult to interpret, and the case wherein quality estimates are unbiased on average but the fallback condition fails is a knife-edge one (Chetty, Friedman, and Rockoff 2014). Moreover, because

---

52In particular, $\eta_i$ is not a structural parameter but rather arises from a complicated statistical relationship between the causal parameters $\beta_j$ and the quality estimates $\alpha_j$.

---

13
my instrument varies only at the geographical level, the inclusion of county fixed effects makes it less likely that the fallback condition is violated (similar to the exclusion restriction).

### 3.2 Quality Estimation Results

The standard deviation of my quality estimates is 0.02, which implies that a resident who goes to a nursing home with one standard deviation higher quality is 2 percent less likely to die within 90 days (all else equal). This is a 27 percent reduction in 90-day mortality compared to the baseline mortality rate of 7.5 percent, suggesting that nursing home choice can have a quantitatively meaningful impact on mortality even in the short run. In addition, Appendix Figure [A.2] shows that the distribution of my quality estimates is roughly bell-shaped, with a notable left tail of low-quality nursing homes.

Next, I validate these quality estimates by estimating the forecast coefficient using IV. Figure [1] provides support for the IV’s first stage and exclusion restriction. There is a clear positive relationship between the instrument and the endogenous variable, which provides strong support for the first stage assumption.\(^{53}\) By contrast, the lack of a clear relationship between the instrument and survival probability (which I use as a proxy for baseline health) lends confidence to the exclusion restriction.\(^{54,55}\)

I provide a visualization of the main IV result in Figure [2] by plotting the second stage of the IV. If the forecast coefficient is one, the best fit line should coincide with the 45-degree line, and this is indeed what we observe for my main quality estimates (shown in green). This figure also illustrates the importance of the rich set of controls in the MDS data: repeating the same validation exercise for quality estimates using no controls or only demographic controls, the best fit lines (in blue and red) are substantially flatter than the 45-degree line, indicating that these alternative quality estimates are likely to be biased.

Appendix Table [A.5] shows IV estimates of the forecast coefficient from different specifications, using the quality of the \(K\) nearest nursing homes to each resident as the instrument(s), for \(K = 1\) to 5. In all the specifications, the first stage F-statistics are relatively large, and estimates of the forecast coefficient are substantial.

\(^{53}\)Although I use the qualities of the \(K\) nearest nursing homes as separate instruments in the main IV specification, I use the average quality of the nearest \(K = 5\) nursing homes for Figure [2] so that it can be plotted more easily. Results from the same figure using only the quality of the nearest nursing home to the resident \((K = 1)\) is similar, albeit slightly noisier.

\(^{54}\)I compute the probability of surviving at least 90 days based on the quality estimation procedure, where I use the predicted values based on resident characteristics but not the nursing home effects.

\(^{55}\)The balance test in Figure [1] checks whether the residualized instrument is correlated with survival probability. There is a stronger version of this test that checks whether the unresidualized instrument is correlated with survival probability. Appendix Figure [A.3] shows the results of this test, revealing that there is no clear relationship between the instrument and survival probability as long as county fixed effects are accounted for. In particular, we observe that there seems to be a positive correlation between the instrument and predicted survival if we fail to include county fixed effects, suggesting that there may be sorting at broader geographical levels, and highlighting the importance of controlling for county fixed effects in my IV specifications.
Figure 1: First Stage Assumption and Exclusion Restriction

Notes: The x-axis in this figure is the average quality of the five nearest nursing homes to each resident. All variables are residualized of baseline resident characteristics and county fixed effects, other than the probability of survival (which is demeaned so that it can be plotted on the same scale).

In addition, the first stage and reduced form estimates for these models (shown in Appendix Figure A.4) demonstrate that the effects of nursing homes closer to a resident’s prior address on both the quality of her chosen nursing home and her outcome are larger, consistent with our intuition about this IV strategy.

A disadvantage of including residents from all payer sources in my sample is that there may be truncation bias due to the fact that I do not observe deaths after discharge for residents who are not covered by Medicare. On the other hand, even if I did observe deaths after discharge for all residents, it is unclear whether nursing homes are responsible for these deaths. Nonetheless, to gauge whether such truncation bias is likely to meaningfully affect my quality estimates, I estimate quality using survival for at least 19, 30, 60, and 180 days after admission as the outcome variables, and compare these

56 It makes sense that the point estimates may be smaller than one due to measurement error. In particular, if the measurement errors in the instrument and the endogenous variable are positively correlated, the IV estimates still suffer from a similar type of attenuation bias that OLS estimates do. See Appendix L for a more detailed explanation.

57 In Appendix O I provide a more general discussion about potential institutional reasons behind differences in reliability of value-added estimates in education and in the nursing home setting.
Figure 2: Second Stage of the Distance IV

Notes: The instruments are the qualities of the 5 nearest nursing homes to each resident. The blue, red, and green lines are the best fit lines for the quality estimates with no controls, only demographic controls, and the full set of controls, respectively. The 45-degree line is shown as a black dashed line. All variables are residualized of resident controls and county fixed effects.

I choose 19-day survival rate as one of the outcomes since Medicare switches from full to partial coverage on day 21 after admission, so nursing homes may have a financial incentive to discharge some residents around this cutoff if there are concerns over potential non-payment by the resident. Hence, there may be fewer sample selection issues when using 19-day survival rate, compared to survival rates over longer time horizons.
3.3 Discussion of Nursing Home Quality Estimates

3.3.1 Relationship Between Various Different Dimensions of Quality

Previously, I argued that different dimensions of quality are likely to be positively correlated. Here, I test whether this is the case by correlating my survival-based quality measure with other quality indicators. Table 2 shows that the correlations between my quality estimates and observable nursing home characteristics have the expected signs (Grabowski et al. 2016; You et al. 2016) – skilled staffing levels and CMS star ratings are positively correlated with my quality measure, whereas for-profit status and number of cited deficiencies are negatively correlated with quality. However, only a small fraction of the variation in my quality measure can be explained by these quality indicators, and I demonstrate in Appendix Table A.8 that my quality measure is a stronger (out-of-sample) predictor of resident outcomes than these quality indicators, suggesting that it contains information about nursing home quality beyond that which can be gleaned through publicly available metrics.

I also explore the relationship between my quality measure and quality estimates based on resident outcomes other than death, such as development of pressure sores, use of physical restraints, and antipsychotic use. Issues of misreporting and truncation bias for these other outcomes aside, Appendix Figure A.6 shows that my quality measure tends to be positively correlated with quality measures based on these other outcomes.

3.3.2 Local Variation in Nursing Home Quality

Finally, while the quality estimates indicate substantial variation in nursing home quality across California, the degree of local variation is also important for consumer choice; in particular, if quality is highly geographically concentrated, then residents living in certain areas may not have access to high-

59 The CMS introduced a five-star rating system for nursing homes in late 2008 in an effort to provide consumers with a simple metric with which to gauge nursing home quality. The introduction of star ratings comes towards the end of my sample period, but to the extent that relative nursing home quality remains roughly stable over time, the association between my quality measures and the star ratings remains informative of how predictive star ratings are of nursing home quality.

60 These correlations also have an implication for nursing home choice: if residents choose nursing homes based on observable nursing home characteristics that are predictive of quality rather than my quality estimates, they will also tend to choose higher-quality nursing homes according to my quality measure. I will revisit this point later when discussing the interpretation of my demand estimates for quality.

61 Correlations between the unshrunk en quality estimates with these nursing home characteristics are shown in Appendix Table A.7 and the results are very similar.

62 The regression of my quality estimates on all these quality indicators has an R-squared of only 0.04.

63 Specifically, Appendix Table A.8 shows results from regressions of survival on leave-year-out quality estimates and nursing home characteristics, wherein I standardize all right-hand-side variables to facilitate comparisons and multiply the coefficients by 100 for better legibility. We observe that the coefficient on the leave-year-out quality estimate in these regressions is highly statistically significant and an order of magnitude larger than the coefficients on the other quality indicators (which are mostly statistically insignificant). Moreover, both the coefficient on my quality measure and the R-squared remain essentially unchanged whether or not I include the other quality indicators.
Table 2: Relationship Between Quality and Nursing Home Characteristics

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009 Star Ratings (s.d.)</td>
<td>0.0783**</td>
<td>0.0518</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0310)</td>
<td>(0.0316)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RN hours per resident day (s.d.)</td>
<td>0.0921***</td>
<td>0.197***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0328)</td>
<td>(0.0461)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPN hours per resident day (s.d.)</td>
<td>0.0977***</td>
<td>0.103***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0205)</td>
<td>(0.0316)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNA hours per resident day (s.d.)</td>
<td>0.0900***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0159)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deficiencies (s.d.)</td>
<td></td>
<td>-0.0457***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0129)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>For-Profit (s.d.)</td>
<td></td>
<td>-0.0846***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0283)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chain (s.d.)</td>
<td></td>
<td>-0.0322</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0277)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>10,103</td>
<td>10,118</td>
<td>10,119</td>
<td>10,112</td>
<td>10,121</td>
<td>10,121</td>
<td>10,121</td>
<td>10,089</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.009</td>
<td>0.013</td>
<td>0.014</td>
<td>0.012</td>
<td>0.003</td>
<td>0.011</td>
<td>0.002</td>
<td>0.040</td>
</tr>
</tbody>
</table>

Notes: This table shows correlations between the nursing home quality estimates and various nursing home characteristics. The unit of observation is a nursing home-year. Observations are weighted such that the total weight each nursing home receives is equal to the number of long-stay residents it has over the sample period. Standard errors are clustered by nursing home.

quality nursing homes. To shed light on this, Figure 3 plots empirical cumulative distribution functions (ECDFs) showing the overall variation in nursing home quality as well as the within- and across-county variations in quality.64 The figure shows that there is almost as much variation in nursing home quality within counties as there is overall,65 implying that there tends to be both high-quality and low-quality nursing homes close to each resident. Whether residents take advantage of this variation is the main topic of the next section.

4 Resident Preferences for Nursing Homes

In this section, I use my survival-based quality estimates to study residents’ demand for quality. I start in section 4.1 by showing evidence that there are unobserved constraints on residents’ choice sets due to selective admissions practices by nursing homes (Gandhi, 2019). In section 4.2, I describe how I address these unobserved choice set constraints by outlining a structural demand model and estimation procedure based on recent advances in the empirical matching literature (Agarwal and Somaini, 2022). Finally, in section 4.3, I present results from the structural demand estimation, and

64 To plot the within-county quality variation, I use the residuals from a regression of my quality estimates on county fixed effects. To plot across-county quality variation, I use average quality within each county.
65 Moreover, within-county variation in quality accounts for far more of the overall variation than across-county variation in quality.
Figure 3: ECDF of Quality Estimates (Overall, Within Countries, and Between Counties)

Notes: This figure plots the empirical CDFs showing the overall variation in nursing home quality, as well as the within-county, and across-county variations in quality, in blue, red, and green respectively. Within-county variation is plotted using the residuals from a regression of the quality estimates on county fixed effects, across-county variation is plotted using the averages of the quality estimates within each county. The p-values from two-sample Kolmogorov-Smirnov tests for equality of distributions comparing the overall distribution of quality, to the within-county and across-county distributions of quality are also shown in the figure.

4.1 Selective Admissions by Nursing Homes

Gandhi (2019) showed that nursing homes may have a financial incentive to reject certain types of residents, even if the resident is profitable and the nursing home still has spare beds. This is due to the option value of waiting for the arrival of a more profitable resident in the future: if the nursing home accepts a resident today, it is more likely that it will not have spare capacity if a more profitable resident arrives in the near future. This also implies that when nursing homes are closer to capacity, the option value is higher, and nursing homes will be more selective about the types of residents they
admit.

If nursing homes engage in these selective admissions practices, it will impose unobserved constraints in residents’ choice sets. This poses a challenge for demand estimation, given that ignoring these constraints may lead to spurious demand estimates. Hence, before discussing my structural demand estimation, I first present evidence that nursing homes do indeed practice selection admissions, by testing two predictions from Gandhi’s (2019) model.

**Prediction 1.** When occupancy within a nursing home is higher than usual, the nursing home is less likely to admit new residents.

**Prediction 2.** Characteristics of the typical resident who is admitted to a nursing home when occupancy is high tend to differ from characteristics of the typical resident admitted when occupancy is low.

Table 3 shows results supporting prediction 1: conditional on nursing home-month fixed effects, nursing homes admit fewer new residents on days when occupancy is higher than usual (and this finding is not sensitive to the precise definition of occupancy used). Figure 4 shows support for prediction 2: when nursing homes are closer to capacity, they are less likely to admit Medicaid residents (whether or not we control for other resident characteristics that may be correlated with Medicaid status), consistent with the low reimbursement rates associated with Medicaid. Appendix Section M contains a more detailed description of the tests for predictions 1 and 2, as well as additional supporting evidence. In summary, these tests reach the same conclusion as Gandhi (2019): nursing homes’ behaviors are consistent with selective admissions.

---

66I focus only on within-nursing home month variation in occupancy so that the correlation between admissions and occupancy is driven by short-run capacity management incentives, rather than by nursing home expansions and contractions (which would typically work in the opposite direction, given that expanding nursing homes will typically admit more residents). Although the total number of beds in nursing homes is reported in the OSCAR data, this is updated only annually, and there is also measurement error in this variable.

67Strictly speaking, prediction 1 only tests whether capacity constraints are relevant (i.e., if nursing homes do not account for the option value of leaving beds open and always admit residents unless they are completely full, prediction 1 will still be satisfied), and thus it is a weaker test of selective admissions than prediction 2 (which will only be fulfilled if nursing homes reject certain types of residents when close to being full). Nonetheless, prediction 1 still implies that there are choice set constraints, which is relevant for demand estimation, and moreover, unless we have perfect data on nursing home capacity (which we do not), even in the absence of selective admissions, choice set constraints due to limited capacity will still be unobserved (or at most, observed with error).

68Appendix Table A.9 shows that this result is robust to alternative measures of new admissions, specifically, a dummy for any new admissions, and the flow of residents (i.e., admissions minus discharges).

69Appendix Table A.10 and Appendix Figure A.2 provide additional evidence in favor of prediction 2.

70It is worth noting that these tests do not necessarily require nursing homes to be profit-maximizing entities: in particular, they allow for selective admissions due to non-financial motives. For example, a purely altruistic nursing home that is close to capacity may give higher priority to residents whom it perceives to be in greater need of care.
Table 3: Effect of Occupancy on Admissions

<table>
<thead>
<tr>
<th></th>
<th>Number of New Residents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Lagged 7-Day Avg. Log Occupancy</td>
<td>-0.480***</td>
</tr>
<tr>
<td></td>
<td>(0.0616)</td>
</tr>
<tr>
<td>Lagged 7-Day Avg. Occupancy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged 7-Day Avg. Occ. Percentile</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Nursing Home-Month Fixed Effects</td>
<td>X</td>
</tr>
<tr>
<td>N</td>
<td>1,103,528</td>
</tr>
</tbody>
</table>

Notes: This table shows regression results at the nursing home-day level wherein the dependent variable is number of new patients, and the independent variables are various measures of nursing home occupancy. Standard errors are clustered at the nursing home level.

Figure 4: Bin Scatters of Medicaid Against Occupancy

(a) Without Controlling for Other Characteristics

(b) Controlling for Other Characteristics

Notes: The unit of observation is a resident, and the occupancy measure is the lagged seven-day average log occupancy for the nursing home as of the date of admission for the resident, residualized of nursing home-month fixed effects. Nursing home fixed effects are included in the bin scatters and regressions.

4.2 Structural Model for Demand Estimation

In this subsection, I describe a structural demand model and estimation procedure that accounts for the unobserved choice set constraints arising from the selective admissions practices documented in the previous subsection.
4.2.1 Identification of Two-Sided Matching Model

Given that nursing homes have rankings over residents and vice versa, the setting is well-approximated by a many-to-one two-sided matching market. Each resident $i \in I$ with characteristics $w_i$ is matched to exactly one nursing home, whereas a nursing home $j \in J$ can be matched with more than one resident. I model this using a random utility model, with residents’ (decision) utility $v_{ij}$ and nursing homes’ profits $\pi_{ij}$ given by:

$$v_{ij} = v_1^j(w_i, \zeta_i) - v_2^j(w_i, dist_{ij}),$$
$$\pi_{ij} = \pi_j(w_i, \zeta_i, occ_{ij}),$$

(5)

where $dist_{ij}$ is the distance between resident $i$ and nursing home $j$, and $occ_{ij}$ is a measure of short-term fluctuations in nursing home $j$’s occupancy when resident $i$ is choosing her nursing home.

Agarwal and Somaini (2022) establish a sharp set of conditions under which these preferences are non-parametrically identified. The key substantive requirement for identification is the existence of demand and supply instruments,\textsuperscript{71} which in my case are $dist_{ij}$ and $occ_{ij}$, respectively.

Similar to standard IV, the validity of these instruments relies on the relevance (i.e., first stage) condition and exclusion restriction, which I state below.

**Assumption D1 (Relevance).** The demand and supply instruments must be relevant (i.e., $\partial v_{ij}/\partial dist_{ij} \neq 0$, and $\partial \pi_{ij}/\partial occ_{ij} \neq 0$).

**Assumption D2 (Exclusion Restriction).** The demand instrument must be excluded from the supply side ($\partial v_{ij}/\partial occ_{ij} = 0$), and the supply instrument must be excluded from the demand side ($\partial \pi_{ij}/\partial dist_{ij} = 0$).

Relevance of the supply instrument $occ_{ij}$ is supported by earlier results in Table 3 and similarly, relevance of the demand instrument $dist_{ij}$ is demonstrated in the first stage of the distance IV.\textsuperscript{72} The exclusion restriction for distance as the demand instrument is also quite intuitive: there is little reason for nursing homes to care where their residents lived prior to admission.

The assumption that merits most discussion is the exclusion restriction for temporary occupancy fluctuations as the supply instrument. Let us first consider why simply using occupancy as the supply

\textsuperscript{71}See Appendix E for the full list of the technical assumptions required for identification, as well as a discussion about how they relate to my setting.

\textsuperscript{72}I also verify that choice and distance are negatively correlated. Specifically, for each resident-nursing home pair, I regress a dummy for nursing home choice on distance. The estimated coefficient on distance is negative and highly statistically significant whether or not I include nursing-home fixed effects (with t-statistics greater than 10).
instrument is likely to violate the exclusion restriction, and how this will affect our demand estimate. Suppose that high-quality nursing homes tend to have higher occupancy rates on average but that, all else equal, residents prefer nursing homes with lower occupancy rates (which violates the exclusion restriction). This will bias our demand estimate downwards because it conflates residents’ preferences for quality and lower occupancy rates.

To avoid such a bias, I use only within-nursing home-month variation for my occupancy measure. This is important for two reasons. First, it is more plausible that demand is insensitive to short-term fluctuations in occupancy rates. Second, this definition of the supply instrument ensures that it does not vary systematically with nursing home quality. Indeed, Appendix Figure A.8 shows that the distributions of $occ_{ij}$ at above-median and below-median quality nursing homes are essentially identical.

### 4.2.2 Estimation Framework

Although preferences of residents and nursing homes are non-parametrically identified, non-parametric estimation is likely to have slow rates of convergence. Hence, I consider the following parametrization of residents’ and nursing homes’ preferences:

$$
v_{ij} = q'_j \kappa^1 + q'_j \kappa^2 w_i + dist'_{ij} \kappa^{dist} + \epsilon_{ij},$$

$$\pi_{ij} = w'_i \psi^1 + q'_j \psi^2 w_i + occ'_{ij} \psi^{occ} + \omega_{ij},$$

(6)

where $\epsilon_{ij}$ and $\omega_{ij}$ follow independent Gaussian distributions. I impose the location normalization for residents’ utility by setting the utility of a nursing home and the intercept to zero, and the location normalization for nursing homes’ admission rules by assuming that nursing home $j$ is willing to accept resident $i$ if and only if $\pi_{ij} \geq 0$. To set the scale normalization, I set the variances of $\epsilon_{ij}$ and $\omega_{ij}$ to one.

Estimation must still address the curse of dimensionality, considering that the number of possible eligible choice sets for each resident grows exponentially with the potential number of nursing homes in her choice set. This makes standard methods such as maximum likelihood computationally infeasible.

---

73 Focusing on within-nursing home-month variation as opposed to simply within-nursing home variation not only ensures that I use only fluctuations relative to the nursing home’s typical occupancy levels, but also accounts for changes in nursing home capacity over time (e.g., if nursing homes expand or contract).

74 Note that all residents in the sample choose a nursing home by construction; by definition, none of the residents in my sample choose the outside option of no nursing home.

75 Denoting the potential number of nursing homes in resident $i$’s choice set by $J_i$, there are $2^{J_i} - 1$ distinct possibilities for the set of nursing homes that are willing to admit her.
sible,76 unless one is willing to make further assumptions to rule out some possibilities,77 or restrict the number of nursing homes in each resident’s choice set.78 To address the curse of dimensionality, I estimate the model using a Gibbs sampler with data augmentation on the latent variables $v_{ij}$ and $\pi_{ij}$, a method that provides a convenient dimension reduction without requiring further substantive assumptions (Agarwal and Somaini 2022).

At a high level, the Gibbs sampling method involves iteratively drawing the structural error terms $(\epsilon', \omega')$ and the parameters $\theta \equiv (\kappa', \psi')'$,79 in a way that respects the matching outcomes.80 Under standard conditions, the draws of $\theta$ will eventually converge to the stationary distribution,81 and I conduct inference using the draws of $\theta$ after the point at which the chain seems to have reached the stationary distribution.82,83 Appendix F describes the full algorithm for the Gibbs sampler in detail.

4.3 Demand Estimation Results

4.3.1 Demand for Quality and Information Frictions

Table 4 presents demand side estimates from the structural demand model described in the previous subsection. Given that nursing homes’ admission decisions are less relevant to the main discussion, supply side estimates are displayed in Appendix Table A.12.84

---

76For my demand estimation sample, the (potential) number of nursing homes in each resident’s choice set is 50 on average (so there are $2^{50} - 1 > 10^{15}$ possible eligible choice sets), and can be larger than 200 (in which case there are $2^{200} - 1 > 10^{60}$ possible eligible choice sets). Hence, any method that requires directly integrating over all $2^{2i} - 1$ possibilities for each resident is computationally infeasible.

77For example, Gandhi (2019) assumes that there are no idiosyncratic differences between nursing homes’ admission rules for a given resident (i.e., he assumes that $\omega_{ij} = \omega_i$ for all $j$ in the notation of our model). This limits the number of possible eligible choice sets for each resident to $2^{2i}$ (down from $2^{2i}$).

78I use nursing homes within 15 miles of each resident as her choice set. Choosing a smaller radius will reduce the computational burden, but at the expense of excluding residents who choose nursing homes further away, and may thus result end up with an unrepresentative sample. The CDF of residents’ distances to their chosen nursing homes, shown in Appendix Figure A.29 indicates that more than 80 percent of residents choose a nursing home within 15 miles, less than 80 percent choose a nursing home within 10 miles, and less than 60 percent choose a nursing home within 5 miles, so using a smaller radius to determine residents’ choice sets risks losing a substantial portion of the sample.

79This procedure follows a Bayesian approach, so the parameters $\theta \equiv (\kappa', \psi')'$ are treated as random, with some prior distribution. The posterior distribution of $\theta$ is updated in each iteration.

80Thus, for the nursing home $\mu(i)$ that resident $i$ is admitted to, during each iteration of the Gibbs sampler she must be eligible for it, and she must also prefer it to all other nursing homes that she is eligible for (i.e., $\pi_{i,\mu(i)} \geq 0$, and $v_{i,\mu(i)} \geq v_{ij}$ for all nursing home $j$ for which $\pi_{ij} \geq 0$). Similarly, if resident $i$ is not admitted to nursing home $j$, then in each iteration we must ensure that either she is not eligible for nursing home $j$, or that she prefers her chosen nursing home $\mu(i)$ to nursing home $j$ (i.e. either $\pi_{ij} < 0$ or $v_{ij} < v_{i,\mu(i)}$).

81Specifically, the transition kernel for the Gibbs sampler has to be irreducible and aperiodic. More primitive conditions have also been derived that ensure the transition kernel for the Gibbs sampler has these properties (Robert and Smith 1994).

82In particular, I visually inspect the Markov chain, check whether the potential scale reduction factor (Brooks and Gelman 1998) is close to one, and consider the effective sample size (Vats, Flegal, and Jones 2019).

83Although the Gibbs sampler is derived following a Bayesian approach, the Bernstein-von Mises theorem implies that the distribution of $\theta$ will converge to the distribution of the maximum likelihood estimator under correct specification. Therefore, our Bayesian inferences based on the results of the Gibbs sampler are also (asymptotically) correct from the frequentist point of view.

84The supply side estimates generally indicate that nursing homes are more likely to admit residents from acute care hospitals (who are often covered by Medicare) and less likely to admit black and hispanic residents. However, the supply
Column 1 contains one of the key results of this paper: the estimate of residents' demand for quality is quantitatively very small. In particular, the demand coefficients on quality and distance imply that the marginal rate of substitution (MRS) between quality (in percentage points) and distance is:

\[ MRS = \frac{d(\text{dist}_{ij})}{d(\text{quality}_j)} = -\frac{\partial v_{ij}}{\partial \text{quality}_j} \cdot 0.01 \approx -\frac{0.00564}{-0.16} \approx 0.035. \]

This MRS estimate implies that in exchange for a one-mile reduction in travel distance, residents are willing to tolerate a \( \frac{1}{0.035} = 28.6 \) percent higher probability of dying within 90 days (more than 14 times the standard deviation in quality, which is 2 percent). Alternatively, if we convert the mortality effects into dollar terms, the MRS implies that residents are willing to give up more than $100,000 in exchange for traveling one less mile.\(^{85}\) Moreover, the MRS estimate of 0.035 is also an order of magnitude smaller than MRS estimates in previous studies on demand for quality in the hospital setting, which range from 1.8 to 8 (Tay 2003; Chandra, Finkelstein, Sacarny, and Syverson 2016).\(^{86,87}\)

The coefficients on the demand and supply side instruments (distance to nursing home and occupancy fluctuations) are highly statistically significant and stable across specifications, so the low demand estimate is unlikely to be due to weak identification. In addition, Appendix Table A.13 shows that the demand estimate remains low whether we use alternative specifications for the instruments (panel A), or different sample restrictions (panel B).

Moreover, this finding of low demand for quality is not limited to my survival-based quality measure:

\(^{85}\)We can convert this to dollar terms using a back-of-the-envelope calculation under a few assumptions. The units for my quality measure are in terms of 90-day survival rate, which cannot be translated directly into life expectancy (and it is also difficult to study life expectancy with my data because I cannot track most residents after they are discharged from their nursing home). Nonetheless, suppose we are willing to assume that a resident who survives at least 90 days after admission goes on to live one year on average. This is a somewhat conservative assumption considering that most residents are admitted for rehabilitation and are expected to make a full recovery; for example, a study found that life expectancy after a hip replacement surgery (a common reason for which residents are admitted to nursing homes) is 8.2, 4.8, and 2.8 years for female patients 70, 80, and 90 years old, respectively (Liu et al. 2021). Taking $400,000 as the value of a statistical life year, the MRS implies that residents are willing to forgo more than $100,000 in exchange for a one-mile reduction in distance to her nursing home on average, given that:

\[ ($400,000/\text{life-year}) \times (0.286 \text{ life-years/mile}) = $114,000/\text{mile}, \]

which is very large (especially considering that most residents only travel to their nursing home once).

\(^{86}\)The MRS estimates from these other studies imply that residents are willing to tolerate roughly a 1/8 = 0.125 to 1/1.8 \( \approx 0.556 \) percent higher probability of death in exchange for a one-mile reduction in travel distance.

\(^{87}\)Although a low MRS estimate in the nursing home setting may in theory reflect either low demand for quality or a strong preference for distance, the latter seems less likely because residents typically only need to travel to the nursing home once. Distance may be more relevant for visiting family members, but this is unlikely to explain the low MRS estimate unless their preferences for distance are an order of magnitude greater than in settings considered in the prior literature.
Table 4: Structural Demand Estimates

<table>
<thead>
<tr>
<th>Resident Preferences</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance to Nursing Home (in miles)</td>
<td>-0.159***</td>
<td>-0.161***</td>
<td>-0.16***</td>
<td>-0.159***</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.008)</td>
<td>(0.011)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Quality</td>
<td>0.564*</td>
<td>0.257*</td>
<td>5.678***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.326)</td>
<td>(0.145)</td>
<td>(0.755)</td>
<td></td>
</tr>
<tr>
<td>Deficiencies</td>
<td>0.008***</td>
<td>0.008***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>For-Profit Nursing Home</td>
<td>0.128*</td>
<td>0.183*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.134)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RN Hours Per Resident-Day</td>
<td>0.294***</td>
<td>0.283***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.032)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPN Hours per Resident-Day</td>
<td>0.241***</td>
<td>0.259***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.042)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNA Hours Per Resident-Day</td>
<td>-0.122***</td>
<td>-0.113***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.023)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chain</td>
<td>0.152***</td>
<td>0.165***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.031)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality x Dementia</td>
<td>-2.424***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality x Age</td>
<td>-0.059***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality x At Least Bachelor's Degree</td>
<td>1.326***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.299)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality x Post-Star Ratings</td>
<td>0.56***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.199)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality x Black</td>
<td>-1.369***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.513)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality x Hispanic</td>
<td>-2.325***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.528)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality x Female</td>
<td>-0.014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.22)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality x Married</td>
<td>-0.323*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.246)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality x Lived Alone</td>
<td>-0.046</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.245)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Supply Side

| Temporary Fluctuations in log(occupancy)                   | -7.34***  | -7.97***  | -6.95***  | -7.38***  |
|                                                           | (1.069)   | (1.24)    | (1.365)   | (1.1)     |

Resident Controls in Supply Side Equation X X X X

Notes: This table shows demand side estimates (as well as the estimated coefficient on the supply side instrument) from the structural estimation using Gibbs sampling. A burn-in period corresponding to the first half of the chain was used.

Residents also do not seem to consistently value publicly available nursing home quality measures. In particular, the estimates in columns 2 and 3 of Table 4 show that residents tend to choose nursing
homes that are cited for more deficiencies or are for-profit, characteristics that are associated with lower
nursing home quality (Grabowski et al. 2016; You et al. 2016). In addition, although estimates
of residents’ demand for skilled staffing are positive, they are quantitatively small and similar to the
estimate of residents’ demand for quality in column 1 if we convert staffing levels into dollar terms
based on wages. The fact that residents do not seem to take full advantage of publicly available information suggests
that information frictions may play a role. As further evidence of this interpretation, column 4 of Table
shows that residents with characteristics associated with less information frictions display a greater
demand for quality: the coefficients on the interactions between quality and resident characteristics
indicate that residents who are free of dementia, are less advanced in age, or have at least a Bachelor’s
degree display a greater demand for quality. Moreover, residents choosing nursing homes after an
information intervention by the CMS (specifically, the introduction of the five-star ratings system at
the end of 2008) are more sensitive to quality. Appendix Table A.14 shows that these demand
patterns are also robust to alternative specifications, such as allowing preferences to vary by the size
of residents’ choice sets, interacting the instruments with resident characteristics, or including nursing
home fixed effects in the utility equation. In addition, Appendix Table A.15 shows that older residents,

\[ \frac{0.16 \text{RN hours daily} \times 180,000/\text{one-hour increase in RN care daily for an average resident} \approx 102,000/\text{mile},}{0.28 \text{ miles}} \]

which is similar to the value of $114,000/mile calculated based on demand for survival-based quality.

There is a strong possibility that residents with cognitive impairments (e.g., dementia) do not choose their own
nursing homes and the choice is instead made by a surrogate (e.g., a family member or staff from the hospital that she
was discharged from). This will change the interpretation of the results on demand heterogeneity slightly, but to the
extent that the resident would have made worse choices if she chose on her own in these instances, my estimation results
will understimate the true association between information frictions (of residents) and demand heterogeneity (strengthening
my results). The identity of the decision maker has implications for the design of effective information interventions, but
this is not the primary objective of the present paper.

This echoes the finding of Chen (2018) that Yelp ratings for physicians have an effect on patients’ demand for them.

The magnitude of these differences in demand for quality are sizable relative to the average effect. For example,
the magnitude of the average difference in demand for quality (2.424) between residents with and without dementia
is four times the average demand for quality (0.564). Similarly, the magnitude of the average differences in demand
between residents who are one standard deviation apart in age (almost 13 years) between residents with and without
Bachelor’s degrees and residents choosing nursing homes before and after the introduction of star ratings are respectively,
12.991 × 0.059/0.564 ≈ 1.36, 1.326/0.564 ≈ 2.35, and 0.56/0.564 ≈ 0.99 times the average demand for quality.
residents with dementia, and residents choosing their nursing homes prior to the introduction of star ratings are also less sensitive to skilled staffing levels, data on which is publicly available.\footnote{The results for other predictors of quality such as deficiencies and for-profit status are more ambiguous. However, this is unsurprising given that average demand for these variables is positive even though these characteristics negatively predict quality, suggesting that most residents do not make effective use of these variables.}

4.3.2 Other Possible Explanations for Demand Patterns

Information frictions is not the only possible explanation for the low demand and patterns of demand heterogeneity I find. So, for the remainder of this section, I consider alternative explanations for these demand patterns, and provide evidence showing that they are less plausible explanations than information frictions.

First, there is a possibility that even if residents had a low demand for my survival-based quality measure, they are in fact highly sensitive to other dimensions of quality. However, the finding in columns 2 and 3 of Table 4 that residents also have a low demand for publicly observable quality measures suggests that strong preferences for other dimensions of quality are unlikely to explain the low demand estimate. Moreover, given my earlier finding in Table 2 and Appendix Figure A.6 that the survival-based quality measure is positively correlated with other quality measures, the demand estimate in column 1 may already partially capture residents’ demand for other dimensions of quality. Supporting this interpretation, we observe that the demand estimate becomes smaller in column 3 when I include publicly available nursing home quality measures in residents’ utility equation.

A second related possible explanation for the low demand is that it is due to omitted variables bias. However, the positive correlation between my survival-based quality measure and other quality measures suggests that if anything, the omitted variables bias is likely to be in the wrong direction — indeed, the demand estimate becomes smaller in column 3 when I include publicly available nursing home quality measures in residents’ utility equation. Moreover, even if we believe that there is an omitted variable that residents value but that is negatively correlated with quality, in terms of magnitude, residents have to value it many times more than they do observable quality measures, and I have ongoing work on selection on unobservables for discrete choice models that provides a way to quantify this. In particular, Appendix Figure A.10 shows that in order for omitted variables bias to completely explain the discrepancy between my demand estimate and previous estimates from the hospital setting, residents need to value the omitted variable more than 100 times more than they do observable quality measures, based on the type of assumption on the proportional selection relationship adopted by Oster.
A third possible explanation of the low demand estimate is that residents may not directly observe nursing home quality $\alpha_j$. A natural way to model this is to assume residents only receive a noisy signal of quality, which they combine with publicly available information about nursing homes $q_j$ to make an optimal forecast about quality $\alpha_j$.\(^{95}\) I illustrate with a simple model in Appendix Section N that if this is the case, residents should value characteristics in $q_j$ that are positively correlated with $\alpha_j$.\(^{96}\) However, Appendix Figure A.11 shows that residents tend to choose nursing homes with more cited deficiencies and for-profit nursing homes even though these characteristics are negatively correlated with $\alpha_j$, which provides evidence against this explanation.\(^{97}\) These patterns also provide evidence against an explanation for the low demand based on search costs (in terms of learning about nursing home quality) and “rational inattention,” as long as we assume that search costs for publicly available information is low.

Fourth, an econometric explanation for the low demand estimate is that it may the result of functional form misspecification, or attenuation bias due to the fact that quality was estimated. To test this possibility, I replace quality with nursing home fixed effects in residents’ utility equation, and plot these mean utilities against nursing home quality in Appendix Figure A.12. This scatterplot shows that there is hardly any relationship between residents’ preferences and nursing home quality (linear or otherwise), suggesting that functional form misspecification is unlikely to be the main explanation for the low demand estimate. Moreover, Abaluck et al. (2021) shows that even if there is estimation noise in quality, the slope of the best fit line in this figure provides an upper bound for demand, and we observe that the implied upper bound for the MRS remains substantially smaller than MRS estimates from the literature.

In addition to the arguments presented above, the first four explanations also do not explain why we observe patterns of demand heterogeneity in column 3 of Table 4 that are consistent with information frictions. So, I next explore other possible explanations for demand heterogeneity.

First, there is a possibility that nursing homes specialize in treating different types of residents.\(^{98}\)

\(^{95}\)For example, we can imagine the noisy quality signal as the resident’s (or her family’s) impression of the nursing home after visiting it, or based on word-of-mouth recommendations.

\(^{96}\)Allowing for estimation noise in my quality measure makes no difference to the model’s predictions.

\(^{97}\)Appendix Figure A.11 is similar in spirit to the graphical tests used by Card and DellaVigna (2020) in the context of editors’ decisions and subsequent paper citations. Specifically, I plot the choice coefficients on nursing home characteristics $q_j$ against the coefficients from a regression of my quality measure $\alpha_j$ on $q_j$. If residents made inferences about nursing home quality based on easily observable characteristics, all points in this graph should lie in the first and third quadrants. However, we observe that this is not the case: the points for cited deficiencies and for-profit status lie in the second quadrant. In other words, residents tend to choose nursing homes with more cited deficiencies and that are for-profit, even though these characteristics are negative predictors of quality.

\(^{98}\)For instance, it may be the case that some nursing homes specialize in providing care for long-stay residents, but
If this is the case, then my use of a single quality measure may be inappropriate, and the estimates of demand heterogeneity may simply be the result of residents reacting to nursing homes’ comparative advantages. To test this explanation, I partition my sample based on several resident characteristics and estimate quality separately for each subsample. Figure 5 shows that the quality estimates based on different subsamples are highly correlated, whether I split the sample by admission origin (which predicts whether a resident is short or long stay), age, dementia status, or education, suggesting that my use of a single quality measure for all residents is a reasonable approximation, and that specialization alone is unlikely to explain the patterns of demand heterogeneity that I find.99

Figure 5: Bin Scatters of Quality Estimates for Different Subpopulations

(a) Whether the Resident was Post-Acute Care
(b) Whether the Resident was Above 80
(c) Whether the Resident has Dementia
(d) Whether the Resident has Bachelor’s Degree

Notes: For each plot, I partition the resident population based on the resident characteristic in the sub-figure title, estimate quality separately for the two subsamples, and plot these two sets of coefficients against each other.

99While the analogy is not exact, this echoes the finding in Grabowski, Gruber, and Angelelli (2008) that quality is common across Medicaid and private-pay residents within a nursing home.
Second, there may be demand heterogeneity due to selection on gains: if quality matters less for the outcomes of certain types of residents and these residents are aware of this, they may “rationally” have a lower demand for quality. To probe this possibility, I check whether patterns of heterogeneity in the predictivity of quality for resident outcomes line up against demand heterogeneity.\footnote{Specifically, to estimate heterogeneity in the predictivity of quality for resident outcomes, I run OLS and IV regressions of resident outcomes on (leave-year-out) quality, and its interactions with several resident characteristics. For my IV specification, I instrument quality of the resident’s chosen nursing home with the quality of the closest nursing home to the resident, and similarly, I instrument the interactions between quality and resident characteristics with the interactions between the quality of the closest nursing home to the resident and these same characteristics.} Instead, the results in Appendix Table A.10 show a “reverse Roy” pattern (Walters, 2012): the estimates imply that quality matters more for the outcomes of older residents and residents with dementia, and yet, results from column 4 of Table 4 show that these residents also tend to have lower demand for quality (which is the opposite of what a model based on selection on gains would predict).

Finally, there is a possibility that the marginal utility associated with an additional life-year may be lower for older residents and residents with dementia, e.g., due to lower quality of life. It is possible to test this directly by calculating quality-adjusted mortality effects of different nursing homes for each resident. However, results from such an exercise may not be reliable, since there is substantial uncertainty around the weights used for quality adjustment (Gold, Stevenson, and Fryback 2002). So, instead of testing this explanation directly, we note that our earlier finding that quality matters more for the mortality outcomes of older residents and residents with dementia should mitigate this force. Moreover, in terms of magnitude, adjusting my survival-based quality estimates to account for quality of life is unlikely to explain the large gap between my demand estimate and existing demand estimates from the hospital setting.\footnote{For example, the disability weight for Alzheimer’s and other dementias is 0.666 (WHO 2013), so even if my entire sample had dementia, my demand estimate for quality after adjusting for dementia will only be 1.5 times greater, and still far smaller than demand estimates from the literature from the hospital setting (0.035 \times 1.5 = 0.0525, compared to estimates ranging from 1.8–8).}

Therefore, the evidence in this subsection points to information frictions being the most likely explanation for residents’ low demand for quality. Quantifying the costs of these information frictions is the topic of the next section.

5 Policy Counterfactuals

In this section, I quantify the costs of the information frictions documented in the previous section via counterfactual simulations. In section 5.1, I describe how I model the elimination of information frictions, and discuss the various treatment effect mechanisms at work. In section 5.2, I describe a
competing risks model for resident deaths and discharges, which is necessary for my counterfactual simulations. In section 5.3, I present my simulation results, which quantify both the short- and long-run effects of eliminating information frictions on resident mortality as well as the distributional consequences. Finally, in section 5.4, I discuss these results in the context of nursing home policies more generally.

5.1 Framework for Policy Counterfactuals

I use a broad definition of information frictions, which includes both frictions that arise from search and processing costs as well as those that are due to psychological distortions (Handel and Schwartzstein, 2018). We can also view this as the treatment effect of an information intervention that eliminates information frictions. I remain largely agnostic about the precise ways in which a policy intervention may be able to achieve this (be it through information provision, or aid in interpreting and processing this information) because this requires a detailed analysis of past policies that is beyond the scope of this paper.

5.1.1 Modeling Policy Counterfactuals

To study the costs of information frictions, I consider counterfactual preferences for residents that are consistent with estimates from the literature of the MRS between quality and distance. For expositional purposes, in the following discussion I refer to this as the perfect information benchmark, which implicitly assumes that consumers in these other healthcare settings do not face information frictions. If consumers in these studies do in fact face information frictions, then, unless these frictions cause them to overvalue healthcare quality (in all of the studies I consider), I am likely to understate the true costs of information frictions for nursing home residents. Alternatively, a more conservative interpretation of my counterfactuals is that they simulate the likely consequences if nursing home residents made choices more similar to those made by consumers in other healthcare settings.

In my short-run analysis, I simulate outcomes based on these counterfactual preferences, holding nursing home quality constant.\textsuperscript{102} This is likely to be a partial equilibrium because nursing homes may respond to changes in demand. Hence, for my long-run analysis, I allow nursing homes to adjust

\textsuperscript{102}For robustness, I run additional simulations under the assumption that the idiosyncratic utility shocks $\epsilon_{ij}$ are also mistakes due to information frictions. For example, if these shocks were due to factors such as the persuasiveness of a particular sales representative (which alters the resident’s perception of the nursing home, but is unrelated to the actual experience the resident will receive if she is admitted), then there is a case to be made that such preference shocks are mistakes.
their quality in response to the elimination of information frictions. To determine precisely how much nursing homes will respond to the change, I consider a simple model of quality competition between nursing homes and calibrate the model using quality and demand changes induced by the introduction of star ratings in 2008.\footnote{The structural demand estimation in section 4 provides evidence on demand changes due to the introduction of star ratings, while Appendix Figure A.19 shows support for the notion that quality may have improved due to the star ratings. Although the use of a single quality measure throughout most of this paper implicitly assumes time-invariant quality, which seems at odds with the results on quality changes due to the introduction of star ratings at the end of 2008, it is the relative differences of nursing home quality in each point in time that matters for my analysis on resident demand for quality, which is the focus of this paper. Appendix Figure A.20 shows that nursing home quality estimated before and after 2008 are well-correlated, so the use of a single quality measure (estimated using data from 2000–2010) is unlikely to cause issues and is more precisely estimated.} The quality competition model and its calibration are described in greater detail in Appendix Section I.

I also compare these effects with the simulated effects of various supply side policies — specifically, a minimum standard mandate and a pay-for-performance scheme.\footnote{Specifically, I assume that under the minimum standard mandate, the lowest-quality nursing homes are forced to improve their quality to a minimum acceptable level (corresponding to the 10th certain percentile of the quality distribution). To simulate the effect of a pay-for-performance scheme, I use estimates from Werner, Konetzka, and Polsky (2013) as a benchmark.} This provides an alternative way to interpret the magnitude of costs due to information frictions, and also sheds light on the potential efficacy of other policies if an effective information intervention is difficult to implement.

It should be noted that these counterfactual simulations require additional assumptions, which I discuss in Appendix Section G.1. At a high level, most of these assumptions restrict nursing homes’ responses to changes in demand (ruling out, for instance, counterfactual expansions/contractions or entries/exports). Because these restrictions on nursing home behavior become increasingly tenuous over longer time horizons, I restrict my counterfactual simulations to the year 2009.

### 5.1.2 Treatment Effect Mechanisms

Because there are multiple market imperfections (specifically, information frictions and selective admissions), the theory of the second best implies that eliminating information frictions may not necessarily improve resident outcomes. Distributional consequences are likewise ambiguous a priori. As a framework for understanding the various forces at work, I consider a simple model of an information intervention’s effect on resident outcomes in Appendix Section I, but here I summarize the main insights from this model with the aid of Figure 6.\footnote{This model only considers the partial equilibrium effect of an information intervention (i.e., it does not account for possible responses by nursing homes).}

Although the model applies to information interventions more generally, for expositional purposes the description here focuses on an information intervention that eliminates information frictions. The
black solid arrows show the channels through which the intervention may improve resident outcomes for treated residents. However, the red solid arrows show that improved nursing home choice by a given cohort of residents imposes negative externalities on future cohorts. This is because they increase crowding at high-quality nursing homes, making it more likely that these nursing homes will reject future residents.\footnote{These negative spillover effects also imply that the Stable Unit Treatment Value Assumption (Cox 1958; Rubin 1978; Angrist, Imbens, and Rubin 1996) may be violated, given that a resident’s outcome may depend not only on her own treatment but also on the treatment received by other residents.} The magnitude of these negative spillovers relative to improvements in choice quality will determine whether the intervention improves resident outcomes.

The dashed lines in the figure show that there are at least three sources of treatment effect heterogeneity. First, the two blue dashed arrows in the bottom left show that there is greater potential for improvements in nursing home choice for the least-informed residents. This is because the elimination of information frictions improves the information of these residents the most and because they would likely have stayed in low-quality nursing homes absent the intervention, and thus that there is greater room for improvement. I call this the \textit{reallocation effect}, and in this instance it is likely to be progressive.\footnote{I refer to an information intervention as progressive if its effect is increasing in baseline information frictions.}

Second, the green dashed arrow in the bottom right shows that there is treatment effect heterogeneity stemming from the fact that quality matters more for certain residents than for others. In this setting, quality improves the outcomes of residents with the greatest baseline mortality risk the most,
and baseline mortality risk is positively correlated with information frictions. Therefore, even though an information intervention that eliminates information frictions may be directly targeted at the least informed residents, it also indirectly targets residents for whom the “treatment effect of quality” is greater. Hence, I call this the indirect tagging effect, and it is also likely to be progressive.

Finally, the red dashed arrow shows that there is heterogeneity induced by negative spillovers, which in contrast to the previous two effects, is likely to be regressive. This is because the least informed residents also happen to be the types of residents that nursing homes find unattractive. Therefore, negative spillovers from increased crowding at higher-quality nursing homes are likely to fall disproportionately on this group of residents.

5.2 Simulation Procedure

5.2.1 Competing Risks Model of Deaths and Discharges

While the structural model in section 4 provides a way to simulate nursing home admissions that accounts for selective admissions by nursing homes, we also need a way to simulate resident exits from nursing homes. This is because nursing homes' admission decisions depend on occupancy, which is a state variable that is a function of both resident admissions and exits. In addition, because the outcome of interest is mortality, we must decompose resident exits into deaths and discharges. Given that deaths and discharges are competing risks,¹⁰⁸ I use a model of competing risks that accounts for this codependency (specifically, a model with cause-specific hazards that depend on resident and nursing home characteristics), which I estimate semiparametrically by maximizing a modified partial likelihood.¹⁰⁹ Finally, simulations of death and discharge times also depend on the nature of the correlations for the mortality and discharge shocks (within resident, across potential nursing homes), and I adopt the assumption that will result in the most conservative treatment effects.¹¹⁰ For details

¹⁰⁸ For a resident who dies in her nursing home, we do not observe when she would have been discharged had she stayed alive. Similarly, for a resident who is discharged from her nursing home, it is unknown whether she would have died if she had stayed in her nursing home.

¹⁰⁹ Standard hazard methods that do not account for competing risks will typically produce misleading estimates. For example, if one simply used the Kaplan-Meier estimates of the survival functions for deaths and discharges separately, the sum of these two estimates would exceed the Kaplan-Meier estimate of the survival function for the composite event (i.e., any type of exit), regardless of whether the two types of events were independent.

¹¹⁰ For example, if a resident suffers from an underlying condition that would have resulted in her death regardless of the nursing home she stayed in (although perhaps not at the exact same time), her mortality shocks across nursing homes will be highly correlated. By contrast, if a resident dies due to an accident (e.g., an accidental fall) and the same accident is unlikely to occur had she stayed in a different nursing home, the mortality shocks for the resident across potential nursing homes may be independent. For my baseline simulations, I assume that mortality and discharge shocks within resident across potential nursing homes are perfectly correlated because under this assumption a resident’s outcome is less likely to change even if she went to a different nursing home, and thus, the simulated treatment effects are likely to be conservative.
about the competing risks model and the procedure for simulating death and discharge times, see Appendix Section G.

### 5.2.2 Estimates of the Competing Risks Model

Appendix Figure A.13 shows the estimated cause-specific baseline hazard functions for death and discharges, revealing that their hazards tend to be greatest early in a resident’s stay and that the baseline hazard for discharges is an order of magnitude greater than for deaths. To check whether these estimates agree with our intuition about nursing home quality, Appendix Figure A.14 plots survival curves for deaths and discharges separately for nursing homes at the 75th and 25th percentiles of quality. Consistent with the construction of the survival-based quality measure, the probability of survival is substantially higher at every point in time for nursing homes at the 75th percentile of the quality distribution relative to nursing homes at the 25th percentile. The gap between the survival functions for resident discharges is much smaller because my quality measure is not based on discharges.\footnote{Moreover, there are different types of discharges, which can represent either a positive or negative outcome. Typically, upstream discharges (e.g., readmission to a hospital from the nursing home) are undesirable, whereas downstream discharges (e.g., discharge back to the community) are desirable, and the number of downstream discharges tend to outnumber upstream discharges (Einav, Mahoney, and Finkelstein 2022).}

### 5.2.3 Outline of Simulation Algorithm

I simulate outcomes for residents in order of date of entry. For a given day, I first simulate admissions using the structural model in section 4: residents are admitted to the nursing home they value the most (based on their counterfactual preferences) among the set of nursing homes willing to admit them (based on counterfactual occupancies). I then simulate their exit time and cause of exit based on the competing risks model, which I use to update future nursing home occupancy when my simulation reaches that stage. Finally, before moving onto the next day in my simulations, I update nursing home occupancies for the start of the next day based on admissions and exits that occur on the current date. For the full simulation algorithm, see Appendix Section G.4.

### 5.3 Simulation Results

#### 5.3.1 Average Treatment Effects

Figure 7 shows my main simulation results for the average treatment effect of eliminating information frictions. The solid black line shows the simulated short-run effects, using different estimates of the
MRS (between quality and distance) from the literature as the perfect information benchmark. The results indicate that even under my most conservative assumption about the MRS,\(^\text{112}\) the short-run effect of eliminating information frictions is an 8 percent reduction in nursing home deaths, and the effect can be as large as a 28 percent reduction in deaths if we use other MRS estimates from the literature. Similarly, if we assume that the idiosyncratic utility shocks \(\epsilon_{ij}\) are also mistakes due to information frictions, the dashed black line shows that the short-run effects will be even larger. Additionally, I explore the plausibility of my assumption about the correlation for discharge and mortality shocks in Appendix Table A.17 and find that my assumption is reasonable.\(^\text{113}\)

Next, I compare the short-run effects of eliminating information frictions to the simulated effects of two supply side policies — a minimum standard mandate and a pay-for-performance policy — which are shown as horizontal red and blue lines. The results show that the effect of a minimum standard mandate that forces nursing homes in the lowest 10 percentiles of the quality distribution to raise their quality to the 10th percentile is similar to the most conservative estimate of the short-run effect of eliminating information frictions,\(^\text{114}\) whereas the simulated effect of a pay-for-performance policy calibrated based on results from a program in Georgia (Werner, Konetzka, and Polsky 2013) is smaller.

Finally, the figure shows that the simulated long-run effect of eliminating information frictions (under the most conservative MRS assumption) is much larger than the other simulated effects, resulting in 44 percent fewer deaths. Appendix Table A.18 shows that the exact magnitude of the long-run effect varies depending on the functional forms for the demand and cost functions in the model of quality competition between nursing homes, but the simulated long-run effects are several times larger than the short-run effect in all cases, suggesting that supply side responses may substantially amplify the

\(^{112}\)Estimates of the MRS between quality and distance in the literature range from Chandra, Finkelstein, Sacarny, and Syverson’s (2016) MRS estimate of 1.8 at the low end, to Tay’s (2003) MRS estimate of 8 at the high end. For my most conservative assumption, I use the estimate from Chandra, Finkelstein, Sacarny, and Syverson (2016), and moreover, I divide their MRS estimate of 1.8 by 3 because their quality measure is based on 30-day mortality, whereas my quality measure is based on 90-day mortality.

\(^{113}\)Specifically, I compare my baseline assumption that mortality shocks are perfectly correlated within resident across nursing homes (and likewise for discharge shocks) to an alternative assumption of independent shocks. In Appendix Table A.17, we observe that under my baseline assumption, the average change in quality of residents’ chosen nursing home when information frictions are eliminated is 0.003, and that the effect on survival rate is 0.005, which is slightly larger. This makes sense because quality is defined based on risk-adjusted 90-day survival rate, whereas for my simulations I consider all deaths within the simulation period of a year so that most residents are “at risk” for a period longer than 90 days in the simulations. By contrast, the table shows that under the assumption of independent shocks, the average increase in quality of residents’ chosen nursing home is somewhat similar, and yet the average increase in survival rate of 0.021 is an order of magnitude greater, suggesting that this alternative assumption yields an overestimate of the treatment effect.

\(^{114}\)While the effect of a minimum standard mandate that affects only 10 percent of nursing homes seems quite large, increasing the proportion of nursing homes affected by this mandate in my simulations does not yield much greater improvements. This is because the bell shape of my quality distribution implies that nursing homes in the lowest 10 percentiles are in the left tail, so moving them to the 10th percentile leads to a large improvement in quality (whereas a given percentile increase at higher quality percentiles will result in smaller quality improvements).
direct demand side effects of an information intervention.

Figure 7: Main Simulation Results

Notes: The figure shows the fraction of nursing home deaths averted under various policy interventions. The labels for the (log of) MRS corresponds to different estimates from the literature, where CFSS (2016) refers to Chandra, Finkelstein, Sacarny, and Syverson (2016) and RG (2011) refers to Romley and Goldman (2011). For my most conservative MRS for the perfect information benchmark, I divide the smallest MRS in the literature (from CFSS) by three because the quality measure in their paper was based on 30-day mortality, whereas the quality measure in the present paper is based on 90-day mortality.
5.3.2 Distributional Consequences

In Figure 8, I display a visualization of the treatment effect heterogeneity by plotting the predicted short-run treatment effect of eliminating information frictions (based on the most conservative assumption about the MRS) as a function of the first two principal components of resident characteristics. The shape of the treatment effect surface suggests that there may be substantial heterogeneity. Moreover, because the base of the plot corresponds to zero treatment effect, the fact that the surface mostly lies above the base suggests that the intervention might be a Pareto improvement (in the ex ante sense, i.e., based on expected treatment effects).\textsuperscript{115}

To gain a better understanding of the types of residents who benefit more from the intervention, I study treatment effect heterogeneity by resident characteristics in Appendix Figure A.18. The results indicate that the intervention has progressive distributional effects: we observe that residents from relatively disadvantaged groups (e.g., older residents, residents likely to be long stay as proxied by non-post-acute care, and residents with dementia) tend to benefit more from the elimination of information frictions.\textsuperscript{116} Consistent with these results, I find that one standard deviation greater baseline information frictions is associated with 12.5 percent greater treatment effect relative to the average treatment effect.

The findings that eliminating information frictions does not harm any subgroup of residents and has progressive distributional consequences suggest that negative spillover effects play a relatively minor role. This is indeed what I find when I decompose the treatment effect and its heterogeneity into various components using the simple model of treatment effect mechanisms described earlier (see Appendix Section H.2 for details on this decomposition).\textsuperscript{117} Interestingly, the decomposition also reveals that the indirect tagging effect is more important than the reallocation effect in explaining the treatment effect heterogeneity.\textsuperscript{118,119}

\textsuperscript{115}Specifically, I regress the individual-level treatment effect (which is equal to 1 if the resident survives in the simulation but dies in practice, –1 in the opposite case, and 0 if the actual and simulated outcomes are the same) on baseline information frictions (which I define using the negative of the estimated demand for quality).

\textsuperscript{116}Appendix Figure A.18 also explores treatment effect heterogeneity by resident characteristics for the long-run effects of eliminating information frictions, the minimum standard mandate, and a pay-for-performance policy. The coefficients for the different policies are scaled by the average treatment effect of the policy to make the magnitudes comparable. The distributional effects for these policies are qualitatively quite similar for most resident characteristics, reflecting to a great extent the fact that the health of these groups of residents tends to be more sensitive to quality differences. The primary exception is that residents with dementia benefit by a much larger extent under the long-run effect of eliminating information frictions. This is because the health of residents with dementia tends to be sensitive to quality differences over a relatively large range of quality, and thus they benefit relatively more from the large quality improvements in the long-run analysis.

\textsuperscript{117}Specifically, negative spillovers explain less than 5 percent of the treatment effect and its heterogeneity compared to the reallocation effect.

\textsuperscript{118}In particular, the indirect tagging effect is roughly 1.6 times as important as the reallocation effect.

\textsuperscript{119}One reason for this finding of limited externalities may be because the simulations are limited to a single year.
Notes: This figure shows predicted reduction in mortality if information frictions were eliminated, as a function of the first two principal components of resident characteristics. The predicted reduction in mortality are obtained by fitting a third-order polynomial in the two principal components to the change in outcome. The change in outcome is defined to be one if the counterfactual outcome is survival and the observed outcome is death, negative one in the opposite case, and zero if the counterfactual and observed outcomes are equal. The base of the plot corresponds to no reduction in mortality.

Finally, the finding that negative externalities are small is not at odds with the evidence on selective admissions presented earlier. In particular, although higher-quality nursing homes are more crowded in the simulations and reject residents more frequently, there are typically numerous relatively high-quality nursing homes in residents’ choice sets. Thus, it is rare that a resident is rejected by all of them and forced to go to a low-quality nursing home.

Capacity at high-quality nursing homes may become more strained in the long-run as the number of long stay residents increases over time at these nursing homes. By contrast, unmodeled expansion and entry of high-quality nursing homes are likely to mitigate these capacity strains over longer time horizons.
5.4 Discussion

The short-run effects of information frictions are a conservative lower bound for a number of reasons. First, we observed that allowing endogenous quality choice by nursing homes may lead to larger reductions in mortality, and there is reason to believe that other unmodeled dynamics such as entry and expansions (respectively, exit and contractions) of higher-quality (lower-quality) nursing homes may lead to further gains. Second, by using MRS estimates from the previous literature as the perfect information benchmark, I have implicitly assumed that there are no information frictions in these other settings. Third, I observed that under alternative modeling assumptions (e.g., about the correlation between mortality and discharge shocks, and whether idiosyncratic utility shocks are mistakes), the simulated effects are even larger. Finally, mortality is a narrow measure of quality, and to the extent that it is correlated with other dimensions of quality, the true cost of information frictions will also be larger than my estimate.

While the cost of information frictions may be large, it may be challenging to implement effective information interventions. A vast literature examining the effects of information interventions in other settings have found mixed evidence, but there seems to be a general theme that even when the average effect of information interventions is positive, they often do not have desirable targeting properties unless they are designed explicitly with targeting in mind.

An information intervention by the CMS — the introduction of the five-star ratings system — seems to suffer from this same shortcoming. In particular, although we estimated that the introduction of star ratings increased average demand for quality, additional results in Appendix Table A.11 suggests that it may have also increased disparities. In addition, research studying the time period after the end of my sample suggests that the usefulness of the five-star ratings system may have declined over time, as nursing homes become increasingly adept at gaming the ratings (Han, Yaraghi, and

\footnote{For example, there is work studying the effects of information interventions on enrollment in various public programs such as the Earned Income Tax Credit (Manoli and Turner 2014; Bhargava and Manoli 2015; Guyton et al. 2017; Barr and Turner 2018), Social Security Disability Insurance (Armour 2018), and SNAP (Daponte, Sanders, and Lowell 1999; Finkelstein and Notowidigdo 2019), as well as in education settings (Hastings and Weinstein 2008; Bettinger, Long, Oreopoulos, and Sanbonmatsu 2012; Barr and Turner 2018; Dynarski, Libassi, Michelmore, and Owen 2018; Berman, Denning, and Manoli 2019; Cohodes, Corcoran, Jennings, and Sattin-Baja 2022).}

\footnote{For instance, the marginal individuals affected by information interventions are often of relatively high socioeconomic status, for whom treatment effects may be relatively small (Finkelstein and Notowidigdo 2019).}

\footnote{In particular, I estimate differential changes in preferences post-star ratings by including triple interactions between quality, a dummy for post-star ratings, and various resident characteristics. The results shown in Appendix Table A.11 suggest that demand for quality among highly educated residents may have increased to a greater extent compared to residents who were black or who had dementia.}

\footnote{Similarly, Konetzka et al. (2015) find that the five-star ratings system exacerbated disparities in quality by payer source.}
Finally, the five-star ratings system is a rather crude measure of quality; although earlier results in Table 2 show that my quality estimates are correlated with star ratings, Appendix Figure A.21 shows that there is substantial overlap even between the quality distributions of one-star and five-star nursing homes.125

Quality measures such as the risk-adjusted survival rate estimated in this paper can potentially provide residents with a better way to evaluate nursing home quality than the five-star ratings for several reasons.126 First, it provides more detailed information about relative quality of nursing homes, and can be validated using quasi-experimental methods (e.g., the distance-based IV strategy).127 Second, this type of quality estimate may be harder for nursing homes to game for several reasons — mortality is rarely misreported, and the complexity of the risk adjustment makes it difficult for nursing homes to determine how to effectively manipulate controls used in the risk adjustment effectively.128 However, a potential concern is that more prominent use of this quality measure may result in nursing homes rejecting high-risk residents at higher rates; specifically, even if the risk adjustment is adequate, risk-averse nursing homes may still have an incentive to reject high-risk residents if there is estimation noise in the quality measure (Dranove, Kessler, McClellan, and Satterthwaite 2003).

The specifics of designing an effective information intervention depend on the nature of information frictions, which for example could be due to lack of relevant information or from the inability to process publicly observable nursing home characteristics that are associated with higher quality suggests that

---

124 For example, although the share of nursing homes that receive five star ratings (the highest possible score) has increased over time, there has been limited corresponding improvements in resident outcomes.

125 This is consistent with the general sentiment in the industry. For example, a primer on nursing home standards by a resident advocacy organization notes that “a high star rating is not necessarily a guarantee of the highest quality, but low ratings are generally a cause for concern.”

126 There are several practical implementation issues that are worth mentioning. First, there may be concerns that consumers will find it more difficult to understand such a quality measure (e.g., it is easier to understand what a five-star rating means than it is to understand a risk-adjusted survival measure of 0.015). Nonetheless, this can be mitigated by existing tools, e.g., the nursing home compare website maintained by the CMS allows consumers to sort nursing homes by star ratings, so the same can be done for a risk-adjusted survival measure. In addition, the CMS can also improve the interpretability of the quality measure by standardizing the measure or reporting the percentile rather than the underlying risk-adjusted survival rate.

Second, if the CMS does use risk-adjusted survival rate as a quality indicator, it is unlikely to risk-adjust based on protected characteristics such as race and gender (whereas these variables are used in the risk adjustment for the quality estimates in this paper). Nevertheless, the estimates of quality I obtain without controlling for protected characteristics are highly correlated with the estimates using the full set of controls (a bivariate regression between these two quality measures has an R-squared of 0.98).

127 Mogstad, Romano, Shaikh, and Wilhelm (forthcoming) develop methods to assess rankings that are based on estimates. Nonetheless, even if there is uncertainty in the exact ranking of nursing home by quality, the more granular quality measure based on risk-adjusted survival rate is likely to convey more information than the five-star ratings, where poor conditions have been found even at five-star nursing homes (New York Times 2021).

128 Interestingly, if publishing a risk-adjusted quality measure does in fact substantially improve resident choice, there is the potential for accompanying changes in selection patterns to lead to selection bias for the quality measure (which was not present before), in the spirit of the type of dynamic selection discussed by Oster (2019). To avoid this, quality estimates should periodically be validated to ensure they remain informative (e.g., using quasi-experimental methods from the value-added literature).
the first explanation plays a role,\textsuperscript{129} and indeed, a small-scale survey by Konetzka and Perraillon (2016) finds that use of the Nursing Home Compare website by residents’ family members is limited by lack of awareness and mistrust of the data. A survey similar to the one conducted by Handel and Kolstad (2015) may help us better understand the sources of these information frictions, although implementing such a survey may be tricky given that it is not always clear who ultimately chooses the nursing home in this setting.

The issue of who the decision-maker is likewise key for the design of an information intervention. For example, the resident and her family members are most likely the ones ultimately choosing the nursing home, but hospital discharge planners also play an important role. While we may expect hospital discharge planners to have a better grasp of nursing home quality given their prior experience, they often do not provide patients about to be discharged with information about nursing home quality, possibly due to legal risks associated with choice requirement laws (Tyler et al. 2017).\textsuperscript{130}

Finally, in the case that it is too difficult to implement an effective information intervention, my simulations suggest that under optimistic assumptions, a minimum standard mandate may be able to achieve gains similar to the most conservative short-run effect estimate of eliminating information frictions. Nonetheless, past evidence suggests that minimum standard mandates face challenges of their own. Federal and state governments have increased nursing home regulations over the past few decades,\textsuperscript{131} and despite these copious regulations,\textsuperscript{132} quality of care is often still lacking. One reason is that nursing homes are often able to find ways around many of these regulations, so for increased regulations to have the desired effect, they must at least be paired with effective means of monitoring compliance.\textsuperscript{133,134}

\textsuperscript{129}One particular type of error in processing information related to mortality risk — probability weighting from prospect theory (Kahneman and Tversky, 1979) — seems less likely to explain the observed patterns. Specifically, under typical probability weighting functions, the baseline mortality rate of 7.5 percent is likely to cause residents to overweight the probability of dying, which should lead to greater demand for quality.

\textsuperscript{130}For example, Tyler et al. write that according to the senior employee at a hospital, “legal experts have said that we’re not able to do that [provide recommendations or data] outside of any sort of Medicare innovation programs, and because...they are very concerned about the requirement for the patient to feel that they have choice in where they go postdischarge.”

\textsuperscript{131}Examples include federal regulations over reporting requirements and the use of physical or chemical restraints as well as state regulations over minimum staffing standards.

\textsuperscript{132}A common refrain among those studying or working in the nursing home industry is that nursing homes are the second most regulated industry after nuclear energy.

\textsuperscript{133}An example of this is the introduction of payroll-based journal reporting of nursing home staffing in 2016 that led to a significant improvement in the reliability of staffing measures relative to earlier self-reported data, which were largely unaudited (Geng, Stevenson, and Grabowski 2019).

\textsuperscript{134}Nursing homes often argue that they are already operating on thin margins and that the costs of complying with additional regulations may force them to shut down. However, it is not easy to verify these claims about their profit margins, considering that nursing homes’ finances are notoriously opaque and profits are often disguised using related party transactions. The Brius nursing home chain in California is a notable example, but this practice is also common elsewhere. For example, a review of 2020 nursing home financial reports gathered by the New York State Department of Health reveals that 72 percent of nursing homes in the state engage in some form of related party transactions (NYS
6 Conclusion

In this paper, I study demand for quality in the nursing home market. I demonstrate that despite substantial variation in nursing home quality and the consequential impact quality can have on resident health, many residents, especially those with greater information frictions, do not take advantage of this. The costs of these information frictions are substantial, accounting for at least 8–28 percent of nursing home deaths, even before considering possible supply side responses by nursing homes.

The finding that information frictions play an important role naturally leads to a number of follow-up questions. In order for policymakers to design effective information interventions, it is critical for us to gain a better understanding of the underlying sources of these information frictions. Moreover, information frictions are also relevant for critical social issues such as racial disparities. For example, in ongoing work, I find that in addition to in-group preferences and discrimination by nursing homes, information frictions is another contributory factor to the disproportionate concentration of minorities in low-quality nursing homes.

Department of Health 2020).
References


[77] Long Term Care Community Coalition, 2021. “Nursing Home Quality Standards.”


Appendix

A  Additional Background on Nursing Homes

A.1  Brief History of Nursing Home Quality

The growth and development of the nursing home industry in its early days was driven in large part by the creation of the Medicare and Medicaid programs in the 1960s. Since then, reimbursements from federal and state governments have been a major source of revenue for the industry, and as such, federal and state oversight bodies have a significant influence on nursing home quality. The large role played by Medicare and Medicaid has also meant that most residents face limited variation in out-of-pocket price when choosing nursing homes.

Despite the entry of nursing homes during this early era, demand often outstripped supply, with many nursing homes operating at maximum capacity, and essentially having to ration care. This “excess demand” limited competition between nursing homes, and dampened financial incentives to compete by providing better quality. The establishment of the Health Care Financing Administration (HCFA) in 1977 was in part a response to the persistently low quality of care at many nursing homes. Process quality indicators were introduced as part of the attempt to increase nursing home accountability, but lobbying efforts by the nursing home industry as well as the increasing complexity of residents’ medical needs limited quality improvements.

These persistent issues led to the Nursing Home Reform Act in 1987 (which refers to parts of the Omnibus Budget Reconciliation Act of 1987 [OBRA-1987] that were specific to nursing homes), which brought about significant and wide-ranging changes in the nursing home industry. These changes included a revision of the quality standards and penalties, and the introduction of the Resident Assessment instrument, which includes the MDS. Over the years, elements and regulations introduced by the Nursing Home Reform Act were revised and updated — for instance, the MDS 2.0 superceded the original MDS in 1995, which was then replaced by the MDS 3.0 in 2010. In addition, changes were made to the way that data on nursing home quality was presented to the public — a major component of these changes was the introduction of the CMS five-star ratings in 2008, which attempted to simplify information on nursing home quality for residents by summarizing a number of different quality indicators in a composite measure. Occupancy rates at nursing homes have also fallen substantially over the years, but remain relatively high.
There is general agreement that nursing home quality has improved over the past few decades, even among resident advocates. For example, one industry insider remarked that whereas chemical restraints were used openly by nursing homes 20 years ago, there is at least now a recognition that this is not an acceptable practice (although some nursing homes may still use it to some extent behind closed doors). However, a plethora of lawsuits as well as pieces from investigative journalism have revealed that quality in some nursing homes remains poor. In addition, these cases have shown that various indicators of nursing home quality (e.g., the five-star ratings) are not always reliable, as nursing homes have found ways to game these quality measures.

Recently, the COVID-19 pandemic has shaken up the nursing home industry. First, the significant number of COVID-related nursing home deaths has increased public scrutiny over nursing home quality.\textsuperscript{135} Second, the pandemic led to a precipitous fall in short-stay residents at nursing homes, as individuals have shied away from medical procedures (e.g., knee surgeries) that may require them to stay in a nursing home. Given that nursing home care for these individuals is typically covered by Medicare, and because Medicare reimbursement rates are substantially higher than Medicaid rates, this trend threatens the financial viability of nursing homes if they were already operating on thin profit margins before the pandemic.

A.2 More Details on Various Nursing Home Quality Measures

Numerous quality measures have been used for nursing homes, and Donabedian (1985) provides a useful classification of these measures. Specifically, Donabedian argued that these quality measures are either based on structures (S), processes (P), or outcomes (O). Structural measures refer to nursing home characteristics that are associated with the provision of care for residents (e.g., staffing levels). Process measures refer to the care received by residents, negative examples of which include the use of physical restraints and inappropriate antipsychotic use. Finally, outcomes measures include mortality (the primary measure used in this paper) and other adverse outcomes such as pressure sores and avoidable falls.

Some quality indicators span several of these categories. For example, regulators conduct inspections of nursing homes annually (or as a result of a complaint), and deficiency citations refer to areas wherein inspectors determine that the nursing home has failed to meet CMS requirements. More than 200 aspects covering 19 categories are examined during inspections, and in addition, deficiency citations

\textsuperscript{135}The Kaiser Family Foundation found that long-term care facility residents and staff account for at least 23 percent of all COVID-19 deaths in the U.S, as of the end of January 2022.
are rated based on severity (level of jeopardy to resident health) and scope (e.g., whether this is an isolated incident or a systemic problem). Some types of deficiency citations (e.g., inadequately trained staff) fall under the structures (S) category of Donabedian’s classification, whereas other types fall under the process (P) category (e.g., abuse or neglect of residents), and the severity of these deficiency citations are often based on resident outcomes (O).

One of the most well-known measures is the five-star rating system, which the CMS introduced at the end of 2008. The goal of the five-star rating system is to allow consumers to assess nursing home quality more easily — even though consumers were able to access data on quality measures such as staffing levels, ownership status, and deficiency citations through the Nursing Home Compare website previously, the concern was that the multitude of indicators made it difficult for consumers to compare nursing homes. Hence, the five-star rating provides a summary index of many of the quality measures discussed here.

Specifically, the star ratings are calculated based on scores on three domains: the health inspection domain, the staffing domain, and the quality measure domain. The health inspection score is determined by results of the inspections of nursing homes by regulators, the staffing score is calculated using case-mix adjusted staffing hours (for different types of staff) per resident day, and the quality measure domain combines performance on 15 types of resident outcomes (10 of which are derived from MDS assessments and five from Medicare claims).

A persistent issue with many of these quality measures is gaming by nursing homes. Staffing levels during the period of my sample were self-reported, rarely audited, and likely to be inflated (Geng, Stevenson, and Grabowski 2019). Outcome-based measures also show signs of manipulation; for example, after the government began publicly releasing information about inappropriate antipsychotic use in 2012, Schizophrenia diagnoses rose sharply, the implication being that some were made not on a clinical basis but simply so that nursing homes could continue administering antipsychotics to residents as a form of chemical restraint without having it counted as inappropriate antipsychotic use (New York Times, 2021). Deficiency citations may not always provide an accurate picture of nursing home quality either, given that nursing homes may behave differently around the time of inspections. For instance, Geng, Stevenson, and Grabowski (2019) find that staffing levels at nursing homes spike during the week of inspections, and there are anecdotal accounts of other strategic behaviors by nursing homes.

\[136\]

As an example, management at a nursing home under private equity ownership had a designated code to covertly alert staff of the presence of inspectors, e.g., “Marilyn Woods, line twelve” (New York Post, 2022).
B Data Appendix

B.1 Sample Construction

For the quality estimation sample, I drop the relatively small number of residents with errors in birth or death dates (e.g., with different birth or death dates recorded across different assessments), or with missing values of the resident characteristics that I control for. I include all possible baseline resident characteristics recorded in the MDS, other than a few variables that are missing for a large proportion of residents (e.g., HIV status). Finally, the provider number for some residents cannot be matched with nursing homes in the OSCAR dataset. I examine each of these unmatched provider numbers, correcting obvious typos (e.g. the digit 0 being replaced with a letter O) and dropping observations when it cannot be determined which nursing home the provider number corresponds to.

For the structural demand estimation sample, I consider only resident-nursing home pairs within 15 miles of each other. In addition, I drop nursing homes that admit fewer than 30 residents over the period of my structural estimation sample (2008–2010) to ensure sufficient power. I include a smaller set of variables in the structural demand estimation (and the sample for hazard estimation) compared to my quality estimation sample for computational reasons and for easier interpretation of their coefficients in residents’ utility equation (and the supply side equation). I do not include expected out-of-pocket expenditures associated with each residents’ utility equation because there is relatively little variation in out-of-pocket costs for residents covered by insurance, and Gandhi (2019) finds very little price sensitivity on the part of residents. Moreover, a large portion of the variation in expected prices comes from differences in the distribution of lengths of stay and when residents will switch to a different payer source (e.g., because Medicare coverage expires in the case of short-stay residents or because residents spend down their assets sufficiently to become eligible for Medicaid in the case of long-stay residents). These are unknown to the residents at the time of their nursing home choice, and it is difficult to determine how accurately they are able to predict these quantities, especially given the substantial information frictions and heterogeneity documented in this paper.

B.2 Payer Source Data

The data on payer source at admission recorded in the MDS is less reliable than claims data. This is in part because nursing homes are sometimes unsure at the time of admission (which is when they complete the initial MDS assessment form) how they are ultimately going to be reimbursed for the
resident, and in fact, claims they submit may be rejected months after. For example, as the CEO of a firm providing billing assistance to nursing homes put it:

Most patients that need Medicaid for long term care don’t actually have Medicaid for long term care in place when when they are coming into the facility... There’s been so many scenarios where a patient comes into the facility, the responsible party or power attorney says, sure Mom and Dad are eligible... but when you get your hands on those accounts, you see that there was $50,000 transferred out of that account 3 months before the patient comes into the nursing home, which is a potential disqualifier.

Nonetheless, given that I use payer source primarily to study/control for selective admissions by nursing homes, it is unclear whether the actual payer source from claims data is preferable. This is because nursing homes’ admission decisions depend on what they expect the payer source to be (which is presumably the payer source recorded in the MDS) rather than what the ultimate payer source ends up being (which is recorded in the claims data). Moreover, given that I am interested in residents from all payer sources, if one insisted on using claims data for verification of payer source we would need claims data from numerous programs (e.g., Medicare, Medicaid, and the Veterans Health Administration) and will likely still not be able to verify payer sources for residents on private insurance or who are private pay.

A separate issue with the payer source recorded in the MDS is that it is often not updated in assessments subsequent to the initial admission assessment (Grabowski, Gruber, and Angelelli 2008). This shortcoming of the data is largely irrelevant to my analysis because I focus on resident characteristics at admission.

B.3 Variable Definitions

Many of the answers to MDS questions are categorical (e.g., questions requiring the assessor to check all boxes that apply). In these cases, I include a dummy variable for each category (other than an omitted category, if it exists). There are also a number of numerical variables, such as weight, height, and age. For residents’ weight, I create dummies for weights in 10-pound intervals starting from 60–69 pounds up to 390–399 pounds (as well as a dummy for less than 60 pounds), and for height I create dummies for heights in 5-inch intervals from 40–44 inches to 70–74 inches (and a dummy for less than 40 inches). Similarly, I create the following dummies for residents’ ages: less than 40, 40–49, 50–59, 60–64, 70–74, 75–79, and 80–84. Finally, the assessor filling in the MDS can include up to five ICD-9 codes for each resident. I create a dummy for each unique first 3 digits of the ICD-9 codes, which
equals one for each resident if any of the 5 ICD-9 codes entered for the resident has the corresponding first three digits, and zero otherwise.

C Model for Quality Estimation

The causal equation \( Y_{ij} = \mu_j + a_i \),

\[
Y_{ij} = \mu_j + a_i, \quad (7)
\]

where \( Y_{ij} \) is the potential outcome for resident \( i \) in nursing home \( j \), \( \mu_j \) is a measure of nursing home \( j \)’s quality (in terms of the nursing home’s causal effect on the outcome \( Y \)), and \( a_i \) is a residual for resident \( i \).

I can relate potential outcomes to realized outcomes by summing across nursing homes in equation \( (7) \) to obtain:

\[
Y_i = Y_{i1} + \sum_{j=2}^{J} (Y_{ij} - Y_{i1})D_{ij} = \mu_1 + \sum_{j=2}^{J} \beta_j D_{ij} + a_i, \quad (8)
\]

where \( D_{ij} \) is a dummy variable for whether resident \( i \) chooses nursing home \( j \), and \( \beta_j \) is the quality of nursing home \( j \) relative to the omitted nursing home, indexed by \( j = 1 \). Finally, I decompose the resident residual, \( a_i \), into a component explained by resident characteristics \( X_i \) and an idiosyncratic component \( u_i \) by projecting \( a_i \) onto \( X_i \):

\[
a_i = X_i' \gamma + u_i, \quad E[X_iu_i] = 0, \quad (9)
\]

and substitute this into equation \( (8) \) to obtain the causal equation \( (1) \) in the main text.

To derive equation \( (2) \), which forms the basis of the IV estimation of the forecast coefficient, I consider the infeasible regression of causal effects \( \beta_j \) on the quality estimates \( \alpha_j \), normalizing \( \eta_j \) to have mean zero:

\[
\beta_j = \lambda \alpha_j + \eta_j, \quad E[\alpha_j \eta_j] = 0. \quad (10)
\]
Substituting equation (10) into the causal equation (1), I obtain equation (2) as desired.

D Alternative Identification Strategies for Estimating the Forecast Coefficient

In addition to distance between residents and nursing homes, temporary occupancy fluctuations provide an additional source of exogenous variation that may be used to estimate the forecast coefficient \( \lambda \). To take advantage of this, I estimate an IV model using variation in quality of residents’ chosen nursing homes induced by distance and temporary occupancy fluctuations. Specifically, I use \( \{ \alpha_i \sim t(i), k, \text{occ}_t(i), k \times \text{occ}_t(i), k \}_{k=1}^K \) as my instruments, where the superscript \( k \) denotes the \( k \)th closest nursing home to the resident.\(^{137}\) The idea behind the interaction terms is that if high-quality nursing homes close to a resident are close to capacity at the time the resident is looking for one, she is more likely to be admitted to a low-quality nursing home. Estimates of this IV specification, shown in Appendix Table A.6, reveal that adding occupancy instruments hardly changes estimates of the forecast coefficient (which are still not statistically different from one).\(^{138}\)

An alternative distance IV strategy is to use the quality of nearby nursing homes \textit{excluding the nursing home actually chosen by the resident} as instruments. The idea underlying the first stage of this strategy is akin to Bayesian updating: if a resident values quality and is well-informed about it, then the fact that there are high-quality nursing homes close to her but that she did not choose suggests that her chosen nursing home is likely to be high quality as well. It turns out that the first stage for this IV specification is violated, which is unsurprising given our finding from the structural demand estimation that average demand for quality is low and that information frictions are likely to be present.

Finally, one might consider estimating \( \lambda \) using an event study analysis. To this end, I use entry by high-quality nursing homes close to a resident as the event, and study the effect this has on the quality of nursing homes chosen by residents (which we can think of as the first stage) as well as on resident outcomes (which can be thought of as the reduced form). I implement the event studies using methods from Borusyak, Jaravel, and Spiess (2022), considering that recent work has highlighted the shortcomings of conventional estimation methods such as two-way fixed effects when treatment effects

\(^{137}\)The definition of the occupancy measure \text{occ} corresponds to the definition used for the structural demand estimation in section 4: the lagged 7-day average of log occupancy, residualized of nursing home-month fixed effects.

\(^{138}\)I did not include results for \( K > 3 \) for these specifications because the F-statistic suggests that these specifications suffer many weak instruments bias.
are heterogeneous and treatment is staggered (de Chaisemartin and D’Haultfœuille 2020; Sun and Abraham 2021; Callaway and Sant’Anna 2021; Borusyak, Jaravel, and Spiess 2022). Unfortunately, this event study is underpowered, and there is suggestive evidence that the parallel trends assumption is violated.

E Identification of the Matching Model

This section discusses how the model of demand and selective admissions maps into the framework in Agarwal and Somaini (2022), and the formal assumptions required for identification. I also briefly comment on how these assumptions relate to the setting in this paper. The few differences in the description below from my setup in section 4 (e.g., scale normalizations) make no difference to the identification argument and avoids the need to introduce additional notation.

Recall that residents’ and nursing homes’ preferences are given by:

\[ v_{ij} = v^1_j(w_i, \zeta_i) - v^2_j(w_i, dist_{ij}), \]
\[ \pi_{ij} = \pi(w_i, \zeta_i, occ_{ij}), \]

where I denote resident-specific preference heterogeneity by \( \zeta_i \). We can set \( |v^2_j(w_i, dist_{ij})| = 1 \) for the scale normalization for resident preferences, and set the utility for an arbitrary nursing home to be zero for the location normalization, e.g., \( v_{i1} = 0 \) (and do not include an intercept term).

I set the location normalization for the supply side equation so that nursing homes’ acceptance decision can be written as:

\[ \sigma_{ij} = \sigma_j(w_i, \zeta_i, occ_{ij}) \]
\[ = \mathbb{I}[\pi(w_i, \zeta_i, occ_{ij}) \geq 0]. \]

**Assumption I1.** Unobserved consumer-specific heterogeneity \( \zeta_i \) is conditionally independent of \( (dist^i_j, occ^i_j) \) given \( w_i \).

This is the formal statement for the exclusion restriction for the demand and supply instruments, which I provide evidence in support of in the main text.

**Assumption I2.** The acceptance decision function \( \sigma_j(w_i, \zeta_i, occ_{ij}) \) is weakly decreasing in \( occ_{ij} \). In addition, for all nursing homes \( j \) and resident characteristics \( w_i \), and unobserved preference hetero-
geneity \( \zeta \), we have \( \lim_{\text{occ} \to -\infty} \sigma_j(w_i, \zeta_j, \text{occ}) = 1 \), and \( \lim_{\text{occ} \to \infty} \sigma_j(w_i, \zeta_j, \text{occ}) = 0 \).

Recall that my occupancy measure represents temporary fluctuations (so negative values of \( \text{occ} \) make sense). The assumption that the acceptance decision function is weakly decreasing in the occupancy measure is the relevance condition for the supply instrument, which I provided evidence for when testing prediction 1 in the main text. The fact that nursing homes have finite capacities is consistent with the requirement that \( \lim_{\text{occ} \to \infty} \sigma_j(w_i, \zeta_j, \text{occ}) = 0 \). Finally, for \( \lim_{\text{occ} \to -\infty} \sigma_j(w_i, \zeta_j, \text{occ}) = 1 \) to be true, the direct costs of caring for each resident (i.e., not accounting for the option value of using up a spare bed) must be less than the marginal revenue the resident brings in. Gandhi (2019) provides evidence that this is true for most residents: even Medicaid reimbursement rates (which are much lower than Medicare rates) are typically sufficient to cover direct costs of care.

We will say that nursing homes \( j \) and \( k \) are strict substitutes in \( \text{dist}_{ij} \) at \( (w_i, \text{dist}_i, \text{occ}_i) \) if 
\[
\partial s_j(w_i, \text{dist}_i, \text{occ}_i) / \partial \text{dist}_{ik} \quad \text{and} \quad \partial s_k(w_i, \text{dist}_i, \text{occ}_i) / \partial \text{dist}_{ij}
\]
both exist and are strictly positive, where \( s_j(w_i, \text{dist}_i, \text{occ}_i) \) is the share of residents with \( (w_i, \text{dist}_i, \text{occ}_i) \) who are matched with nursing home \( j \).

Define \( \Sigma(w_i, \text{dist}_i, \text{occ}_i) \) to be the matrix where the \((j, k)\) entry is one if \( j \) and \( k \) are strict substitutes in \( \text{dist} \) at \( (w_i, \text{dist}_i, \text{occ}_i) \) and zero otherwise, and let \( \Sigma(w_i, \text{dist}_i) \equiv \bigvee_{\text{occ} \in \text{supp}(\text{occ})} \Sigma(w_i, \text{dist}_i, \text{occ}) \) so that the \((j, k)\) entry is one if \( j \) and \( k \) are strict substitutes at some occupancy.

**Assumption I3.** For every \( w_i \) and all but a finite set of \( \text{dist}_{ij} \) in its support, the graph of \( \Sigma(w_i, \text{dist}_i) \) has a path connecting any two nursing homes.

Roughly speaking, this assumption requires a path between any two pairs of nursing homes. An example where this assumption may potentially be violated is if I tried to estimate demand for nursing homes in California and Massachusetts (which are on different coasts of the US) in the same model. If no resident ever substitutes from a nursing home in California to one in Massachusetts and vice versa, we cannot identify how residents rank nursing homes in the two states relative to each other. In my setting where I study only nursing homes in California, given that residents and nursing homes are spread out over California, Assumption I3 seems plausible.

**Assumption I4.** (i) The support of the random vector \( \text{dist}_i \) is rectangular with non-empty interior. (ii) For each \( w_i \) and \( j \), the function \( v_j^2(w_i, \text{dist}_j) \) is continuously differentiable in \( \text{dist}_j \) with \( v_j^2(w_i, \text{dist}_j) \neq 0 \) for all \( \text{dist}_j \).

The second part of Assumption I4 is satisfied if residents’ utility for each nursing home is sufficiently smooth and strictly decreasing in her distance to it for all possible distances, and the first part is a weak requirement on the support of the demand instrument.
Under the assumptions above, residents' preferences and nursing homes' admission rules are non-parametrically identified.

**Theorem 1** (Agarwal and Somaini 2022). If Assumptions I1–I4 hold and \(|J| > 1\), then for every \(w\), (i) the function \(v_j^2(w, \cdot)\) is identified for every \(j \in J\) and \(\text{dist}_j \in \text{supp}(\text{dist}_j)\), and (ii) the joint distribution of \((u_i, \pi_i^{\text{cutoff}})\) is identified for every value \((u, \pi)\) in the interior of \(v^2(w, \text{supp}(\text{dist})) \times \text{supp}(\text{occ}) = \Pi_{j=1}^J v_j^2(w, \text{supp}(\text{dist}_j)) \times \text{supp}(\text{occ})\), where \(\pi_i^{\text{cutoff}}(w_i, \zeta_i) \equiv \sup\{\text{occ} : \pi(w_i, \zeta_i, \text{occ}) \geq 0\}\).

### F Algorithm for the Gibbs Sampler

To ease the computational burden of this estimation, I only consider nursing homes within 15 miles of each resident, dropping the small number of residents who choose a nursing home further away. Even with this restriction, the size of residents' choice sets (before any supply side constraints) tends to be quite large: it has roughly a median of 30 and a mean of 50, and can be greater than 200. Below, I will denote the choice set for resident \(i\) by \(J_i \equiv \{j \in J | \text{dist}_{ij} \leq 15 \text{ miles}\}\). In the following description for the Gibbs sampler, when drawing structural error terms in sequence for \(j \in J_i\), I assume an increasing order (although obviously any other order works as well). In addition, to simplify notation, I denote variables in the resident utility and supply side equations by \(W_{ij}\) and \(Q_{ij}\), respectively (these tend to be components of resident characteristics \(w_i\) and nursing home characteristics \(q_j\), as well as interactions between them), and refer to the nursing home that resident \(i\) ends up in by \(\mu(i)\).

The steps for implementing the Gibbs sampler are as follows (the superscripts represent the iteration of the Gibbs sampler).

1. **Initialization** (\(k = 0\), where \(k\) is the current iteration of the Gibbs sampler): I assume that \((\epsilon_{ij}, \omega_{ij}) \sim \text{i.i.d.} N(0, I_2)\) and set the following conjugate priors for the parameters: \((\kappa', \psi')' \sim N(0, 100I)\).

   (a) Set the initial values of the parameters \(\theta^0 = (\kappa^{0r}, \psi^{0r})\) at their prior mean.

   (b) Initial data augmentation: For each resident \(i\), draw the vector \(\epsilon_i^0\) such that \(v_{i,\mu(i)}^0 \geq v_{ij}^0\) for all \(j \in J_i\).

      i. Draw \(\omega_{i,\mu(i)}^0\) such that \(\omega_{i,\mu(i)}^0 \geq -Q'_{ij} \psi^0\) and for \(j \neq \mu(i)\) draw \(\omega_{ij}^0\) from the unconditional distribution.

      ii. Set \(\epsilon_{i,\mu(i)}^0\) equal to three times the standard deviation of the prior. For \(j \neq \mu(i)\), draw \(\epsilon_{ij}^0\) such that \(\epsilon_{ij}^0 \leq (W_{i,\mu(i)} - W_{ij})'k^0 + \epsilon_{i,\mu(i)}^0\) if \(\pi_{ij}^0 \geq 0\) or draw \(\epsilon_{ij}^0\) unconditionally.
2. For \( k + 1 = 1, \ldots, K \) (where \( K + 1 \) is the total number of iterations):

(a) Draw the profit shocks \( \omega_i^{k+1} \) in sequence for \( j \in J_i \).

i. If \( v_{ij}^k < v_{i,\mu(i)}^k \), draw \( \omega_{ij}^{k+1} \) unconditional on assignment (given that even if \( i \) is eligible for \( j \), \( i \) would not choose \( j \)).

ii. If \( v_{ij}^k > v_{i,\mu(i)}^k \), draw \( \omega_{ij}^{k+1} \) from a truncated normal with mean and variance given by the conditional distribution and truncation point \( \omega_{ij}^{k+1} < -Q_{ij}^k \psi_k \) (given that otherwise \( i \) would choose \( j \) over \( \mu(i) \)).

iii. Finally, if \( j = \mu(i) \), draw from the conditional distribution with truncation point given by \( \omega_{ij}^{k+1} \geq -Q_{ij}^k \psi_k \) (given that \( i \) must always be eligible for the facility she was ultimately assigned to).

(b) Update \( \pi_i^{k+1} \) according to \( \pi_i^{k+1} = Q_{ij}^k \psi_k + \omega_{ij}^{k+1} \).

(c) Draw the utility shocks \( \epsilon_i^{k+1} \) in sequence, for \( j \in J_i \).

i. If \( \pi_{ij}^{k+1} < 0 \), draw \( \epsilon_{ij}^{k+1} \) unconditionally (given that \( i \) would not choose such a facility even if she were eligible for it).

ii. If \( \pi_{ij}^{k+1} \geq 0 \) and \( j \neq \mu(i) \), draw \( \epsilon_{ij}^{k+1} \) from the conditional distribution with truncation point given by \( v_{ij}^{k+1} < W_{ij}^k \kappa^k \).

iii. For \( j = \mu(i) \), draw \( \epsilon_{ij}^{k+1} \) such that \( v_{ij}^{k+1} \) is larger than the current values of \( v_{i,j'} \) for \( j' \neq j \) and \( \pi_{ij'} \geq 0 \).

(d) Update \( v_i^{k+1} \) according to \( v_i^{k+1} = W_{ij}^k \kappa^k + \epsilon_{ij}^{k+1} \).

(e) Update the parameters \( \theta \) based on the new indirect utilities \( v^{k+1} \) and profits \( \pi^{k+1} \).

i. First, we update \( \kappa \). Denote the design matrix in the equation for indirect utilities by \( W \). In matrix notation, we have:

\[
v = W_\kappa + \epsilon, \quad \epsilon \sim N(0, I).
\]

We have a normal conjugate prior for \( \kappa \), with mean \( \mu_0^\kappa \) and covariance matrix \( \Sigma_0^\kappa \). The posterior distribution of \( \kappa \) conditional on \( v \) and \( W \) is:

\[
\kappa(v, W) \sim N(\bar{\mu}_\kappa, \bar{\Sigma}_\kappa),
\]
where the posterior mean and covariance matrix are given by:
\[
\tilde{\mu}_\kappa = \left( \frac{W'W}{\sigma^2_\varepsilon} + \left( \Sigma^0_\kappa \right)^{-1} \right)^{-1} \left( \left( \Sigma^0_\kappa \right)^{-1} \mu^0_\kappa + W'\kappa \right),
\]
\[
\tilde{\Sigma}_\theta = \left( \frac{W'W}{\sigma^2_\varepsilon} + \left( \Sigma^0_\kappa \right)^{-1} \right)^{-1} = \left( W'W + \left( \Sigma^0_\kappa \right)^{-1} \right)^{-1}.
\]

We then set \( \kappa^{k+1} \) by drawing from this posterior distribution.

A. Next, we will update \( \psi \). Denote the design matrix in the equation for firm profitability by \( Q \). In matrix notation, we have:
\[
\pi = Q\psi + \omega, \quad \omega \sim N(0, I).
\]

We have a normal prior for \( \psi \), with mean \( \mu^0_\psi \) and covariance matrix \( \Sigma^0_\psi \), so the posterior distribution of \( \theta_\pi \) conditional on \( \pi \) and \( X_\pi \) is:
\[
\psi | (\pi, Q) \sim N(\tilde{\mu}_\psi, \tilde{\Sigma}_\psi),
\]
with posterior mean and covariance matrices given by:
\[
\tilde{\mu}_\psi = \left( \frac{Q'Q}{\sigma^2_\omega} + \left( \Sigma^0_\psi \right)^{-1} \right)^{-1} \left( \left( \Sigma^0_\psi \right)^{-1} \mu^0_\psi + \frac{Q'\psi}{\sigma^2_\omega} \right),
\]
\[
\tilde{\Sigma}_\psi = \left( \frac{Q'Q}{\sigma^2_\omega} + \left( \Sigma^0_\psi \right)^{-1} \right)^{-1} = \left( Q'Q + \left( \Sigma^0_\psi \right)^{-1} \right)^{-1}.
\]

We then set \( \psi^{k+1} \) by drawing from this posterior distribution.
G Simulation Details

G.1 Assumptions for Counterfactual Simulations

To simulate what happens under various counterfactuals, we need to make several assumptions, which I discuss below.

Assumption C1. Decisions made by nursing homes to increase or reduce capacity do not change in the counterfactuals.

Assumption C2. Entry and exit decisions by nursing homes do not change in the counterfactuals.

Assumption C3. Nursing homes’ quality of care does not change with temporary fluctuations in occupancy.

Assumption C4. Residents’ most preferred nursing home among those willing to accept them in the counterfactual is preferable to the outside option of not going to a nursing home.

Assumption C5. Nursing homes’ discharge behavior does not change in the counterfactual.

Assumptions C1, C2, and C3 are necessary because modeling nursing homes’ capacity choices, entry and exit decisions, and the way in which quality varies with occupancy fluctuations is out of the scope of this paper. Assumptions C1 and C2 will be violated, for instance, if high-quality nursing homes add beds in response to greater demand, or if low-quality nursing homes exit the market due to insufficient demand (over the counterfactual period). To increase the plausibility of Assumptions C1 and C2, I restrict my counterfactual simulations to the single year of 2009, given that the myriad regulations make it more difficult to make large adjustments to capacity, and exits in any given year is a relatively rare event. Indeed, only 0.8 and 0.4 percent of nursing homes entered and exited the market, respectively, in 2009, and the average percent change in number of beds reported by nursing homes from 2008–2009 is only 1.1 percent.

The main threat to Assumption C3 is that nursing homes that are experiencing a temporary spike in occupancy may provide poorer care during this period (e.g., because they are short-staffed). To test this hypothesis, in Appendix Figure A.16 I show a binscatter of resident outcomes against my occupancy measure, controlling for resident characteristics and nursing home fixed effects. If care provided by nursing homes deteriorates when occupancies are temporarily elevated, we would expect a negative relationship between resident outcomes and occupancy. Instead, Appendix Figure A.16 shows the lack of a clear relationship between outcomes and occupancy, which provides support for Assumption C3.
Assumption C4 is required to ensure that residents do not choose the outside option (of not going to any nursing home) if no nursing home at least as desirable as her chosen nursing home is available to her in the counterfactual. This assumption cannot be tested directly because we only have data on admitted residents (and thus cannot estimate the relative value of the outside option). Nonetheless, several qualitative facts support the assumption that these residents will still prefer going to a nursing home in these counterfactuals. First, nursing home residents discharged from an acute care hospital (which comprise the majority of my sample) typically require some rehabilitative support before they are discharged to the community, and nursing homes provide most of such rehabilitative care. Second, long-stay residents are often admitted to nursing homes when most other options are exhausted, considering that living in a nursing home is typically considered an unattractive option. Third, most residents end up in a higher-quality nursing home in the counterfactuals that I consider, so the possibility that they would prefer the outside option is relatively unlikely.

Finally, Assumption C5 is required because modeling nursing homes’ discharge decisions is outside the scope of this paper. This condition may be violated, for instance, if nursing homes expedite discharges of their existing residents when they are close to capacity and a more desirable resident applies. Nonetheless, discharging residents on short order is presumably more difficult than it is for a nursing home to reject (or dissuade) an applicant, and thus there is reason to believe that selective admissions are more important than selective discharges as a means for nursing homes to manage their occupancy level. Moreover, nursing homes face legal liabilities if they try to forcefully evict residents who are not ready to be discharged (§483.10, §483.21).

G.2 Background on the Cause-Specific Hazard Model

In my modeling of nursing home exits, there are two competing risks: death and discharge. In particular, if a resident dies in the nursing home, we do not know when she would have been discharged if she had remained alive; and similarly, if a resident is discharged, we do not know when she would have died if she had stayed in the nursing home instead.

---

139 Long-term care hospitals also admit patients from acute care hospitals, but these are typically more clinically intensive cases.
140 For example, more than half of US adults responded that they would rather die than live in a nursing home, according to a 2017 Harris poll and a 2019 poll by the Nationwide Retirement Institute.
141 There is a lengthy list of steps that nursing homes must follow in the discharge planning process (§483.21).
142 Valid reasons for eviction are if the needs of the resident are greater than the nursing home can provide, refusal to pay for nursing home care in spite of “reasonable and appropriate notice” (pending Medicaid applications do not justify eviction), nursing home care no longer being necessary for the resident, the resident’s presence jeopardizing the health or safety of other residents, and nursing home closure.
The presence of competing risks means that standard tools for analyzing survival data may not suffice. For example, the Kaplan–Meier estimator is commonly used to estimate the survival function \( S(t) \) nonparametrically, and is defined by:

\[
\hat{S}(t) = \prod_{i, t_i \leq t} \left( 1 - \frac{d_i}{n_i} \right),
\]

where \( t_i \) is the event time for individual \( i \) and \( n_i \) is the number of individuals surviving at least until time \( t \). One may be tempted to simply use the Kaplan–Meier estimator separately for each event type, i.e.,

\[
\hat{S}_{\text{death}}(t) = \prod_{i, t_i \leq t} \left( 1 - \frac{d_{i, \text{death}}}{n_i} \right), \quad \hat{S}_{\text{discharge}}(t) = \prod_{i, t_i \leq t} \left( 1 - \frac{d_{i, \text{discharge}}}{n_i} \right).
\]

However, this has the undesirable property that the sum of these separate survivor function estimates will generally exceed the survivor function estimate for the composite outcome:

\[
\hat{S}_{\text{death}}(t) + \hat{S}_{\text{discharge}}(t) \geq \hat{S}_{\text{death or discharge}}(t).
\]

This is the case even if the competing risks are independent.\(^{143}\)

Competing risks models avoid this problem. In particular, a key concept in competing risks model is the cause-specific cumulative incidence function, defined as \( CIF_c(t) \equiv Pr(T \leq t, C = c) \). Cumulative incidence functions will have the natural property that:

\[
CIF_{\text{death}}(t) + CIF_{\text{discharge}}(t) = CIF_{\text{death or discharge}}(t),
\]

in contrast to the Kaplan–Meier estimates in the presence of competing risks. If we introduce covariates \( W \) into the model, we can define the cumulative incidence function analogously as being conditional on \( W \), i.e., \( CIF(t|W) \).

Two commonly used models of competing risks are the cause-specific hazard model and the Fine–Gray subdistribution hazard model. One of the main reasons practitioners prefer one over the other involves the interpretation of the estimated hazard ratios: a greater ratio for a particular covariate’s coefficients for cause \( c \) can be interpreted as higher values of the covariate being associated with greater

\(^{143}\)It may be difficult to think of independence of death and discharge in the present context. An example that illustrates this more clearly is a case with two totally unrelated events, e.g. time that I have my first coffee in the morning today and the time that a particular rabbit halfway across the world first ventures out of its rabbit hole today. Even in this case, the above property holds.
hazard rates for cause $c$ in the former model, and with greater cumulative incidence for cause $c$, i.e., $CIF_c(t|W)$, in the latter. It turns out that these two interpretations are not equivalent, which may be surprising, given that one might have expected the effect a covariate has on hazard rates and cumulative incidence should always have the same sign. However, while this is true for survival analysis with a single type of event, this is not the case when there are competing risks, as I will explain below.

Given that the primary purpose of my model is for prediction rather than interpretation purposes, it does not matter much which model I pick, so I use the cause-specific hazard model. This method involves estimating the cause-specific hazard functions, which are defined as:

$$h_c(t|W) = \lim_{\Delta t \to 0} \frac{Pr(t \leq T < t + \Delta t, C = c|T \geq t, W)}{\Delta t}, (11)$$

for each cause $c$. I model these functions semi-parametrically:

$$h_c(t|W) = h_{c,0}(t) exp(W' \beta_{c,haz}), (12)$$

letting the cause-specific baseline hazard $h_{c,0}(t)$ be non-parametric. Defining the cumulative hazard function as:

$$H_c(t|W) = \int_0^t h_c(u|W) du,$$

we can show that the cumulative incidence function for cause $c$ is given by:

$$CIF_c(t|W) = \int_0^t S(u|W) h_c(t|W) du,$$

where $S(t|W) = exp(- \sum_c H_c(t|W))$ is the survivor function for the composite event.

This formula also reveals why the sign of the hazard ratio $\beta_{c,haz}^k$ associated with the $k$th covariate in $W$ alone does not reveal whether increasing the value of this covariate would increase or decrease $CIF_c(t|W)$. This is because the $CIF_c(t|W)$ depends on the survivor function $S(u|W)$, which is itself a function of the hazard functions for other causes. Hence, $CIF_c(t|W)$ depends on the $k$th covariate in $W$ not only through $\beta_{c,haz}^k$ but also through the hazard ratios associated with this covariate in the other cause specific hazards, $\beta_{c',haz}^k$ (which of course may have a different sign).

We can estimate the parameters $\beta_{c,haz}^k$ for the cause-specific hazard functions by maximizing the
modified partial likelihood for each cause \(c\), which is given by:

\[
L(\beta_k^{c,haz}) = \prod_i \left( \frac{\exp(W_i' \beta_k^{c,haz})}{\sum_{i' \in R_i} \exp(W_{i'}' \beta_k^{c,haz})} \right)^{I[C_i = c]},
\]

where \(I[C_i = c]\) is an indicator equal to one if and only if resident \(i\) exits a nursing home due to cause \(c\) before the end of the sample period, \(W_i\) is the vector of covariates associated with resident \(i\) and the nursing home she went to, and \(R_i\) contains residents who do not exit their nursing home and are not censored before the minimum of resident \(i\)'s censoring and exit times.

### G.3 Simulating Exit Times from Cause-Specific Hazard Model

I simulate event times for cause \(c\) using the formula from Bender, Augustin, and Blettner (2015):

\[
T_{ij}^c = H_{c,0}^{-1}(-\log(U_{ij}^c)\exp(-W_{ij}' \beta_k^{c,haz})), \tag{13}
\]

where \(U_{ij}^c\) is a uniformly distributed random variable, and \(c \in \{\text{death, discharge}\}\).

A practical difficulty in implementing this formula is that \(H_{c,0}^{-1}(\cdot)\) is non-parametric in my model, and thus there is no closed form formula for \(H_{c,0}^{-1}(\cdot)\). Hence, I approximate \(H_{c,0}^{-1}(\cdot)\) as follows. Examining the estimates of the cumulative baseline hazard functions \(H_{c,0}(\cdot)\) in Appendix Figure [A.10], we observe that they are approximately of the form \(H_{c,0}(t) \approx \hat{H}_{c,0}(t) = a^c \cdot \log(1 + b^c t)\). I use this approximation, and estimate the values of \((a^c, b^c)\) using nonlinear least squares. I then use the inverse of this function as my approximation for \(H_{c,0}^{-1}(\cdot)\), i.e., \(\hat{H}_{c,0}^{-1}(x) = \frac{1}{b^c} \cdot \left[ \exp \left( \frac{x}{a^c} \right) - 1 \right] \).

As mentioned briefly in the main text, I consider various assumptions about whether \(U_{ij}^c\) is correlated with \(U_{ij'}^c\), in my simulations. Exactly how I implement this is described in the next subsection.

### G.4 Simulation Algorithm

In the counterfactuals, I consider what happens when we either change residents' demand for quality from \(\kappa_i^{\text{quality}}\) to \(\tilde{\kappa}_i^{\text{quality}}\) via a targeted information intervention (where \(\kappa_i^{\text{quality}}\) is resident \(i\)'s demand for quality taking into account her characteristics, from the structural demand model with heterogeneous preferences), or change nursing homes’ quality from \(\alpha\) to \(\tilde{\alpha}\) (which could be due to minimum standard mandates, a pay-for-performance scheme, or endogenous nursing home responses to changes in demand).
Given that residents’ choices (and outcomes) will generally differ from their original ones in the counterfactuals, so will my measure of nursing home occupancy, which I denote by \( \tilde{occ}_{ij} \) in the counterfactual. Finally, for residents’ and nursing homes’ preferences, I use the final \( (K) \)th draw of the latent variables and preference parameters from the Gibbs sampler. This is because the values of \((\epsilon_{ij}, \omega_{ij})\) in the Gibbs sampler for structural demand estimation reveals some “private” information about residents and nursing homes. For example, if a high-quality nursing home \( j \) is available to resident \( i \) who happens to have relatively strong preferences for quality and yet \( i \) chooses a different nursing home that seems worse on observables (e.g., lower quality and further away), this suggests that \( i \) dislikes \( j \) for certain reasons not captured by observables (which I call idiosyncratic tastes). By using the values of \((\epsilon_{ij}, \omega_{ij})\) from the Gibbs sampler (rather than setting them to zero or just drawing them i.i.d. without regard to matching outcomes), I am able to retain these idiosyncratic tastes in the simulations.

To model the short-run effects of an information intervention that eliminates information frictions, I take the MRS estimate of quality with respect to distance of 1.8 from Chandra, Finkelstein, Sacarny, and Syverson (2017) as a benchmark. Given that the outcome measure for this MRS estimate is 30-day mortality instead of 90-day mortality, I take a conservative approach and divide it by three, and assume that all residents’ demand for quality is such that the MRS with respect to distance is 0.6. I model the minimum standard mandate by changing the quality \( \alpha_j \) of nursing homes with quality below the 10th percentile to be equal to the 10th percentile of quality, and the pay-for-performance scheme by increasing \( \alpha \) by five percent of the baseline mortality rate based on one of the more optimistic estimates of the effect of pay-for-performance schemes in Polsky, Konetzka, and Werner (2013).144

As a robustness check, I additionally consider a version of the information intervention that assumes idiosyncratic preference shocks are also part of the information frictions and thus sets \( \epsilon_{ij} = 0 \) in addition to changing residents’ preferences for quality (although I will not introduce additional notation for this in the description below). Finally, the long-run effect of the information intervention involves changing nursing homes’ quality in a way that depends on the counterfactual changes in preferences, and exactly how I do this is described in Appendix Section II.

Denote the occupancy of nursing home \( j \) at time \( t \) (in levels) by \( o_{jt} \), and define the flow at nursing \( j \) and time \( t \) by \( flow_{jt} = o_{jt} - o_{j,t-1} \), so we have \( o_{jt} = o_{j,t-1} + flow_{jt} \). Let \( t = 1 \) be the first day of the sample, and \( t = T^{sim} \) be the last day. We can order residents who are admitted to a nursing home

---

144In particular, Polsky, Konetzka, and Werner (2013) examine the effect of Medicaid pay-for-performance schemes in a number of states, and generally find inconsistent results for most states (e.g., the program seemed to improve one resident outcome but not another). The exception is Georgia, where the scheme seemed to improve several resident outcomes by roughly five percent.
on day $t$ arbitrarily, from $i_t = 1, \ldots, I_t$. Also, let $T_{i_t}$ by the death or discharge date of resident $i_t$. The simulation proceeds as follows:

1. Initialize the counterfactual flows to be the same as the original flows: $\tilde{f}_{\text{Flow}}_{jt} = f_{\text{Flow}}_{jt}$.

2. For $t = 1, \ldots, T_{\text{sim}}$:

(a) For resident $i_t = 1, \ldots, I_t$:

i. Simulate residents’ and nursing homes’ counterfactual preferences, as given by:

$$\tilde{v}_{ij} = v_{ij}^{(K)} - \kappa_i^{\text{quality}} \alpha_j + \check{\kappa}_i^{\text{quality}} \check{\alpha}_j,$$
$$\tilde{\pi}_{ij} = \pi_{ij}^{(K)} + \psi^{\text{occ}} (\check{\omega}_{ij} - \omega_{ij}).$$

ii. Find the nursing home the resident is assigned to given these counterfactual values of the latent variables, which I denote by $\tilde{\mu}(i_t)$.\footnote{I drop the small number of residents who are ineligible for any nursing home in the counterfactual.} This is given by:

$$\tilde{\mu}(i_t) = \argmax\{ j : \pi_{ij} \geq 0, \text{ and } \tilde{v} \geq v_{ij'} \forall j' \text{ s.t. } \pi_{ij'} \geq 0 \}.$$

iii. If $\tilde{\mu}(i_t) = \mu(i_t)$, and $\alpha_{\mu(i_t)} = \check{\alpha}_{\mu(i_t)}$, set the counterfactual death or discharge date $\tilde{T}_{i_t}$ for resident $i_t$ to her original one $\tilde{T}_{i_t} = T_{i_t}$.

iv. If $\tilde{\mu}(i_t) \neq \mu(i_t)$, or $\tilde{\mu}(i_t) = \mu(i_t)$ but $\alpha_{\mu(i_t)} = \check{\alpha}_{\mu(i_t)}$, and we are assuming that unobserved risks are independent:

A. Simulate her event time $\tilde{T}_{i_t, \tilde{\mu}(i_t)}$ for cause $c \in \{\text{death, discharge}\}$ by drawing $U_{i_t, \tilde{\mu}(i_t)}$ from the standard uniform distribution and calculating:

$$\tilde{T}_{i_t, \tilde{\mu}(i_t)} = \hat{H}_{c,0}^{-1} \left( -\log(U_{i_t, \tilde{\mu}(i_t)}) \exp(-W' \beta_{c, \text{haz}}) \right).$$

Set $\tilde{T}_{i_t} = \lfloor \min\{ \tilde{T}_{\text{death}}, \tilde{T}_{\text{discharge}} \} \rfloor$.

v. If $\tilde{\mu}(i_t) \neq \mu(i_t)$, or $\tilde{\mu}(i_t) = \mu(i_t)$ but $\alpha_{\mu(i_t)} = \check{\alpha}_{\mu(i_t)}$ and we are assuming that unobserved risks are perfectly correlated:

A. If she exits the nursing home due to cause $c$ at event time $T_{i_t, \mu(i_t)}$, then compute her unobserved risk for cause $c$ by inverting the equation $T_{i_t, \mu(i_t)} = \hat{H}_{c,0}^{-1} \left( -\log(U_{i_t}) \exp(-W' \beta_{c, \text{haz}}) \right)$,
i.e., compute:

\[ U^c_i = \exp \left( -(\hat{H}_{c,0}(T_{i,\mu(i)})\exp(W'\beta_{c,haz}) \right). \]

Next, because we know that \( T_{c,i,\mu(i)}' \geq T_{c,i,\mu(i)} \), it must be the case that:

\[ U'^c_i \leq g(U^c_i, W, \beta_{c,haz}, \beta_{c',haz}) \equiv \exp \left( -\hat{H}_{c,0}^{-1} \left( -\log(U^c_i)\exp(-W'\beta_{c,haz}) \right) \right) \exp(W'\beta_{c',haz}). \]

Thus, I draw \( U'^c_i \) from a uniform distribution on \([0, g(U^c_i, W, \beta_{c,haz}, \beta_{c',haz})] \).

B. If she does not exit the nursing home due to either cause before the end of the simulation period \( T_{sim} \), we know that \( \min \left\{ T_{\text{death}}(i), T_{\text{discharge}}(i) \right\} > T_{\text{sim}} - t \). This implies that:

\[ U^c_i \leq h_c(W_{ij}, \beta_{c,haz}, T_{\text{sim}} - t) \equiv \exp \left( -\hat{H}_{c,0}(T_{\text{sim}} - t)\exp(W'\beta_{c,haz}) \right), \]

for \( c \in \{\text{death, discharge}\} \). So, I draw \( U^c_i \) from uniform distributions on \([0, h_c(W_{ij}, \beta_{c,haz}, T_{\text{sim}} - t)] \) for \( c \in \{\text{death, discharge}\} \).

C. Compute

\[ \hat{T}^c_i = \hat{H}_{c,0}^{-1} \left( -\log(U^c_i)\exp(-\hat{W}'\beta_{c,haz}) \right), \]

for \( c \in \{\text{death, discharge}\} \), and set \( \hat{T}_i = \lceil \min(\hat{T}_{\text{death}}, \hat{T}_{\text{discharge}}) \rceil \).

vi. Update the flow at her counterfactual and original nursing homes and record her outcomes.

A. Add one to \( f_{\text{flow}}_{\mu(i),t+1} \) because the counterfactual nursing home admitted one more resident at this time, and subtract one from \( f_{\text{flow}}_{\mu(i),t+1} \) because her original nursing home admitted one less resident in the counterfactual.

B. If her original outcome was not censored (i.e., \( T_{it} < T_{\text{sim}} \)), add one to \( f_{\text{flow}}_{\mu(i),t+T_{it}+1} \) because her original nursing home had one less discharge (or death) on her original discharge date (given that she is no longer assigned to that nursing home in the counterfactual).

C. If \( \hat{T}_{\text{death}} < \hat{T}_{\text{discharge}} \) and \( \hat{T}_{\text{death}} \leq T_{\text{sim}} \), record her counterfactual outcome as \( \text{death} \).
D. If $\hat{T}_{\text{death}} \geq \hat{T}_{\text{discharge}}$ and $\hat{T}_{\text{discharge}} \leq T^{\text{sim}}$, record her counterfactual outcome as discharge.

E. If $\hat{T}_{it} > T^{\text{sim}}$, record her counterfactual outcome as censored.

F. If her counterfactual outcome is not censored, update the flow at her counterfactual nursing home due to her death or discharge, i.e., subtract one from $\hat{\mu}_{it} + \hat{T}_{it} + 1$ because she exits her counterfactual nursing home at $t + \hat{T}_{it}$.

(b) Update the occupancy measure for nursing homes at time $t + 1$, i.e., set:

$$\hat{o}_{jt+1} = \hat{o}_{jt} + \hat{\ell}_{jt+1},$$

and compute the counterfactual occupancy measure using the formula:

$$\hat{occ}_{jt+1} = \left( \frac{1}{t} \sum_{s=t-6}^{t} \log(\hat{o}_{js}) \right) - \hat{occ}_{jt,m(t)},$$

where $\hat{occ}_{jt,m(t)}$ is the average value of $\log(\hat{o}_{jt})$ in the month of $t$.

H \quad \textbf{Simple Model of an Information Intervention’s Mechanisms}

H.1 \quad \textbf{Model}

I first consider a base case with no spillovers, i.e., ignoring the effects that changes in a resident’s nursing home choice may have on the choice set constraints of other residents. I denote residents’ health by $Y_i$ and assume that is affected by her baseline health $H^0_i$ and the quality of the nursing home she goes to $\alpha_i$. Her nursing home quality $\alpha_i$ depends on how informed she is $I_i$, which in turn depends on how informed she was pre-intervention $I^0_i$ and whether she is treated by the information intervention $D^i_{\text{info}}$. Finally, her baseline health is correlated with the amount of information she has pre-intervention. Under this setup, her potential outcome can be written as:

$$Y \left( H(I^0_i), Q_i(I(I^0_i, D^i_{\text{info}})) \right),$$

where $\alpha_i = Q_i(I(I^0_i, D^i_{\text{info}}))$.

Additionally, I make the following intuitive assumptions:

\textbf{Assumption M1.} Outcomes improve as $H^0_i$ and $\alpha_i$ increase, $Y_H > 0$ and $Y_\alpha > 0$. 

74
Assumption M2. Better information for a resident results in better nursing home choice, \( Q_i' \geq 0 \), but at a diminishing rate, \( Q_i'' \leq 0 \).

Assumption M3. The more informed the resident is initially the more informed she will be in the end, \( \partial I / \partial I^0 > 0 \), the effect of initial information on her final information is smaller under the information intervention, \( \partial I(I^0, 1) / \partial I^0 \leq \partial I(I^0, 0) / \partial I^0 \), that being treated by the information intervention makes her more informed, \( I(I^0, 1) - I(I^0, 0) > 0 \), and finally that the information set of a resident who is not treated is equal to her initial information \( I(I^0, 0) = I^0 \).

Assumption M4. Outcomes are weakly concave in quality, \( Y_{\alpha \alpha} \leq 0 \).\(^{146}\)

Assumption M5. Quality matters more for the outcomes of sicker residents, \( Y_{H_\alpha} \leq 0 \).

Assumption M6. Baseline health is positively related with baseline information, \( H'(I^0) > 0 \).

Proposition 1. Under Assumptions (i)–(vi), the treatment effect is positive and decreasing in baseline information.

Proof. The treatment effect for resident \( i \) is:

\[
\tau_i = \tau(I^0_i) = Y(H(I^0_i), Q(I(I^0_i, 1))) - Y(H(I^0_i), Q(I(I^0_i, 0))).
\]

The fact that \( \tau_i > 0 \) is a straightforward consequence of assumptions (i) and (ii). Taking the derivative with respect to baseline information \( I^0_i \), I obtain:

\[
\tau'(I^0_i) = \left( \frac{\partial Y(H(I^0_i), Q(I(I^0_i, 1)))}{\partial H} - \frac{\partial Y(H(I^0_i), Q(I(I^0_i, 0)))}{\partial H} \right) \cdot H'(I^0_i)
\]

\[
+ \left( \frac{\partial Y(H, Q(I(I^0_i, 1)))}{\partial Q} \cdot Q'(I(I^0_i, 1)) \cdot \frac{\partial I(I^0_i, 1)}{\partial I^0} - \frac{\partial Y(H, Q(I(I^0_i, 0)))}{\partial Q} \cdot Q'(I(I^0_i, 0)) \cdot \frac{\partial I(I^0_i, 0)}{\partial I^0} \right).
\]

\( ^{146} \)According to the linear model in the paper, average outcomes should typically increase one-for-one with quality, in which case \( Y_{\alpha \alpha} = 0 \). I allow for concavity here to accommodate a potential ceiling effect (e.g., if \( Y \) represents survival probability, it cannot increase beyond 1).
in $D_i^{info}$ due to Assumptions M2 and M4, and $Q'(I_i^0, D_i^{info})$ is decreasing in $D_i^{info}$ due to Assumptions M2 and M3, this implies that the second line is also positive.

For the general case with spillovers, the quality of resident $i$’s nursing home $\alpha_i$ depends not only on her own information $I_i$ but also on other residents’ information $I_{-i}$. Other residents’ information $I_{-i}$ depend on their initial information $I_{0,-i}$ as well as whether they are treated $D_{-i}^{info}$. For simplicity, I abstract from the temporal dimension (e.g., there are no spillover effects for the first residents to receive the intervention, but spillovers tend to affect subsequent cohorts). Under this setup, the outcome for a resident $i$ under the policy $D_i^{info}$ can be written as:

$$Y_i = \left( H_i, Q_i \left( I_i, D_i^{info}, I_{-i}, D_{-i}^{info} \right) \right).$$

I make an additional assumption that captures the idea of negative spillover effects due to occupancy constraints.

**Assumption M7.** The quality of a resident’s chosen nursing home is decreasing in the information of other residents, $\partial Q_i / \partial I_j \leq 0$ for $j \neq i$, and moreover, the increase in quality due to better information is smaller when other residents are more informed, $\partial^2 Q_i / \partial I_i \partial I_j \leq 0$.

**Proposition 2.** Consider an intervention $D_i^{info}$ that weakly increases all residents’ information. Under Assumptions M1–M7, the sign of the average treatment effect is ambiguous. Moreover, treatment effects may not be increasing in residents’ initial information frictions.

**Proof.** To prove that the signs of the average treatment effect as well as $\tau'(I_i^0)$ are ambiguous, it suffices to provide examples that satisfy Assumptions M1–M7 and under which these quantities have different signs.

Consider a simple case with two residents, $i = 1, 2$. First, suppose $Y_i = log(H_i) + Q_i(I_i, I_j)$, $Q_i(I_i, I_j) = log(I_i) / I_j$, and $I_i(I_i^{0}, D_i) = I_i^0 + aD_i$ for $i \neq j$, and where $a > 0$. It is easy to show that this setup satisfies Assumptions M1–M7, and moreover, $\tau_i(I_i^0) > 0$, and $\tau_i'(I_i^0) < 0$.

Now, suppose instead that $Y_i = M H_i + H_i Q_i$, where $M$ is a large number such that $\partial Y_i / \partial H_i > 0$, and suppose $H_i(I_i^0) = -I_i^1$. Also, as before, let $Q_i(I_i, I_j) = log(I_i) / I_j$, and $I_i(I_i^{0}, D_i) = I_i^0 + aD_i$. It is straightforward that $\tau_i(I_i^0) < 0$, and if $I_i^0 < 1$ and $a$ is sufficiently small, then $\tau_i'(I_i^0) > 0$.

While the examples suffice for the proof, it is also useful to consider the general case to understand
the mechanisms. The effect of the intervention on resident \(i\)'s outcome can be written as:

\[
\tau_{D^{in_{fo},i}} = Y \left( H_i(I^0_i), Q_i \left( I_i(I^0_i), 1, I_{-i}(I^0_{-i}, D^{in_{fo},-i}_i) \right) \right) - Y \left( H_i, Q_i \left( I_i, I^0_i \right) \right).
\]

Adding and subtracting her outcome when the resident is not treated but without changing other residents' treatments, we obtain the following expression:

\[
\tau_{D^{in_{fo},i}} = \left[ Y \left( H_i(I^0_i), Q_i \left( I_i(I^0_i), 1, I_{-i}(I^0_{-i}, D^{in_{fo},-i}_i) \right) \right) - Y \left( H_i, Q_i \left( I_i, I^0_i \right) \right) \right] + \left[ Y \left( H_i(I^0_i), Q_i \left( I_i(I^0_i, I_{-i}(I^0_{-i}, D^{in_{fo},-i}_i)) \right) \right) - Y \left( H_i(I^0_i), Q_i \left( I_i(I^0_i, 1, I^0_{-i}) \right) \right) \right].
\]

The term in the first line is the standard individual level treatment effect holding all other residents' information constant, so it corresponds to the treatment effect under no spillovers, which is positive (as shown in Proposition 1). However, given that resident \(i\)'s nursing home quality decreases weakly when others are treated from assumption (vii), the spillover term in the second line is negative. It is ambiguous which of these effects dominates, and so the sign of the average treatment effect

\[
\bar{\tau}_{D^{in_{fo}}} = \int_{i \in I} \tau_{D^{in_{fo},i}} dF_i(i)
\]

is ambiguous as well.

To simplify notation, I denote \(Q_{D,i}^{D_{-i}^{in_{fo}}} \equiv Q_i \left( I_i(I^0_i, D_i), I_{-i}(I^0_{-i}, D^{in_{fo},-i}_i) \right) \). Turning now to the distributional effects, I take the derivative of a resident’s treatment effect \(\tau_{D^{in_{fo},i}}\) with respect to her initial information \(I^0_i\):

\[
\frac{d\tau_{D^{in_{fo},i}}}{dI^0_i} = \frac{d}{dI^0_i} \left[ Y \left( H_i(I^0_i), Q_i \left( I_i(I^0_i, 1, I_{-i}(I^0_{-i}, D^{in_{fo},-i}_i)) \right) \right) - Y \left( H_i(I^0_i), Q_i \left( I_i(I^0_i, 1, I^0_{-i}) \right) \right) \right] + \frac{d}{dI^0_i} \left[ Y \left( H_i(I^0_i), Q_i \left( I_i(I^0_i, I_{-i}(I^0_{-i}, D^{in_{fo},-i}_i)) \right) \right) - Y \left( H_i(I^0_i), Q_i \left( I_i(I^0_i, 1, I^0_{-i}) \right) \right) \right] \tag{15}
\]

The first line is the standard individual-level treatment effect heterogeneity as in equation (14) in the proof of Proposition 1, which is negative and hence progressive. However, the same is not true for the second line, and hence, the sign of \(d\tau_{D^{in_{fo},i}} / dI^0_i\) is ambiguous.
H.2 Empirical Decomposition

In this subsection, I describe how I decompose the treatment effects in the simulations into the various components described above.

First, to decompose the average treatment effect of an information intervention into components representing the treatment effect without spillovers, and the effect of spillovers, I run an additional simulation wherein I modify residents’ preferences as before but keep the supply side constant by ignoring the effect occupancy changes may have on nursing homes’ admissions policies. In this subsection, I refer to the simulations for the information intervention with and without updating the supply side simulations A and B, respectively. The average treatment effect without spillovers is represented by average treatment effect in simulation B, whereas the spillover effects are the difference in (counterfactual) resident outcomes under simulations A and B.

To decompose the treatment effect heterogeneity \( \tau' (I^0) \) into the reallocation and indirect tagging effects, as well as the negative spillover effects, I rely once again on simulations A and B. I regress the change in outcome on the change in quality and baseline health in simulation B and denote the coefficients by \( \beta_{\tau Q} \) and \( \beta_{\tau H} \) respectively. The indirect tagging effect is given by the expression:

\[
\tau_{\text{tag}}' (I^0_i) \equiv \left( \frac{\partial Y(H(I^0_i), Q(I(I^0_i, 1), I_{-i}^{D_{inf}^o}))}{\partial H} \right) - \left( \frac{\partial Y(H(I^0_i), Q(I(I^0_i, 0), I_{-i}^{D_{inf}^o}))}{\partial H} \right)
\]

and thus, to approximate this, I use the product of \( \beta_{\tau Q} \) by \( \beta_{Q I^0} \), where the latter term is the coefficient from a regression of baseline health on baseline information. The reallocation effect is given by:

\[
\tau_{\text{realloc}}' (I^0_i) \equiv \left( \frac{\partial Y(H(I^0_i), Q(I(I^0_i, 1), I_{-i}^{D_{inf}^o}))}{\partial Q} \right) \frac{\partial Q(I(I^0_i, 1), I_{-i}^{D_{inf}^o})}{\partial I^0_i} \left( \frac{\partial Y(H(I^0_i), Q(I(I^0_i, 0), I_{-i}^{D_{inf}^o}))}{\partial Q} \right) \frac{\partial Q(I(I^0_i, 0), I_{-i}^{D_{inf}^o})}{\partial I^0_i}
\]

and I approximate this by multiplying \( \beta_{\tau Q} \) on \( \beta_{Q I^0} \), where the second term is the coefficient from a regression of change in quality on baseline information. Finally, heterogeneity due to spillover effects, which I denote by \( \tau_{\text{spill}}' (I^0_i) \), is approximated by the coefficient of a regression of the difference in outcomes between simulations A and B on the baseline information of residents.

Next, I decompose the average treatment effect \( \bar{\tau} \) into components due to reallocation, indirect
tagging, and spillovers. To do so, I apply the fundamental theorem of calculus to the expressions derived for these effects when analyzing treatment effect heterogeneity. Therefore, we have:

\[
\hat{\tau} = \int_{I_{0\min}}^{I_{0\max}} \tau(I^0)dF(I^0)
= \int_{I_{0\min}}^{I_{0\max}} \left[ \int_{I_{0\min}}^{I_{0\max}} \tau'(I)dI \right] dF(I^0)
= \int_{I_{0\min}}^{I_{0\max}} \left[ \int_{I_{0\min}}^{I_{0\max}} \tau_{tag}'(I)dI \right] dF(I^0)
+ \int_{I_{0\min}}^{I_{0\max}} \left[ \int_{I_{0\min}}^{I_{0\max}} \tau_{realloc}'(I)dI \right] dF(I^0)
+ \int_{I_{0\min}}^{I_{0\max}} \left[ \int_{I_{0\min}}^{I_{0\max}} \tau_{spill}'(I)dI \right] dF(I^0) + \tau(I_{0\max}^0),
\]

where I also reverse the bounds of the integration so that the effect of the constant term at the end is minimized.\(^{147}\) I use the same linear approximation for the derivative of the treatment effect function as when analyzing treatment effect heterogeneity.\(^{148}\)

I Simulating Endogenous Quality Adjustments by Nursing Homes

I.1 Simple Model of Quality Choice by Nursing Homes

In this section, I describe a simple model that provides an empirical framework for simulating quality adjustments by nursing homes in response to the elimination of information frictions. Consider the model of profit maximization in Gaynor, Ho, and Town (2014):

\[
\max_{\alpha_j} \Pi_j = \bar{p}N_j - c(N_j, \alpha_j),
\]

where \(\Pi_j\) is the profit of nursing home \(j\), \(\bar{p}\) is the reimbursement rate for each resident, and \(c(\cdot, \cdot)\) is a cost function that depends on the number of residents choosing nursing home \(j\), \(N_j\), and the quality chosen by the nursing home \(\alpha_j\).\(^{149}\) The number of residents who choose nursing home \(j\) is given by

\(^{147}\)If I instead integrated from \(I_{0\min}\) up to \(I\) in the inner integral, the constant term would be \(\tau(I_{0\min})\), which is larger than \(\tau\), making the three components difficult to interpret. In contrast, the interpretation is easier if the constant term is close to zero, which is the case with \(\tau(I_{0\max})\).

\(^{148}\)One can obtain a more accurate approximation by fitting a nonlinear function to the treatment effect function and using the slope of that function. Nonetheless, the calculations here are more proof of concept, therefore I do not take this additional step.

\(^{149}\)I ignore fixed costs considering that I hold entry/exit constant in my simulations.
the demand function $N_j(\bar{p}_{\text{res}}, \alpha_j, \alpha_{-j})$ which depends on the price the resident has to pay, $\bar{p}_{\text{res}}$ (e.g. a copay) and the quality of all nursing homes, $\{\alpha_j\}_j$. In addition, I assume that the cost of quality is given by $c(N_j, \alpha_j) = \left(\frac{c_j^2}{2} \alpha_j^2 + c_j^N\right) N_j$.

For my benchmark, I make the conservative assumption that the total number of residents is fixed at $\bar{N}$, so that nursing homes are only competing on the share of residents they get. I abstract away from differences between nursing homes in their distances to potential surrounding residents, and assume that utilities are given by the sum of $\kappa_i^{\text{quality}} \alpha_j$, a nursing home fixed effect $\xi_j$, and an i.i.d. error term with a type-II extreme value distribution. Although capacity constraints are not explicitly incorporated into this simplified model, it is captured in a reduced-form manner by the fact that marginal cost of quality is increasing in the number of residents (as well as in quality). Under this setup, firm $j$’s market share is given approximately by:

$$s_j(\kappa^{\text{quality}}, \alpha) = \frac{\exp(\kappa_j^{\text{quality}} \alpha_j + \xi_j)}{\sum_l \exp(\kappa_l^{\text{quality}} \alpha_l + \xi_l)},$$

where I define the firm-specific demand parameters $\kappa_j^{\text{quality}}$ as the average demand for quality among its potential residents: $\kappa_j^{\text{quality}} \equiv \frac{\sum_i \kappa_i^{\text{quality}} \cdot \mathbb{1}[j \in J_i]}{(\sum_i \mathbb{1}[j \in J_i])}$.

Maximizing profits by choosing quality, nursing home $j$’s first-order conditions are given by:

$$\frac{\kappa^{\text{quality}} \bar{p}_j}{s_j(1 - s_j) \bar{N}} = \frac{\kappa_j^{\text{quality}} \alpha_j s_j \bar{N}}{c_j^{\text{quality}} \alpha_j s_j \bar{N}} + \kappa_j^{\text{quality}} \left(\frac{1}{2} c_j^{\text{quality}} \alpha_j^2 + c_j^N\right) s_j(1 - s_j) \bar{N},$$

where I suppressed the dependence of $s_j$ on demand for quality and the entire vector of quality choices for notational simplicity. The left-hand side of this equation corresponds to the marginal revenue from increasing quality, whereas the right-hand side is the marginal cost. The two terms in the marginal cost expression correspond to the fact that increasing quality to attract additional residents results in greater costs for the inframarginal residents, and the second term reflects the additional costs of providing care for marginal residents attracted by the increase in quality.

### I.2 Calibration Using the Introduction of Five-Star Ratings

To calibrate this model, I consider changes in demand for quality and nursing homes’ quality before and after the introduction of the CMS five-star ratings system. In particular, with estimates of quality
and quality demand before and after the introduction of star ratings, I obtain:

\[
(\gamma_j)(\Delta^*\kappa_j^{quality}) \approx \Delta^*\alpha_j,
\]

where \(\gamma_j \equiv (\bar{p} - c^N)/c_j^{quality}\) is a supply-side parameter that governs the “profitability” of providing quality, and I use the approximation that \(1 - s_j^{pre*} \approx 1 - s_j^{post*} \approx 1\) given that \(s_j\) is small for most nursing homes.\(^{150}\) Alternatively, the assumption that \(1 - s_j \approx 1\) can be interpreted as nursing homes not internalizing the effects of its own quality choice on nursing homes. The notation \(\Delta^*\) and (later on) \(\tilde{\Delta}\) denotes the difference before and after the star ratings, and before and after the counterfactual information intervention, respectively. I already estimated the change in demand for quality \(\Delta^*\kappa_j^{quality}\) in the structural demand estimation, and I will describe my estimation of the change in quality \(\Delta^*\alpha_j\) in the next subsection.

After computing \(\gamma_j\) using the changes induced by the star ratings system (specifically, using estimated values of \(\Delta^*\kappa_j^{quality}\) and \(\Delta^*\alpha_j\)), I can then simulate nursing homes’ quality choice \(\tilde{\alpha}_j\) for the counterfactual demand for quality \(\tilde{\kappa}_j^{quality}\) using the same equation, by replacing the pre- and post-star ratings demand and quality with the observed and counterfactual values:

\[
(\gamma_j)(\tilde{\Delta}\kappa_j^{quality}) \approx \tilde{\Delta}\alpha_j,
\]

the difference being that \(\gamma_j\) is now known, and the unknowns are the counterfactual quality choices by nursing homes \(\tilde{\alpha}_j\). Hence, the counterfactual quality choices by firms are given by:

\[
\tilde{\alpha}_j \approx \alpha_j + \gamma_j\tilde{\Delta}\kappa_j,
\]

or equivalently,

\[
\tilde{\alpha}_j \approx \alpha_j + \left(\frac{\tilde{\Delta}\kappa_j}{\Delta^*\kappa_j}\right)\Delta^*\alpha_j.
\]

Due to power issues, I conduct this calibration exercise using aggregate changes in quality and demand rather than nursing home-specific changes.

\(^{150}\)I also make use of the approximations that \(\alpha^2 \approx 0\) and \(\Delta\alpha^2 \approx 0\) in the first order conditions when obtaining this formula because quantitatively these second-order terms are much smaller than other terms in the formula. To the extent that these terms are non-negligible, my estimate of the profitability of quality, \(\gamma_j\) is an underestimate.
I.3 Estimating Change in Quality Due to the Introduction of the Five-Star Ratings

I estimate quality using two-year bins for the time period before and after the introduction of star ratings, and excluding 2008.\footnote{I exclude the year 2008 because although the introduction of star ratings comes at the end of the year, I find evidence of anticipatory responses by nursing homes in 2008. In particular, when I estimate time-specific quality separately for every single year, I observe a jump in quality in 2008. However, this could also be due to noise given the small samples, which is why I chose to use two-year bins instead.} I then estimate $\Delta^{\text{5-star}} \alpha$ using a method akin to a regression discontinuity (RD) design, with these time-specific quality measures as the dependent variable and time as the running variable, using a linear time trend. Figure A.19b shows the RD plot, where I choose the left bandwidth using a conservative approach by picking the year that will result in the smallest treatment effect estimate.\footnote{Conventional bandwidth selection methods (Imbens and Kalyanaraman, 2012; Calonico, Cattaneo, and Titianik, 2014) do not work because the running variable (year) has discrete support. There are methods for RD designs with discrete running variables (Lee and Card, 2008; Kolesar and Rothe, 2018), but there are too few values in the support of the running variable to use these reliably.} The RD plot confirms that there does seem to be a discrete jump in quality around the time the star ratings were introduced.

Another way to visualize this is by plotting the distribution of quality for the different time periods. The kernel density plot in Figure A.19a shows that while the distributions for quality in two-year windows seem quite similar in the years leading up to the introduction of the star ratings, the distribution of quality from 2009–2010 is shifted substantially to the right.

The fact that quality seems to change in this latter period may raise questions about the usefulness of my quality estimates (which nets out year effects). However, Appendix Figure A.20 shows that quality estimates based on the few years before and after 2008 are highly correlated, which (importantly for my demand estimation) suggests that rankings of nursing home quality are relatively stable over time.

I.4 Alternative Functional Forms

For robustness, I consider other functional forms for demand. In particular, I consider cost functions that are linear in quality $\left(c^\text{qual}_j \cdot \alpha_j + c^N_j N_j\right)$ or where cost of quality is not multiplicative in quantity $\frac{c^\text{quality}_j}{2} \alpha^2_j + c^N_j N_j$, as well as ad hoc demand functions (i.e., demand functions that are not microfounded) that are either a linear or logarithmic function of quality.

Using the first-order conditions, similar to my main specification, I can derive a parameter that controls the profitability of providing quality, which I calibrate using the changes induced by the change in quality to account for the introduction of star ratings.
introduction of the five-star rating system. This parameter then determines nursing homes’ quality choices in the counterfactual for a given change in residents' demand for quality. Expressions and derivations of these formulas are available from the author upon request.

Appendix Table A.18 shows the results of the simulations for the main specification and for the other specifications. The results indicate that the exact magnitude of the long-run effect varies depending on the functional forms of the demand and cost functions, but that in all cases they are several times larger than the short run effect. Specifically, based on the assumed demand for quality in the long-run effect simulations (which comes from the low end of the demand estimates in the literature), the short-run effect estimate is an 8 percent reduction in mortality, whereas all of the long-run effect estimates are at least a 26 percent reduction in mortality.

J Variable Selection Procedure Motivated by Double Machine Learning

The validity of my quality estimates relies on the selection on observables assumption, so it helps that my data contains more than 500 baseline resident characteristics. However, including the full set of controls may raise concerns of “overfitting” — especially because many of the control variables correspond to medical conditions that are quite rare and because the sample sizes for some nursing homes are relatively small.\footnote{An extreme example of overfitting in this context would be if some of the controls (that one need not control for to obtain consistent estimates of quality) end up being perfectly collinear with some of the nursing home choice dummies so that it becomes impossible to estimate quality for these nursing homes.} Hence, I use a variable selection method motivated by the post-double-selection procedure used by Belloni, Chernozhukov, and Hansen (2014).

The standard post-double-selection procedure involves running Lasso regressions of both the outcome $Y_i$ and the endogenous variable $D_i$ on the full set of controls $X_i$, then taking the union of the variables selected in these two Lasso regressions for the final estimation of the treatment effect.\footnote{Although the omitted variables bias is zero if the omitted variable is either unrelated to the endogenous variable or unrelated to the outcome, taking the union of the selected variables makes the procedure robust to “modest” errors in the variables selection process.} The complication with applying post-double-selection in the present setting is that the vector of treatment variables is relatively high-dimensional; specifically, it contains $J - 1 > 800$ nursing home choice dummies, in contrast to most settings where post-double-selection is used that only involve a single treatment variable (or, at most, a few). Running Lasso regressions of each nursing home choice dummy $D_{ij}$ on the full set of controls is computationally infeasible.
Hence, instead of running Lasso regressions of each nursing home choice dummy (and the outcome) on the full set of controls, I create a linear index summarizing the “type” of nursing home a resident chooses based on the leave-out mean of the outcome variable, i.e., \( \bar{Y}_{\mu(i),-i} = \frac{1}{N_{\mu(i)}-1} \sum_{i' \neq i} Y_{i'} D_{i',\mu(i)} \), where \( \mu(i) \) denotes the nursing home chosen by resident \( i \), and \( N_j \equiv \sum_{i=1}^{N} D_{ij} \). We can motivate the use of \( \bar{Y}_{\mu(i),-i} \) in the post-double-selection procedure based on a correlated effects approach, which involves modeling the causal effects \( \beta_j \) as a function of observables (resident outcomes in this case), and I also use the leave-one-out mean to avoid a mechanical relationship between \( Y_i \) and \( \bar{Y}_{\mu(i),-i} \). I then apply the standard post-double-selection procedure but using \( \bar{Y}_{\mu(i),-i} \) in place of \( D_i \). Finally, I take the union of the variables selected by the two Lasso regressions (of \( Y_i \) on \( X_i \) and \( \bar{Y}_{\mu(i),-i} \) on \( X_i \)) and estimate an empirical Bayes model using this set of controls.

**K Details on Empirical Bayes Implementation**

Recall that we are estimating the model:

\[
Y_i = \mu_1 + \sum_{j=2}^{J} \beta_j D_{ij} + X_i' \gamma + \epsilon_i,
\]

where the parameters of interest are \( \beta_j \). We can rewrite this in a more familiar form for panel data:

\[
Y_{ji'} = \mu_0 + X_{ji'}' \gamma + \beta_j + \epsilon_{ji'},
\]

so that nursing homes (indexed by \( j \)) correspond to the members of the (unbalanced) panel, and residents in nursing homes (indexed by \( i' \)) are akin to the time dimension in panel data.

Adopting the Bayesian perspective, I treat the parameters \( \beta_j \) as random and as being drawn from a prior distribution \( N(0, \sigma^2_\beta) \). Under the approximation that \( \epsilon_{ji'} \) are drawn from a mean zero normal distribution with variance \( \sigma^2_\epsilon \), we can derive the likelihood function for MLE. In particular, the log-
likelihood for the \(j\)th nursing home is given by:

\[
l_j = \frac{1}{2} \left( \frac{1}{\sigma^2_X} \sum_{i'=1}^{N_j} (Y_{ji'} - \mu_1 - X_{ji'}'\gamma)^2 \right.
- \frac{\sigma^2}{N_j \sigma^2_\beta + \sigma^2_t} \left( \sum_{i'=1}^{N_j} (Y_{ji'} - \mu_1 - X_{ji'}'\gamma) \right)^2
+ \log \left( N_j \frac{\sigma^2}{\sigma^2_\beta} + 1 \right) + N_j \log(2\pi \sigma^2_t),
\]

where \(N_j\) denotes the total number of residents in nursing home \(j\) (over the sample period), and the log-likelihood is minimized over \((\mu_1, \gamma', \sigma^2_\beta, \sigma^2_\epsilon)\). Finally, the empirical Bayes estimates of nursing home quality are given by:

\[
\hat{\alpha}_j = \frac{\hat{\sigma}^2_\beta}{\hat{\sigma}^2_t + \hat{\sigma}^2_\epsilon / N_j} \left( \frac{1}{N_j} \sum_{i'=1}^{N_j} (Y_{ji'} - \hat{\mu}_1 - X_{ji'}'\hat{\gamma}) \right)
+ \frac{\hat{\sigma}^2_t / N_j}{\hat{\sigma}^2_\beta + \hat{\sigma}^2_t / N_j} \left( \frac{1}{N} \sum_{j=1}^{J} \sum_{i'=1}^{N_j} (Y_{ji'} - \hat{\mu}_1 - X_{ji'}'\hat{\gamma}) \right), \tag{16}
\]

where \(N\) is the total number of observations.\(^{155}\)

### L Simple Example of Attenuation Bias in IV Due to Positively Correlated Measurement Errors

Consider the exactly identified case for our IV estimation of the forecast coefficient \(\lambda\). Recall that the endogenous variable is the quality estimate of resident \(i\)'s chosen nursing home \(\alpha_i\), and the instrument \(Z_i\) is the quality estimate of the nursing home closest to the resident.\(^{156}\) Given that we use a finite-sample estimate \(\alpha_i = \alpha^*_i + e_\alpha\), in place of the estimate \(\alpha^*_i = \lim_{N_j(i) \to \infty} \alpha_i\) that we will obtain in large samples, it suffers from measurement error, and similarly for \(Z_i = Z^*_i + e_{Z_j(i)}\) where \(Z^*_i = \lim_{N_j \to \infty} Z_i\) since it is defined using the \(\alpha_j\)'s. I assume the measurement error is classical, so that \(e_{\alpha,j}(i) \sim \text{i.i.d.} N(0, \text{Var}(e_\alpha))\) and \(e_{Z_j(i)} \sim \text{i.i.d.} N(0, \text{Var}(e_Z))\).

I will now show that the IV coefficient for my setting can be written as \(\frac{\text{Var}(\alpha^*)}{\text{Var}(\alpha^*) + \text{Var}(e_\alpha)} \cdot \lambda\), and thus it suffers from the same attenuation bias as the OLS estimate does. First, we write the IV

\(^{155}\)This MLE can be implemented in Stata via the “xtreg, mle” command. However, as a sidenote, the postestimation command “predict, u” in Stata to recover the random effects is somewhat misleading. In particular, it yields the unshrunked estimates \(\sum_{i'=1}^{N_j} (Y_{ji'} - \hat{\mu}_1 - X_{ji'}'\hat{\gamma}) / N_j\), instead of the shrunked estimates (which confusingly is the result when one runs the same postestimation command after using a different random effects estimator “xtreg, re”). Therefore, shrinkage must be done manually after running “xtreg, mle”.

\(^{156}\)For simplicity, I omit the notation indicating that this is a leave-year-out estimate.
coefficient as the ratio of the reduced-form and first-stage coefficients: \( \frac{Cov(Y_i, Z_i)}{Cov(\alpha_i, Z_i)} \). Using the formula for classical attenuation bias, as well as the basic properties of the covariance operator, we can write the IV coefficient as \( \left( 1 + \frac{Cov(e_{\alpha,j(i)}e_{Z,j(i)})}{Cov(\alpha^*, Z^*)} \right)^{-1} \cdot \lambda \). So, we observe that there will be bias if the measurement errors in the endogenous variable and instrument are correlated.

Denote by \( p \) the probability that the resident chooses the closest nursing home to her, and let \( D_{\text{nearest}} \) be the indicator variable for this event. Then, we can write:

\[
Cov(e_{\alpha,j(i)}e_{Z,i}) = p \cdot E[e_{\alpha,j(i)}e_{Z,j(i)}|D_{\text{nearest}} = 1] + (1 - p) \cdot E[e_{\alpha,j(i)}e_{Z,j(i)}|D_{\text{nearest}} = 0]
\]

\[
= p \cdot Var(\alpha),
\]

using the fact that \( \alpha_i = Z_i \) if the resident chooses her nearest nursing home, and that \( e_{\alpha,j} \) and \( e_{\alpha,j'} \) are independent if \( j \neq j' \). Using similar computations, we find that \( Cov(\alpha_i, Z_i) = p \cdot Var(\alpha^*) \). Therefore, the IV coefficient can be written as:

\[
\left( 1 + \frac{Cov(e_{\alpha,j(i)}e_{Z,j(i)})}{Cov(\alpha^*, Z^*)} \right)^{-1} \cdot \lambda = \frac{Var(\alpha)}{Var(\alpha) + Var(e_{\alpha})} \cdot \lambda,
\]

as desired.

Nonetheless, this attenuation bias will be relatively small if the measurement error in nursing home quality is small relative to the variation in nursing home quality. This seems plausible, given that the estimated standard deviation of nursing home quality is 0.02, and the standard error of this estimate is only 0.001, as seen in Figure A.2.

### M Reduced Form Evidence of Selective Admissions

In this section, I describe my tests of selective admissions in more detail (Gandhi 2019), and also present additional test results. I test prediction 1 in section 4.1 by running the following regression at the nursing home-day level:

\[
\text{admit}_{jdt} = \gamma_0^a + \gamma_1^a \bar{e}c_{jdt} + \delta_{jdt}^a + \xi_{jdt}^a,
\]
where \( admit_{jdmt} \) is a measure of nursing home \( j \)'s admissions on day \( d \) of month \( m \) in year \( t \), \( occ_{jdmt} \) is the average log occupancy over the seven days preceding this date,\(^{157}\) and \( \delta^a_{jm} \) are nursing home-month fixed effects.\(^{158}\) A negative estimate for \( \gamma'^a_1 \) would be in line with prediction 1.

To test prediction 2, I run the following resident-level regressions:

\[
x_i = \gamma'^x_0 + \gamma'^x_1 occ_{\mu(i),d(i),m(i),t(i)} + \hat{x}_i \gamma'^x_2 + \delta^x_{\mu(i)} + \xi^x_i.
\]

for different resident characteristics \( x_i \), controlling for other resident characteristics \( \hat{x}_i \) and nursing home fixed effects \( \delta^x_{\mu(i)} \), where \( \mu(i) \) is the nursing home that resident \( i \) is admitted to, and similarly \( d(i), m(i), \) and \( t(i) \) indicate the date that resident \( i \) was admitted. Evidence that \( \gamma'^x_1 \neq 0 \) would be in line with prediction 2, and we would expect \( \gamma'^x_1 \) to be positive for characteristics considered desirable by nursing homes (e.g., those associated with higher profitability).

Table 3 and Appendix Table A.9 show regression estimates testing prediction 1. I consider different measures for admissions behavior \( admit_{jdmt} \) (number of new residents, a dummy for any new residents, and flow of residents) and the occupancy measure (lagged seven-day average occupancy, log occupancy, and occupancy percentile). Regardless of the specification, the estimate of \( \gamma'^x_1 \) remains negative and statistically significant (at the five percent significance level), consistent with prediction 1.

Figure 4, Appendix Figure A.7, and Appendix Table A.10 show results from tests of prediction 2. Figure 4 and Appendix Figure A.7 show that when nursing homes are closer to capacity, they are less likely to admit Medicaid residents, regardless of the precise measure of occupancy used, and whether one controls for other resident characteristics that may be associated with Medicaid status. This is in line with prediction 2, considering the low Medicaid reimbursement rates. Appendix Table A.10 shows similar tests for other resident characteristics, and although it is less clear a priori which of these characteristics nursing homes may find more desirable,\(^{159}\) the results show that nursing homes admit

\(^{157}\) Although I observe the date each resident’s stay begins, this often does not correspond to the actual day the nursing home makes the admission decision, which often precedes the start of each stay by several days. Therefore, I use the average over the seven days leading up to the day of admission for the occupancy measure. I also consider occupancy measures other than log occupancy, such as occupancy in levels and occupancy percentile.

\(^{158}\) A practical challenge with testing these model predictions is that nursing homes may expand or contract over time (due to reasons outside of the model), and thus we must control for time-varying nursing home capacity. While nursing homes report their total number of beds, this figure is only updated annually, and there is substantial measurement error. Hence, in my analysis I only use variation from short-term fluctuations in nursing home occupancy. Related to this point, for my main sample I drop the relatively small number of nursing homes with occupancies that changed by a factor of more than two over the period of 2008–2010, in case there are underlying issues connected to these large expansions or contractions that are not accounted for by the nursing home-month fixed effects. Nonetheless, I also estimate demand for quality without dropping these nursing homes as a robustness check in panel B of Appendix Table A.13 and I still obtain a very low demand estimate.

\(^{159}\) For example, while certain health conditions make residents more costly to care for, these conditions may also be associated with higher reimbursement rates.
different types of residents during periods of high and low occupancy \((\delta^x_a \neq 0)\), which is consistent with prediction 2.

**N Simple Model of Imperfect Information About Quality**

Suppose that residents do not directly observe nursing home quality \(\alpha_j\), but only a noisy signal of it, \(\tilde{\alpha}_j = \alpha_j + e_j^\alpha\), as well as publicly available information about the nursing home \(q_j\), which includes reported staffing levels, ownership status, and number of cited deficiencies.\(^{160}\) Assume also that \(e_j^\alpha \sim \text{i.i.d. } \mathcal{N}(0, \sigma^2_{e^\alpha})\), and is independent of everything else.

Suppose that nursing home quality is drawn from a normal distribution, with a mean that is linear in publicly available information about nursing homes, specifically, \(q_j^\beta q\). In this case, residents' conditional expectation of nursing home \(j\)'s quality given the signal is:

\[
\tilde{\alpha}_j \equiv \mathbb{E}[\alpha_j | \tilde{\alpha}_j] = \sigma^2_{e^\alpha} \frac{q_j^\beta q\alpha}{\sigma^2_{e^\alpha} + \sigma^2_\alpha} \tilde{\alpha}_j + \frac{\sigma^2_\alpha}{\sigma^2_{e^\alpha} + \sigma^2_\alpha} \tilde{\alpha}.
\]  (17)

If we assume that residents' decision utility is given by \(v_{ij} = \kappa^\tilde{\alpha}_j \tilde{\alpha}_j + \kappa^{dist} dist_{ij} + \tilde{v}_{ij}\), then substituting in the expression for \(\tilde{\alpha}_j\) derived above, we obtain:

\[
v_{ij} = q_j^\beta q \left[\left(\frac{\kappa^\tilde{\alpha}_j \sigma^2_{e^\alpha}}{\sigma^2_{e^\alpha} + \sigma^2_\alpha}\right) \cdot \beta^q_\alpha \right] + \frac{\sigma^2_\alpha}{\sigma^2_{e^\alpha} + \sigma^2_\alpha} \kappa^\tilde{\alpha}_j \tilde{\alpha}_j + \kappa^{dist} dist_{ij} + \tilde{v}_{ij},
\]  (18)

where \(\tilde{v}_{ij} \equiv \kappa^\tilde{\alpha}_j e_j^\alpha + v_{ij}\) is a composite error term that is conditionally independent of the instruments given the covariates. Then, it is clear from the expression for \(\kappa^\tilde{\alpha}_j\) that if residents have positive demand for expected quality \((\kappa^\tilde{\alpha} > 0)\), each component of \(\kappa^\tilde{\alpha}_j\) should have the same sign as the corresponding component in \(\beta^q_\alpha\). Therefore, when we estimate residents' preferences as a linear function of \(q_j, \alpha_j,\) and \(dist_{ij}\), we should estimate that residents place positive weight on nursing home characteristics in \(q_j\) that positively predict quality, and negative weight on those that negatively predict quality.

The results are qualitatively similar if we allow for noise in the estimated quality measure \(\alpha_j\). It is easy to show that residents' preferences for components of \(q_j\) have the same sign as above — the only difference is that the weights that residents put on the quality signal and prior based on observables will now include a term for the variance of the estimation noise for \(\alpha_j\).

\(^{160}\)For purposes of exposition, I initially assume that estimation noise in nursing home \(j\) is negligible, at least relative to the noise in the residents' signal \(e_j^\alpha\). I will also explain at the end of this section that relaxing this assumption does not change the model's predictions.
Comparison of Nursing Home Quality Estimates with Value-Added Estimates in Other Settings

The IV results in section 3 support the validity of my estimates of nursing home value-added. Tests of the validity for value-added estimates in other healthcare settings have often found similar results (Fletcher, Horwitz, and Bradley 2014; Abaluck, Bravo, Hull, and Stac 2021). This contrasts somewhat with the mixed evidence on the reliability of value-added estimates in education settings. For instance, Rothstein (2010) shows in a falsification test that teacher value-added estimates predict their students' past achievement (a violation of an exclusion restriction for value-added models), whereas Chetty, Friedman, and Rockoff (2014) find evidence consistent with their teacher value-added estimates being unbiased. Below, I outline several institutional reasons for why estimates of nursing home value-added may be less subject to selection bias than value-added estimates in education settings.

First, nursing home residents may tend to make less informed choices relative to individuals in education settings. This is supported by the structural demand estimation results in section 4, which suggest that many nursing home residents face substantial information frictions, as well as the simple fact that many residents suffer from cognitive impairments. Moreover, prior studies have found that even choices made by relatively healthy healthcare consumers are often affected by behavioral elements such as information frictions and inertia (Abaluck and Gruber 2011, 2016; Handel 2013). It is possible that these choice frictions may limit the scope for selection bias in healthcare settings, in contrast to the more complicated patterns of student sorting sometimes found in education settings (Walters 2012; Angrist, Hull, Pathak, and Walters 2017; Mountjoy and Hickman 2019).

A second source of selection bias comes from the supply side. In education, be it schools choosing students or principals/teachers assigning students to classrooms, agents on the supply side tend to have a high degree of control over the assignment process, which can give rise to complicated patterns of selection. In contrast, while nursing homes can influence their resident mix through selective admissions as shown by Gandhi (2019) and in the present paper, they exhibit far less control over their resident mix as compared to principals and teachers in the context of classrooms.

---

161 For instance, Rothstein (2009) document that principals assign students to different classrooms based on a variety of criteria, e.g., assigning students with difficulties in a particular subject to teachers with comparative advantage in that area, and spreading high- and low-ability students and students known to create trouble across classrooms.
A Appendix Figures and Tables

Figure A.1: Nursing Home Occupancy Rates

Notes: This figure contains a histogram of nursing home occupancy rates, based on data from the OSCAR data set. The unit of observation is a nursing home-year, and observations are weighted by the number of residents admitted to the nursing home for their first stay during that year.

Figure A.2: Kernel Density Plot of Nursing Home Quality Estimates

Notes: This figure contains a kernel density plot of the main quality estimates, using an Epanechnikov kernel. The standard error for the standard deviation of nursing home quality displayed in the figure is calculated based on the square root of the variance from the inverse of the Fisher information matrix from the maximum likelihood estimation of the empirical Bayes model in equation [I].
Figure A.3: Robustness of First Stage Assumption and Exclusion Restriction

(a) County Fixed Effects

(b) No County Fixed Effects

Notes: The x-axis in these figures is the average quality of the five nearest nursing homes to each resident. Variables in panel A are residualized of county fixed effects, whereas variables in panel B are not residualized of county fixed effects.
Figure A.4: First Stage and Reduced Form for Different IV Specifications

(a) First Stage

(b) Reduced Form

Notes: These figures plot the first-stage and reduced-form coefficients and the associated 95 percent confidence intervals for IV specifications that use the quality of the $K$ nearest nursing homes to each resident as the instrument(s), for $K$ ranging from one to five.
Figure A.5: Relationship Between Quality Estimates Based on Mortality Over Different Time Horizons

(a) 19-Day Versus 90-Day Mortality

(b) 30-Day Versus 90-Day Mortality

(c) 60-Day Versus 90-Day Mortality

(d) 180-Day Versus 90-Day Mortality

Notes: The x-axis of these figures correspond to quality estimates using survival for at least 19, 30, 60, or 180 days as the outcome. The binscatters are weighted by the number of observations for each nursing home, and the standard errors are clustered by nursing home.
Figure A.6: Relationship Between Main Quality Estimates and Quality Estimates Based on Other Outcomes

(a) Stage 1 Pressure Sore

(b) Stage 2 Pressure Sore

(c) Stage 3 Pressure Sore

(d) Stage 4 Pressure Sore

(e) Physical Restraints

(f) Antipsychotic Use

Notes: The x-axis of these figures correspond to quality estimates using other resident outcomes (instead of 90-day survival rate) as the dependent variable. The estimation procedure is the same as for my main quality estimates, except that I do not use double machine-learning to select controls for computational reasons – specifically, instead of a dummy for surviving at least 90 days, I use dummies for not developing a stage $S$ (for $S \in \{1, 2, 3, 4\}$) or higher pressure sore, no use of physical restraints, and no use of antipsychotics during the first 90 days after admission as the dependent variables. The binscatters are weighted by the number of observations for each nursing home, and the standard errors are clustered by nursing home.
Figure A.7: Bin Scatters of Medicaid Against Alternative Measures of Occupancy

(a) Occupancy (No Controls)  
(b) Occupancy Percentile (No Controls)

(c) Occupancy (Controlling for Other Characteristics)  
(d) Occupancy Percentile (Controlling for Other Characteristics)

Notes: The unit of observation for these bin scatters and regressions is a resident, and the occupancy measure is the lagged 7-day average of either occupancy or occupancy percentile (as indicated in the subfigure title) as of the date of admission for the resident, residualized of nursing home-month fixed effects. Occupancy percentiles are computed based on the occupancy distribution within each nursing home. Nursing home fixed effects are included in these bin scatters and regressions.
Figure A.8: Kernel Density Plot of $occ_{ij}$ for Nursing Homes with Above-Median and Below-Median Quality

Notes: This figure contains the kernel density plot for the distributions of the supply side instrument $occ_{ij}$ at above-median and below-median quality nursing homes, based on the Epanechnikov kernel. The unit of observation is a resident-nursing home pair (for nursing homes within 15 miles of each resident).

Figure A.9: Empirical CDF of Distance Between Residents and Their Chosen Nursing Homes

Notes: This figure shows the empirical CDF of distances between residents and their chosen nursing home. The unit of observation is a resident.
Figure A.10: Conditions Under Which Selection on Unobservables Can Explain the Low Demand Estimate

Figure A.11: Plots of Choice Coefficients Against Quality Coefficients on Nursing Home Characteristics

Notes: The y-axis corresponds to estimates of the choice coefficients from the structural demand model where utility is a function of nursing home characteristics (as well as nursing home quality $\alpha_j$ in panel B, but not in panel A). The x-axis corresponds to coefficient estimates from a regression of quality on the same nursing home characteristics. Error bars correspond to the 95 percent confidence intervals.
Figure A.12: Scatterplot of Mean Utility Against Quality

Notes: The unit of observation is a nursing home, with the mean utilities net of distance preferences from the structural demand estimation (i.e. the estimated nursing home fixed effects in residents’ preferences) on the y-axis, and the quality estimates on the x-axis. The regression for the best fit line is weighted by the number of observations corresponding to the nursing home in the entire sample.

Figure A.13: Cause-Specific Baseline Hazard Functions, $h_{c,0}(t)$

Notes: These figures plot the estimated cause-specific baseline hazard functions for death in panel A, and discharge in panel B.
Figure A.14: Survival Curves for Cause-Specific Hazard Model (Split by Nursing Home Quality)

Notes: These figures plot the estimated cause-specific survival curves for a resident at a nursing home at the 25th and 75th percentiles of the quality distribution, for death in panel A, and discharge in panel B (where the “survival” curve is defined based on discharge status, rather than death).

Figure A.15: Cause-Specific Cumulative Baseline Hazard Functions, $H_{c,0}(t)$

Notes: These figures plot the estimated cause-specific cumulative baseline hazard functions for death in panel A, and discharge in panel B.
Figure A.16: Relationship Between Occupancy Fluctuations and Resident Outcomes

Notes: This figure shows a bin scatter of 90-day resident survival against my occupancy measure, controlling for resident characteristics and nursing home fixed effects. My occupancy measure is defined as the average log occupancy over the 7 days preceding admission residualized of nursing home-month fixed effects, and the sample is limited to California and the year 2009. The unit of observation is a resident.

Figure A.17: Relationship Between Baseline Health and Information Frictions

Notes: This figure shows a bin scatter of baseline information (as proxied by estimated demand for quality), against baseline mortality risk (as proxied by the relative risk in the cause-specific hazard model for death, i.e. \( \exp(W'\beta_{death, haz}) \) in equation (12)). The unit of observation is a resident.
Figure A.18: Treatment Effect Heterogeneity by Resident Characteristics

Notes: This figure shows coefficients from bivariate regressions of changes in outcome (i.e., individual-level treatment effects) for different policy interventions on various resident characteristics, where the coefficients are scaled by the average treatment effect for each policy. The change in outcome is defined as one if the resident’s counterfactual outcome is survival whereas the observed outcome is death, negative one in the opposite case, and zero if the counterfactual and observed outcomes are the same.
Figure A.19: Effect of Introduction of Five-Star Ratings on Quality

(a) Distribution of Quality Before/After Star Ratings

Notes: This figure contains kernel density plots of the quality estimates using observations from different periods of the sample, using Epanechnikov kernels. The estimation procedure is the same as for the main quality estimates $\alpha_j$. The unit of observation is a nursing home, and nursing homes are weighted based on the number of first admissions during the entire sample period.

(b) Discontinuity in Quality Trend

Notes: This figure plots a time series of the average quality estimates for nursing homes during different periods of the sample. A linear trend is added, with the bandwidth chosen so that the estimated discontinuity in the quality before and after the star ratings was introduced at the end of 2008 is smallest. The unit of observation is a nursing home, and nursing homes are weighted based on the number of first admissions during the entire sample period.
Figure A.20: Relationship Between Quality Estimates Before and After 2008

Notes: This figure shows a bin scatter of quality estimates based on residents admitted between 2009 and 2010, against quality estimates based on residents admitted between 2000 and 2007. The unit of observation is a nursing home, and nursing homes are weighted based on the number of first admissions during the entire sample period.

Figure A.21: Distributions of Quality for Nursing Homes with Different Star Ratings

Notes: This figure contains kernel density plots of the quality estimates using for nursing homes with either a one-star or a five-star rating, using Epanechnikov kernels. The unit of observation is a nursing home, and nursing homes are weighted based on the number of first admissions during the entire sample period.
Table A.1: Controls for Resident Characteristics at Admission

<table>
<thead>
<tr>
<th>Basic demographics:</th>
<th>Psychosocial well-being:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, race, gender, and marital status</td>
<td>Sense of initiative/involvement</td>
</tr>
<tr>
<td>Insurance status</td>
<td>Unsettled relationships</td>
</tr>
<tr>
<td><strong>Cognitive patterns:</strong></td>
<td>Feelings about past roles</td>
</tr>
<tr>
<td>Whether resident is comatose</td>
<td></td>
</tr>
<tr>
<td>Short-term and long-term memory</td>
<td></td>
</tr>
<tr>
<td>Memory/recall ability:</td>
<td></td>
</tr>
<tr>
<td>Current season, location of own room, staff names/faces, etc.</td>
<td></td>
</tr>
<tr>
<td>Cognitive skills for daily decision-making</td>
<td></td>
</tr>
<tr>
<td>Disordered thinking/awareness</td>
<td></td>
</tr>
<tr>
<td>Change in cognitive patterns</td>
<td></td>
</tr>
<tr>
<td><strong>Communication/hearing patterns:</strong></td>
<td></td>
</tr>
<tr>
<td>Hearing</td>
<td>Independence in bathing</td>
</tr>
<tr>
<td>Communication devices/techniques</td>
<td>Test for balance</td>
</tr>
<tr>
<td>Modes of expression</td>
<td>Functional limitation in range of motion</td>
</tr>
<tr>
<td>Ability to make self understood</td>
<td>Modes of locomotion</td>
</tr>
<tr>
<td>Speech clarity</td>
<td>Modes of transfer</td>
</tr>
<tr>
<td>Ability to understand others</td>
<td>Task segmentation</td>
</tr>
<tr>
<td>Change in communication/hearing</td>
<td>ADL functional rehabilitation potential</td>
</tr>
<tr>
<td><strong>Vision patterns:</strong></td>
<td>Change in ADL function</td>
</tr>
<tr>
<td>Vision adequacy</td>
<td></td>
</tr>
<tr>
<td>Visual limitations/difficulties</td>
<td></td>
</tr>
<tr>
<td>Use of visual appliances</td>
<td></td>
</tr>
<tr>
<td><strong>Mood and behavioral Patterns:</strong></td>
<td></td>
</tr>
<tr>
<td>Indicators of depression, anxiety, and sad mood;</td>
<td></td>
</tr>
<tr>
<td>Mood persistence</td>
<td></td>
</tr>
<tr>
<td>Change in mood</td>
<td></td>
</tr>
<tr>
<td>Behavioral symptoms:</td>
<td></td>
</tr>
<tr>
<td>Wandering, verbally/physically abusive, etc.</td>
<td></td>
</tr>
<tr>
<td>Recent change in behavioral symptoms</td>
<td></td>
</tr>
<tr>
<td><strong>Physical functioning and structural problems:</strong></td>
<td></td>
</tr>
<tr>
<td>ADL self-performance and support provided: walking, dressing, eating, etc.</td>
<td></td>
</tr>
<tr>
<td>Walking</td>
<td></td>
</tr>
<tr>
<td>Dressing</td>
<td></td>
</tr>
<tr>
<td>Eating, etc.</td>
<td></td>
</tr>
<tr>
<td>Change in communication/hearing</td>
<td></td>
</tr>
<tr>
<td>Continence:</td>
<td></td>
</tr>
<tr>
<td>Continence self-control</td>
<td></td>
</tr>
<tr>
<td>Bowel elimination pattern</td>
<td></td>
</tr>
<tr>
<td>Appliances and programs</td>
<td></td>
</tr>
<tr>
<td>Change in urinary continence</td>
<td></td>
</tr>
<tr>
<td><strong>Disease diagnoses:</strong></td>
<td></td>
</tr>
<tr>
<td>Indicators for various diseases</td>
<td></td>
</tr>
<tr>
<td>Indicators for various infections infections</td>
<td></td>
</tr>
<tr>
<td>ICD-9 codes for other diagnoses</td>
<td></td>
</tr>
</tbody>
</table>
Table A.2: Controls for Resident Characteristics at Admission (continued)

<table>
<thead>
<tr>
<th>Health conditions:</th>
<th>Activity pursuit patterns:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluid condition:</td>
<td>Time awake</td>
</tr>
<tr>
<td></td>
<td>Time involved in activities</td>
</tr>
<tr>
<td>Pain symptoms</td>
<td>Preferred activity settings</td>
</tr>
<tr>
<td>Pain site</td>
<td>General activity preferences:</td>
</tr>
<tr>
<td></td>
<td>Cards/other games</td>
</tr>
<tr>
<td></td>
<td>Crafts/arts</td>
</tr>
<tr>
<td></td>
<td>Exercise/sports</td>
</tr>
<tr>
<td></td>
<td>Music</td>
</tr>
<tr>
<td></td>
<td>Reading/writing, etc.</td>
</tr>
<tr>
<td>Stability of conditions</td>
<td>Preferences on change in daily routine</td>
</tr>
<tr>
<td>Oral/nutritional status:</td>
<td>Medications:</td>
</tr>
<tr>
<td>Oral problems:</td>
<td>Number of medications</td>
</tr>
<tr>
<td></td>
<td>New medications</td>
</tr>
<tr>
<td>Chewing</td>
<td>Injections</td>
</tr>
<tr>
<td>Swallowing</td>
<td>Days receiving various medications (antipsychotic, antidepressant, etc.)</td>
</tr>
<tr>
<td>Pain</td>
<td>Special treatments and procedures:</td>
</tr>
<tr>
<td>Height and weight</td>
<td>Special care:</td>
</tr>
<tr>
<td>Weight change</td>
<td>Chemotherapy</td>
</tr>
<tr>
<td>Nutritional problems:</td>
<td>Dialysis</td>
</tr>
<tr>
<td>Complaints about taste and/or hunger</td>
<td>IV medication, etc.</td>
</tr>
<tr>
<td>Leftover food</td>
<td>Intervention programs for mood, behavior, cognitive loss</td>
</tr>
<tr>
<td>Nutritional approaches:</td>
<td>Nursing rehabilitation/restorative care</td>
</tr>
<tr>
<td>Parental/IV</td>
<td>Devices and restraints</td>
</tr>
<tr>
<td>Feeding tube, etc.</td>
<td>Hospital stay(s)</td>
</tr>
<tr>
<td>Parenteral or enteral intake</td>
<td>ER visit(s)</td>
</tr>
<tr>
<td>Oral/dental status:</td>
<td>Physician visits</td>
</tr>
<tr>
<td>Oral status and disease prevention:</td>
<td>Physician orders</td>
</tr>
<tr>
<td>Dentures</td>
<td>Abnormal lab values</td>
</tr>
<tr>
<td>Problems with teeth</td>
<td></td>
</tr>
<tr>
<td>Inflamed gums, etc.</td>
<td></td>
</tr>
<tr>
<td>Skin condition:</td>
<td></td>
</tr>
<tr>
<td>Ulcers</td>
<td></td>
</tr>
<tr>
<td>Type of Ulcer</td>
<td></td>
</tr>
<tr>
<td>History of resolved ulcers</td>
<td></td>
</tr>
<tr>
<td>Other skin problems or lesions present</td>
<td></td>
</tr>
<tr>
<td>Skin treatments</td>
<td></td>
</tr>
<tr>
<td>Foot problems and care</td>
<td></td>
</tr>
</tbody>
</table>
Table A.3: Additional Summary Statistics

(a) Additional Summary Statistics for Residents

<table>
<thead>
<tr>
<th></th>
<th>Residents (N=653,946)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>77.541</td>
<td>0.618</td>
</tr>
<tr>
<td></td>
<td>(12.991)</td>
<td>(0.486)</td>
</tr>
<tr>
<td>Female</td>
<td>0.611</td>
<td>0.128</td>
</tr>
<tr>
<td></td>
<td>(0.488)</td>
<td>(0.334)</td>
</tr>
<tr>
<td>Married</td>
<td>0.329</td>
<td>0.889</td>
</tr>
<tr>
<td></td>
<td>(0.470)</td>
<td>(0.314)</td>
</tr>
<tr>
<td>White</td>
<td>0.735</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td>(0.441)</td>
<td>(0.273)</td>
</tr>
<tr>
<td>Black</td>
<td>0.068</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>(0.252)</td>
<td>(0.264)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.113</td>
<td>0.476</td>
</tr>
<tr>
<td></td>
<td>(0.317)</td>
<td>(0.499)</td>
</tr>
<tr>
<td>High School/Some College</td>
<td>0.634</td>
<td>0.278</td>
</tr>
<tr>
<td></td>
<td>(0.482)</td>
<td>(0.448)</td>
</tr>
<tr>
<td>At Least Bachelor's Degree</td>
<td>0.131</td>
<td>0.233</td>
</tr>
<tr>
<td></td>
<td>(0.338)</td>
<td>(0.423)</td>
</tr>
</tbody>
</table>

Notes: This table contains summary statistics for residents who had their first stays in a nursing home in California between 2000 and 2010.

Table A.4: Summary Statistics for Nursing Homes (Not Weighted by Admissions)

<table>
<thead>
<tr>
<th></th>
<th>Nursing Homes (J=840)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Beds</td>
<td>105.065</td>
</tr>
<tr>
<td></td>
<td>(48.194)</td>
</tr>
<tr>
<td>Occupancy Rate</td>
<td>86.958</td>
</tr>
<tr>
<td></td>
<td>(7.493)</td>
</tr>
<tr>
<td>Chain</td>
<td>0.600</td>
</tr>
<tr>
<td></td>
<td>(0.411)</td>
</tr>
<tr>
<td>For-Profit</td>
<td>0.877</td>
</tr>
<tr>
<td></td>
<td>(0.309)</td>
</tr>
<tr>
<td>Deficiencies</td>
<td>6.259</td>
</tr>
<tr>
<td></td>
<td>(3.541)</td>
</tr>
<tr>
<td>RN hours per resident day</td>
<td>0.331</td>
</tr>
<tr>
<td></td>
<td>(0.281)</td>
</tr>
</tbody>
</table>

Notes: This table contains summary statistics for nursing homes that were in the sample for at least one year between 2000 and 2010.
Table A.5: IV Specification Using Both Variation in Distance and Temporary Occupancy Fluctuations as Instruments

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable: 90-Day Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td><strong>TSLS</strong></td>
<td></td>
</tr>
<tr>
<td>Forecast Coefficient, $\lambda$</td>
<td>0.880</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
</tr>
<tr>
<td>Controls for Resident Characteristics</td>
<td>X</td>
</tr>
<tr>
<td>County Fixed Effects</td>
<td>X</td>
</tr>
<tr>
<td>Number of Nearest Nursing Homes</td>
<td>1</td>
</tr>
<tr>
<td>Kleibergen-Paap rk Wald F statistic</td>
<td>116</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>632,178</td>
</tr>
</tbody>
</table>

Notes: This table presents IV estimates of the effect of the nursing home quality estimate on resident outcomes, instrumenting quality of chosen nursing home with the quality of the $K$ nearest nursing homes to the resident's prior address, for $K$ ranging from 1 to 5. Standard errors clustered at the nursing home level are shown in parentheses.

Table A.6: IV Specification Using Both Variation in Distance and Temporary Occupancy Fluctuations as Instruments

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable: 90-Day Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td><strong>TSLS</strong></td>
<td></td>
</tr>
<tr>
<td>Forecast Coefficient, $\lambda$</td>
<td>0.909</td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
</tr>
<tr>
<td>Demographic Controls</td>
<td>X</td>
</tr>
<tr>
<td>Health Controls</td>
<td>X</td>
</tr>
<tr>
<td>County Fixed Effects</td>
<td>X</td>
</tr>
<tr>
<td>Number of Nearest Nursing Homes</td>
<td>1</td>
</tr>
<tr>
<td>Kleibergen-Paap rk Wald F statistic</td>
<td>39</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>586,483</td>
</tr>
</tbody>
</table>

Notes: This table presents IV estimates of the effect of the nursing home quality estimate on resident outcomes, instrumenting quality of chosen nursing home with the quality of the $K$ nearest nursing homes to the resident's prior address, and occupancy measure for the $K$ nearest nursing homes and its interaction with quality, for $K$ ranging from 1 to 3. Standard errors clustered at the nursing home level are shown in parentheses.
### Table A.7: Relationship Between Unshrunk Quality Estimates and Nursing Home Characteristics

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009 Star Ratings (s.d.)</td>
<td>0.0816**</td>
<td>0.0528</td>
<td>0.0347</td>
<td>0.0355</td>
<td>0.209***</td>
<td>0.0355</td>
<td>0.007</td>
<td>0.0298</td>
</tr>
<tr>
<td>RN hours per resident day (s.d.)</td>
<td>0.0970***</td>
<td>0.209***</td>
<td>0.0353</td>
<td>0.0505</td>
<td>0.112***</td>
<td>0.0351</td>
<td>0.007</td>
<td>0.0296</td>
</tr>
<tr>
<td>LPN hours per resident day (s.d.)</td>
<td>0.107***</td>
<td>0.112***</td>
<td>0.0226</td>
<td>0.0287</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNA hours per resident day (s.d.)</td>
<td>0.0962***</td>
<td>-0.0346</td>
<td>0.0181</td>
<td>0.0287</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deficiencies (s.d.)</td>
<td>-0.0512***</td>
<td>-0.0196</td>
<td>0.0144</td>
<td>0.0137</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>For-Profit (s.d.)</td>
<td>-0.0825***</td>
<td>-0.0615**</td>
<td>0.0314</td>
<td>0.0286</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chain (s.d.)</td>
<td>-0.0555*</td>
<td>-0.0368</td>
<td>0.0308</td>
<td>0.0296</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

** Notes: This table shows correlations between the nursing home quality estimates and various nursing home characteristics. The unit of observation is a nursing home-year. Observations are weighted such that the total weight each nursing home receives is equal to the number of long-stay residents it has over the sample period. Standard errors are clustered by nursing home.

### Table A.8: Predictivity of Quality and Nursing Home Characteristics for Resident Outcomes

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leave-Year-Out Quality Estimate (s.d.)</td>
<td>1.354***</td>
<td>1.339***</td>
</tr>
<tr>
<td></td>
<td>(0.0388)</td>
<td>(0.0400)</td>
</tr>
<tr>
<td>Deficiencies (s.d.)</td>
<td>0.0294</td>
<td>0.0284</td>
</tr>
<tr>
<td>Chain (s.d.)</td>
<td>0.0361</td>
<td>0.0505</td>
</tr>
<tr>
<td>For-Profit (s.d.)</td>
<td>0.0710</td>
<td>0.0436</td>
</tr>
<tr>
<td>RN hours per resident day (s.d.)</td>
<td>-0.0112</td>
<td>(0.0508)</td>
</tr>
<tr>
<td>LPN hours per resident day (s.d.)</td>
<td>0.0338</td>
<td>(0.0517)</td>
</tr>
<tr>
<td>CNA hours per resident day (s.d.)</td>
<td>-0.0246</td>
<td>(0.0304)</td>
</tr>
<tr>
<td>2009 Star Ratings (s.d.)</td>
<td>0.0464</td>
<td>0.0333</td>
</tr>
</tbody>
</table>

** Notes: This table shows regression results of standardized observational nursing home quality on facility characteristics. The unit of observation is a facility-year. Observations are weighted such that the total weight each nursing home receives is equal to the number of long-stay residents it has over the sample period. Standard errors are clustered by nursing home.
Table A.9: Effect of Occupancy on Admissions (Other Measures of New Admissions)

(a) Dependent Variable: Any New Residents

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged 7-Day Avg. Log Occupancy</td>
<td>-0.350***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0451)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged 7-Day Avg. Occupancy</td>
<td></td>
<td>-0.00518***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.000281)</td>
<td></td>
</tr>
<tr>
<td>Lagged 7-Day Avg. Occ. Percentile</td>
<td></td>
<td></td>
<td>-0.00120***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(5.15e-05)</td>
</tr>
<tr>
<td>Nursing Home-Month Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>N</td>
<td>1,103,528</td>
<td>1,103,528</td>
<td>1,103,528</td>
</tr>
</tbody>
</table>

Notes: This table shows regression results at the nursing home-day level wherein the dependent variable is a dummy for any new residents, and the independent variables are various measures of nursing home occupancy. Standard errors are clustered at the nursing home level.

(b) Dependent Variable: Flow of Residents

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged 7-Day Avg. Log Occupancy</td>
<td>-4.614***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.550)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lagged 7-Day Avg. Occupancy</td>
<td></td>
<td>-0.0757***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.00150)</td>
<td></td>
</tr>
<tr>
<td>Lagged 7-Day Avg. Occ. Percentile</td>
<td></td>
<td></td>
<td>-0.0153***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.000339)</td>
</tr>
<tr>
<td>Nursing Home-Month Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>N</td>
<td>1,103,528</td>
<td>1,103,528</td>
<td>1,103,528</td>
</tr>
</tbody>
</table>

Notes: This table shows regression results at the nursing home-day level wherein the dependent variable is the flow of residents (difference between number of residents today and yesterday), and the independent variables are various measures of nursing home occupancy. Standard errors are clustered at the nursing home level.
Table A.10: Types of Residents Admitted at Different Occupancies

(a) Not Controlling for Other Resident Characteristics

<table>
<thead>
<tr>
<th>Medicaid</th>
<th>Post-Acute Care</th>
<th>Dementia</th>
<th>Age</th>
<th>Female</th>
<th>Married</th>
<th>Black</th>
<th>Hispanic</th>
<th>At Least Bachelor's</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
</tr>
<tr>
<td>Residualized 7-Day</td>
<td>-0.277***</td>
<td>0.0594</td>
<td>-0.0631</td>
<td>4.960**</td>
<td>0.0611</td>
<td>0.247***</td>
<td>-0.0432</td>
<td>-0.0747</td>
</tr>
<tr>
<td>Avg. Log Occ.</td>
<td>(0.0459)</td>
<td>(0.0385)</td>
<td>(0.0657)</td>
<td>(1.945)</td>
<td>(0.0799)</td>
<td>(0.0775)</td>
<td>(0.0392)</td>
<td>(0.0502)</td>
</tr>
<tr>
<td>Nursing Home Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>N</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.218</td>
<td>0.126</td>
<td>0.043</td>
<td>0.145</td>
<td>0.023</td>
<td>0.028</td>
<td>0.180</td>
<td>0.166</td>
</tr>
</tbody>
</table>

Notes: All regressions include nursing home fixed effects, and are at the resident level. Standard errors are clustered by nursing home.

(b) Controlling for Other Resident Characteristics

<table>
<thead>
<tr>
<th>Medicaid</th>
<th>Post-Acute Care</th>
<th>Dementia</th>
<th>Age</th>
<th>Female</th>
<th>Married</th>
<th>Black</th>
<th>Hispanic</th>
<th>At Least Bachelor's</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
</tr>
<tr>
<td>Residualized 7-Day</td>
<td>-0.241***</td>
<td>0.0269</td>
<td>-0.0889</td>
<td>3.994**</td>
<td>0.121</td>
<td>0.244***</td>
<td>-0.0346</td>
<td>-0.0522</td>
</tr>
<tr>
<td>Avg. Log Occ.</td>
<td>(0.0450)</td>
<td>(0.0380)</td>
<td>(0.0631)</td>
<td>(1.849)</td>
<td>(0.0767)</td>
<td>(0.0747)</td>
<td>(0.0386)</td>
<td>(0.0492)</td>
</tr>
<tr>
<td>Nursing Home Fixed Effects</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>N</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
<td>159,529</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.247</td>
<td>0.148</td>
<td>0.118</td>
<td>0.228</td>
<td>0.100</td>
<td>0.097</td>
<td>0.207</td>
<td>0.202</td>
</tr>
</tbody>
</table>

Notes: All regressions include nursing home fixed effects, and are at the resident level. Standard errors are clustered by nursing home.
Table A.11: Heterogeneous Impact of Star Ratings

<table>
<thead>
<tr>
<th>Resident Preferences</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance to Facility</td>
<td>-0.159***</td>
</tr>
<tr>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td>6.879***</td>
</tr>
<tr>
<td>(1.264)</td>
<td></td>
</tr>
<tr>
<td>Quality x Alzheimer’s</td>
<td>-2.293***</td>
</tr>
<tr>
<td>(0.419)</td>
<td></td>
</tr>
<tr>
<td>Quality x Age</td>
<td>-0.071***</td>
</tr>
<tr>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>Quality x Female</td>
<td>-0.257</td>
</tr>
<tr>
<td>(0.353)</td>
<td></td>
</tr>
<tr>
<td>Quality x Married</td>
<td>-0.192</td>
</tr>
<tr>
<td>(0.371)</td>
<td></td>
</tr>
<tr>
<td>Quality x Black</td>
<td>-0.663*</td>
</tr>
<tr>
<td>(0.69)</td>
<td></td>
</tr>
<tr>
<td>Quality x Hispanic</td>
<td>-2.614***</td>
</tr>
<tr>
<td>(0.68)</td>
<td></td>
</tr>
<tr>
<td>Quality x At Least Bachelor’s Degree</td>
<td>0.997*</td>
</tr>
<tr>
<td>(0.509)</td>
<td></td>
</tr>
<tr>
<td>Quality x Lived Alone</td>
<td>-0.478*</td>
</tr>
<tr>
<td>(0.419)</td>
<td></td>
</tr>
<tr>
<td>Quality x 2009-10</td>
<td>-1.409*</td>
</tr>
<tr>
<td>(1.311)</td>
<td></td>
</tr>
<tr>
<td>Qual. x Post-Star Ratings x At Least Bachelor’s Degree</td>
<td>0.522</td>
</tr>
<tr>
<td>(0.635)</td>
<td></td>
</tr>
<tr>
<td>Qual. x Post-Star Ratings x Black</td>
<td>-1.159*</td>
</tr>
<tr>
<td>(0.795)</td>
<td></td>
</tr>
<tr>
<td>Qual. x Post-Star Ratings x Dementia</td>
<td>-0.16</td>
</tr>
<tr>
<td>(0.495)</td>
<td></td>
</tr>
<tr>
<td>Qual. x Post-Star Ratings x Age</td>
<td>0.02*</td>
</tr>
<tr>
<td>(0.016)</td>
<td></td>
</tr>
<tr>
<td>Qual. x Post-Star Ratings x Female</td>
<td>0.397</td>
</tr>
<tr>
<td>(0.447)</td>
<td></td>
</tr>
<tr>
<td>Qual. x Post-Star Ratings x Married</td>
<td>-0.199</td>
</tr>
<tr>
<td>(0.469)</td>
<td></td>
</tr>
<tr>
<td>Qual. x Post-Star Ratings x Hispanic</td>
<td>0.467</td>
</tr>
<tr>
<td>(0.627)</td>
<td></td>
</tr>
<tr>
<td>Qual. x Post-Star Ratings x Lived Alone</td>
<td>0.699*</td>
</tr>
<tr>
<td>(0.52)</td>
<td></td>
</tr>
</tbody>
</table>

Supply Side

| Temporary Fluctuations in log(occupancy)               | -7.493*** |
| (1.081)                                                |        |

Resident Controls in Supply Side Equation X
Quality x Demographic Variables X

Notes: This table shows results from the structural estimation using Gibbs sampling. A burn-in period corresponding to the first half of the chain was used.
Table A.12: Supply Side Estimates for Specifications in Table 4

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporary Fluctuations in log(occupancy)</td>
<td>-7.34***</td>
<td>-7.97***</td>
<td>-6.95***</td>
<td>-7.38***</td>
</tr>
<tr>
<td></td>
<td>(1.069)</td>
<td>(1.24)</td>
<td>(1.365)</td>
<td>(1.1)</td>
</tr>
<tr>
<td>Dementia</td>
<td>0.006</td>
<td>-0.069***</td>
<td>-0.068***</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.021)</td>
<td>(0.016)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>Age</td>
<td>0.027***</td>
<td>0.026***</td>
<td>0.02***</td>
<td>0.027***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Female</td>
<td>0.091*</td>
<td>0.074*</td>
<td>0.027*</td>
<td>0.097***</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.046)</td>
<td>(0.028)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Married</td>
<td>0.184***</td>
<td>0.204***</td>
<td>0.164***</td>
<td>0.19***</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.071)</td>
<td>(0.058)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>Race (Black)</td>
<td>-0.71***</td>
<td>-0.671***</td>
<td>-0.594***</td>
<td>-0.688***</td>
</tr>
<tr>
<td></td>
<td>(0.176)</td>
<td>(0.173)</td>
<td>(0.187)</td>
<td>(0.177)</td>
</tr>
<tr>
<td>Race (Hispanic)</td>
<td>-0.145***</td>
<td>-0.182***</td>
<td>-0.126***</td>
<td>-0.18***</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
<td>(0.045)</td>
<td>(0.024)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>Education (Bachelor's)</td>
<td>-0.098***</td>
<td>-0.141***</td>
<td>-0.118***</td>
<td>-0.111**</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.035)</td>
<td>(0.025)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Admitted from Acute Care Hospital</td>
<td>0.057</td>
<td>0.228**</td>
<td>0.179**</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.083)</td>
<td>(0.091)</td>
<td>(0.086)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Lived Alone</td>
<td>0.076***</td>
<td>0.082**</td>
<td>0.052***</td>
<td>0.094***</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.033)</td>
<td>(0.011)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Year (2009)</td>
<td>-0.01</td>
<td>-0.041*</td>
<td>-0.001</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.023)</td>
<td>(0.011)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Year (2010)</td>
<td>0.052**</td>
<td>0.063***</td>
<td>0.066***</td>
<td>0.035*</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.019)</td>
<td>(0.02)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.307**</td>
<td>-1.346***</td>
<td>-1.064**</td>
<td>-1.322**</td>
</tr>
<tr>
<td></td>
<td>(0.56)</td>
<td>(0.483)</td>
<td>(0.459)</td>
<td>(0.536)</td>
</tr>
</tbody>
</table>

Quality in Residents' Utility Equation | X | X | X |
Nursing Home Characteristics in Residents' Utility Equation | X | X |
Qual. Interacted with Res. Chars. in Residents' Utility Equation | X |

Notes: This table shows the supply side estimates from the structural estimation using Gibbs sampling. A burn-in period corresponding to the first half of the chain was used.
Table A.13: Resident Preferences: Robustness Checks

<table>
<thead>
<tr>
<th>Panel A: Different Specifications for Instruments</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand for Quality</td>
<td>0.684*</td>
<td>0.752*</td>
<td>0.681***</td>
</tr>
<tr>
<td></td>
<td>(0.416)</td>
<td>(0.431)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>Instruments Interacted with Resident Characteristics</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Distance</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>4th Order Polynomial in Distance</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Panel B: Different Sample Definitions

<table>
<thead>
<tr>
<th>Panel B: Different Sample Definitions</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand for Quality</td>
<td>0.647*</td>
<td>0.272*</td>
<td>-0.411*</td>
</tr>
<tr>
<td></td>
<td>(0.436)</td>
<td>(0.284)</td>
<td>(0.215)</td>
</tr>
<tr>
<td>Choice Set Includes Nursing Homes Within X Miles</td>
<td>15 miles</td>
<td>10 miles</td>
<td>5 miles</td>
</tr>
<tr>
<td>Drop Rapidly Expanding/Contracting Nursing Homes</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table shows results from the structural estimation using Gibbs sampling. A burn-in period corresponding to the first half of the chain was used.

---

Table A.14: Demand Heterogeneity: Robustness Checks

<table>
<thead>
<tr>
<th>Resident Preferences</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility (Quality x Dementia)</td>
<td>-2.424***</td>
<td>-2.221***</td>
<td>-2.430***</td>
<td>-2.003***</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.277)</td>
<td>(0.273)</td>
<td>(0.353)</td>
</tr>
<tr>
<td>Utility (Quality x Age)</td>
<td>-0.059***</td>
<td>-0.067***</td>
<td>-0.052***</td>
<td>-0.058***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.01)</td>
<td>(0.012)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Utility (Quality x Bachelor's)</td>
<td>1.326***</td>
<td>1.372***</td>
<td>1.376***</td>
<td>1.802***</td>
</tr>
<tr>
<td></td>
<td>(0.299)</td>
<td>(0.313)</td>
<td>(0.299)</td>
<td>(0.347)</td>
</tr>
<tr>
<td>Utility (Qual x Lived Alone)</td>
<td>-0.046</td>
<td>-0.276*</td>
<td>-0.067</td>
<td>-0.187</td>
</tr>
<tr>
<td></td>
<td>(0.245)</td>
<td>(0.259)</td>
<td>(0.254)</td>
<td>(0.287)</td>
</tr>
<tr>
<td>Utility (Quality x Post-Star Ratings)</td>
<td>0.56***</td>
<td>0.518**</td>
<td>0.519**</td>
<td>0.498**</td>
</tr>
<tr>
<td></td>
<td>(0.199)</td>
<td>(0.202)</td>
<td>(0.218)</td>
<td>(0.242)</td>
</tr>
</tbody>
</table>

| Demand and Supply Instruments | X      | X      | X      | X      |
| Quality in Utility Equation   | X      | X      |        |        |
| Quality Interacted with Other Resident Characteristics | X      | X      | X      | X      |
| Quality Interacted with Size of Choice Set | X      |        |        |        |
| Instruments Interacted with Resident Characteristics | X      |        |        |        |
| Nursing Home Fixed Effects in Utility Equation | X      |        |        |        |

Notes: This table shows results from the structural estimation using Gibbs sampling. A burn-in period corresponding to the first half of the chain was used.
Table A.15: Heterogeneous Demand for Quality and Other Nursing Home Characteristics

<table>
<thead>
<tr>
<th>Resident Preferences</th>
<th>Dementia</th>
<th>Resident Characteristic</th>
<th>Age</th>
<th>Bachelor’s Degree</th>
<th>Post-Star Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility (Quality x Resident Characteristic)</td>
<td>-2.523***</td>
<td>(1)</td>
<td>-0.077***</td>
<td>1.021***</td>
<td>0.841***</td>
</tr>
<tr>
<td></td>
<td>(0.298)</td>
<td></td>
<td>(0.01)</td>
<td>(0.334)</td>
<td>(0.235)</td>
</tr>
<tr>
<td>Utility (RN Staffing x Resident Characteristic)</td>
<td>-0.12***</td>
<td>(2)</td>
<td>-0.004***</td>
<td>0.002</td>
<td>0.023*</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td></td>
<td>(0)</td>
<td>(0.014)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Utility (LPN Staffing x Resident Characteristic)</td>
<td>-0.126***</td>
<td>(3)</td>
<td>-0.001*</td>
<td>0.066***</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td></td>
<td>(0.001)</td>
<td>(0.012)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Utility (CNA Staffing x Resident Characteristic)</td>
<td>0.05***</td>
<td>(4)</td>
<td>0.003***</td>
<td>0.017***</td>
<td>0.015***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td></td>
<td>(0)</td>
<td>(0.006)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Utility (Deficiencies x Resident Characteristic)</td>
<td>-0.005***</td>
<td>(5)</td>
<td>-0.001***</td>
<td>-0.01***</td>
<td>-0.015***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td></td>
<td>(0)</td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Utility (For-Profit x Resident Characteristic)</td>
<td>0.024*</td>
<td>(6)</td>
<td>-0.013***</td>
<td>-0.248***</td>
<td>0.062***</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td></td>
<td>(0.001)</td>
<td>(0.026)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>Utility (Chain x Resident Characteristic)</td>
<td>-0.051***</td>
<td>(7)</td>
<td>0.001*</td>
<td>-0.006</td>
<td>0.01*</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td></td>
<td>(0.001)</td>
<td>(0.016)</td>
<td>(0.009)</td>
</tr>
</tbody>
</table>

Demand and Supply Instruments: X X X X
Quality in Utility Equation: X X X X

Notes: This table shows results from the structural estimation using Gibbs sampling. Coefficient estimates between various resident characteristics and nursing home variables are shown, with the resident characteristic being interacted with the nursing home variables being indicated at the top of each column. A burn-in period corresponding to the first half of the chain was used.

Table A.16: Predictivity of Quality for Outcomes of Residents with Different Characteristics

<table>
<thead>
<tr>
<th>Dependent Variable: 90-Day Survival</th>
<th>OLS</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Leave-Year-Out Quality</td>
<td>0.692***</td>
<td>0.501***</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.138)</td>
</tr>
<tr>
<td>Leave-Year-Out Quality x Above 80</td>
<td>0.122**</td>
<td>0.400***</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.137)</td>
</tr>
<tr>
<td>Leave-Year-Out Quality x Dementia</td>
<td>0.260***</td>
<td>0.619***</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.118)</td>
</tr>
<tr>
<td>Leave-Year-Out Quality x At Least Bachelor’s Degree</td>
<td>-0.034</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.160)</td>
</tr>
<tr>
<td>Leave-Year-Out Quality x Post-Star Ratings</td>
<td>-0.298***</td>
<td>-0.562***</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.133)</td>
</tr>
</tbody>
</table>

Demographic Controls: X X
Health Controls: X X
County Fixed Effects: X X
Number of Observations: 632,223 632,178

Notes: This table presents estimates of treatment effect heterogeneity. The IV specification is based on the leave-year-out quality of the closest nursing home to each resident, interacted with the same characteristics as in the endogenous variables. Standard errors clustered at the nursing home level are shown in parentheses.
Table A.17: Short-Run Effects of Eliminating Information Frictions Under Different Assumptions on Correlation Between Shocks

<table>
<thead>
<tr>
<th>Assumption on Correlation Between Shocks</th>
<th>Change in Quality</th>
<th>Change in Survival Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfectly Correlated</td>
<td>0.003</td>
<td>0.005</td>
</tr>
<tr>
<td>Independent</td>
<td>0.003</td>
<td>0.021</td>
</tr>
</tbody>
</table>

Number of Observations 57,636

Notes: This table presents simulation results for the short-run effects of eliminating information frictions, under the assumption that mortality and discharge shocks across potential nursing home-resident pairs within resident are either perfectly correlated, or independent.

Table A.18: Long Run Effect of Eliminating Information Frictions (Alternative Functional Form Specifications)

<table>
<thead>
<tr>
<th>Demand Function</th>
<th>Cost Function</th>
<th>Reduction in Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{\exp(\kappa_j^{\text{quality}} \alpha_j + \xi_j)}{\sum_l \exp(\kappa_l^{\text{quality}} \alpha_l + \xi_l)} \cdot \bar{N} )</td>
<td>( \left( c_j^{\text{quality}} + \frac{\alpha_j^2 + c_j^N}{2} \right) N_j )</td>
<td>44%</td>
</tr>
<tr>
<td>( \frac{\exp(\kappa_j^{\text{quality}} \alpha_j + \xi_j)}{\sum_l \exp(\kappa_l^{\text{quality}} \alpha_l + \xi_l)} \cdot \bar{N} )</td>
<td>( \left( c_j^{\text{qual}} \cdot \alpha_j + c_j^N \right) N_j )</td>
<td>85%</td>
</tr>
<tr>
<td>( \frac{\exp(\kappa_j^{\text{quality}} \alpha_j + \xi_j)}{\sum_l \exp(\kappa_l^{\text{quality}} \alpha_l + \xi_l)} \cdot \bar{N} )</td>
<td>( \frac{c_j^{\text{quality}}}{2} \alpha_j^2 + c_j^N N_j )</td>
<td>49%</td>
</tr>
<tr>
<td>( A_{\kappa^{\text{quality}}} \alpha_j + N \varphi(\bar{p}_{res}) )</td>
<td>( \left( c_j^{\text{quality}} + \frac{\alpha_j^2 + c_j^N}{2} \right) N_j )</td>
<td>26%</td>
</tr>
<tr>
<td>( B_{\kappa^{\text{quality}}} \log(\alpha_j) + N \varphi(\bar{p}_{res}) )</td>
<td>( \left( c_j^{\text{quality}} + \frac{\alpha_j^2 + c_j^N}{2} \right) N_j )</td>
<td>26%</td>
</tr>
</tbody>
</table>

Notes: In these simulations, in order for additional quality investments to be costly, I use a location normalization for quality to ensure that it is always positive (specifically, setting the lowest quality to be equal to zero, or in the logarithm specification, equal to one).