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

By Alden Cheng*

This paper studies consumers' demand for quality in the nursing home market, where information frictions are a potential concern. Using administrative data on the universe of nursing home residents, I estimate substantial variation in quality of nursing homes in California. Yet, structural demand estimates reveal that average demand for quality is very low, with residents who were younger, highly educated, free from dementia, and who made their choices after the introduction of the star rating system being more responsive to quality. Counterfactual simulations suggest that eliminating information frictions may reduce deaths by at least 8 to 28 percent.

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).

Despite the importance of information frictions about quality of care, limited empirical work 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.[1] 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 Spinnewijn 2021) and treatment decisions (Kolstad 2013; Chan, Gentzkow, and Yu 2022; Mullainathan and Obermeyer 2022).[2] However, it is unclear whether lessons learned from these settings apply to

* Postdoctoral Research Associate (starting in August), National Bureau of Economic Research/Center for Business and Public Policy (CBPP) at Gies College of Business, University of Illinois (email: alden15@nber.org). 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, Martin Hackmann, Daniel Keniston, Kyoo il Kim, Kurt Lavetti, Jetson Leder-Luis, Haizhen Lin, Justin Marion, Philip Marx, Martin Mattsson, Nidhiya Menon, David Molitor, Francesc Ortega, Elena Prager, Julian Reif, Nina Roussille, Akshar Saxena, Frank Schilbach, Mark Shepard, Maggie Shi, Ben Shiller, George Siddarth, David Simon, 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 National Bureau of Economic Research for his assistance. All mistakes are my own.

[1]It 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.

[2]In 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.

consumer choice over healthcare quality, given that the nature of information frictions is very different qualitatively.[3] 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 (Cenziper, Jacobs, Crites, and Englund 2020). Second, poor nursing home quality is an issue of great interest to policymakers; for example, President Biden 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 (National Center for Health Statistics 2017), and more than half of those aged 57–61 today will spend some time in a nursing home (Hurd, Michaud, and Rohwedder, 2017).

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.[4] In addition, to address concerns about selection on unobservables, I conduct a validation test using a distance-based instrumental variables (IV) strategy (Card 1993; Card, Fenizia, and Silver 2019). Specifically, I show that variation in the quality of residents' chosen nursing homes induced by differential distances to their prior addresses predicts resident outcomes one-for-one (in expectation).

In the second step, I use these quality estimates to study residents' demand for quality, taking unobserved supply-side constraints arising from selective admissions practices by nursing homes into account (Gandhi 2019).[5] I use distance and temporary occupancy fluctuations as demand and supply instruments respectively in my structural demand model, and I estimate the model using Gibbs sampling with data augmentation to avoid the curse of dimensionality (Agarwal and Somaini 2022).[6]

In the final step, I simulate the consequences of information frictions,[7] the sign of which are a priori

[3]In 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.

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

[5]More precisely, by a resident's "demand for quality", I refer to the demand for quality of the decision maker(s) (who is not necessarily the resident), assuming that they have full information. The identity of the decision maker is not important for the main findings of this paper, but may be relevant for follow-up questions (e.g., regarding the design of effective information interventions).

[6]The number of possible distinct sets of nursing homes willing to admit each resident increases exponentially with the number of nursing homes in her vicinity, exceeding 10^{60} for some residents. This makes methods such as maximum likelihood that require directly integrating over each distinct possibility computationally infeasible.

[7]I 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 decision-making process

unclear due to the theory of the second best.[8,9] To do so, I combine the structural demand model (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). I also consider a simple model that endogenizes nursing homes' quality investment decisions 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.

I estimate substantial variation in nursing home quality across California (both overall and locally), and the distance-based IV strategy supports the validity of these quality estimates. The estimates imply that a resident who goes to a nursing home with one standard deviation higher quality is 2 percentage points less likely to die within 90 days of admission (all else equal), a 27 percent reduction relative to the baseline mortality rate. Moreover, the correlations between my quality measure and publicly available nursing home characteristics have the expected signs, but my quality measure is a far stronger predictor of resident outcomes than these other 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,[10] and that residents also do not seem to take full advantage of publicly available information about nursing home quality. Patterns of demand heterogeneity support the interpretation that information frictions are 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.

Finally, counterfactual simulations suggest that eliminating information frictions may reduce nursing home deaths in California by at least 8–28 percent (or 250–850 deaths annually) in the short run (i.e., keeping quality fixed), and potentially more in the long run after accounting for endogenous quality adjustments by nursing homes. 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 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

(Handel and Schwartzstein 2018).

[8]Specifically, 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, especially disadvantaged residents. Handel (2013) provides an example of how eliminating only one of several market imperfections may worsen outcomes in the context of inertia and adverse selection.

[9]More 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.

[10]The demand estimate implies that residents are willing to tolerate a 28.6 percentage point higher probability of death within 90 days in exchange for a one-mile reduction in travel distance.

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.[11] 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),[12] and also provides one of the first applications of double machine learning in this area (Belloni, Chernozhukov, and Hansen, 2014; Chernozhukov et al. 2018).[13]

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), 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 1, 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 2, before using them to study residents' demand for quality and information frictions in section 3. In section 4, I quantify the costs of information frictions via counterfactual simulations, and I conclude in section 5.

1 Background

1.1 Nursing Home Industry

There are roughly 15,000 nursing homes in the US providing care for about 1.3 million Americans (National Center for Health Statistics 2017), and an estimated 56 percent of Americans currently aged 57–61 are expected to spend at least one night in a nursing home during their lifetimes (Hurd, Michaud, and Rohwedder 2017).[14] Nursing home residents vary widely in their medical conditions

[11] 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), consumer search (Abaluck and Compiani 2020), or two-sided matching (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.

[12] 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).

[13] 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).

[14] More precisely, throughout this paper I colloquially refer to skilled nursing facilities (SNFs) that are certified by the CMS as nursing homes.

and needs, but as a crude approximation, they fall into 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.[15] By contrast, long-stay residents often suffer from chronic conditions (e.g., cognitive decline), and are unlikely be discharged in the short term. Many of these residents are covered by Medicaid, which has substantially lower reimbursement rates than Medicare.[16] 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 nursing homes (Comondore et al. 2009), which account for the majority of nursing homes in the US.[17] 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.

1.2 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 inspections by regulators,[18] 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.[19] 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,[20] resident outcomes,[21] deficiency citations,[22] and the five-star ratings by nursing homes has been well documented over the years.[23] Moreover, even unintentional errors in misreporting can be problematic:

[15] Medicare 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.

[16] Specifically, 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).

[17] 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).

[18] Inspections are conducted annually, but complaints may also trigger additional inspections.

[19] Evidence of this is shown in section 2.3.1.

[20] 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).

[21] Schizophrenia diagnoses rose sharply in the years following 2012 when the government began publicly releasing information about inappropriate antipsychotic use (Thomas, Gebeloff, and Silver-Greenberg 2021).

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

[23] 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 (Silver-Greenberg and Gebeloff 2021).

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 before having another assessment, in which case the adverse outcome may not be recorded in the data.[24,25]

Third, death shortly after admission is an undesirable outcome for most types of residents.[26] This is important given that I am studying residents from all payer sources, and goals of care vary substantially across different types of residents.[27]

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.[28]

1.3 Data

The primary data source for this paper is the Minimum Data Set 2.0 (MDS), and all nursing homes that receive federal funding are required to fill out MDS assessment forms at regular intervals (42 CFR §483.20).[29,30] 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), oral health, nutrition, dental status, skin conditions, activity pursuit patterns, medications, special treatments and procedures, and discharge potential,[31] and several studies on the accuracy of MDS data have found it to be fairly reliable (Shin and Scherer 2009).[32]

[24]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.

[25]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.

[26]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.

[27]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.

[28]In contrast, it is more difficult to interpret the magnitudes of demand based on various other quality measures. For example, while development of pressure sores is a commonly used quality measure in the health services literature on nursing homes, the economic cost of a pressure sore is not well-established, which makes it difficult to determine whether a given level of demand for this quality measure is "too large" or "too small."

[29]The set of nursing homes receiving federal funding account for roughly 96 percent of all nursing homes (Grabowski, Gruber, and Angelelli 2008).

[30]Assessment forms must completed upon admission, at discharge (or death), quarterly in between, and whenever there is a significant change in status.

[31]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.

[32]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

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.[33]

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 2 and 3. 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.[34] To calculate nursing homes' occupancy levels over time, I use 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.[35,36] 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 3.[37] Finally, I consider the single year of 2009 for my counterfactual simulations in section 4, given that various assumptions that the simulations depend on (e.g., counterfactual nursing home entries and exits) are more suspect over longer time horizons. Additional details about the data can be found in Appendix B.

1.4 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

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.

[33] 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.

[34] 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.

[35] 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).

[36] 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.

[37] 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.

at admission, and 7.5 percent die within 90 days of admission.[38] Appendix Table A.3a contains more detailed summary statistics for nursing home residents.

Panel B summarizes nursing home statistics,[39] showing that nursing homes have 120 beds on average, occupancy rates tend to be quite high (86 percent on average)[40], 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)

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

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.

2 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 explore the relationship between my quality estimates and other potential quality measures, as well as the degree of geographical concentration in nursing home quality.

2.1 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:

$$Y_i = \mu_1 + \sum_{j=2}^{J} \beta_j D_{ij} + X_i'\gamma + u_i,\ \mathbb{E}[X_i u_i] = 0, \tag{1}$$

[38]The MDS also records the payer source for residents at admission; although, as I discuss in Appendix B, this is less accurate than claims data.

[39]These 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).

[40]A histogram of nursing homes' occupancy rates is shown in Appendix Figure A.1.

where Y_i is a dummy for whether resident i survives at least 90 days after admission, β_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),[41] and denote the quality estimates by $\{\alpha_j\}$.

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).[42] While controlling for lagged outcomes is impossible in the present setting with survival being the outcome variable, I control for residents' baseline mortality risk using more than 500 health and demographic variables (even without interaction terms) that are recorded for each resident upon admission.

An issue with using the full set of controls for quality estimation is the risk of overfitting, especially because some of the health controls correspond to relatively rare medical conditions, and the sample sizes for nursing homes are not huge.[43] Hence, I use a modification of the post-double-selection method in Belloni, Chernozhukov, and Hansen (2014) to select an appropriate set of controls for my quality estimation (see Appendix J for details).[44,45,46]

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 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 $\{\alpha_j\}$ are valid, then exogenous variation in the estimated quality of residents' nursing homes should predict outcomes one-for-one. So, to obtain exogenous variation in the quality of residents' chosen nursing homes, I leverage residents' preferences for closer nursing homes.

[41] 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 the main result in this paper is that estimated 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.

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

[43] 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.

[44] 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 depends on the relationship between the omitted variable and the dependent and treatment variables, one runs two lasso regressions of the dependent variable and the treatment variable on the controls. To account for the possibility of modest variable selection errors in this first step, one takes the union of the two sets of controls selected by lasso in the two lasso regressions, and estimates the treatment effect controlling for these variables.

[45] 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.

[46] It turns out that my quality estimates are not sensitive to the variable selection method. 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.

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 λ 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 η_i being a "forecast residual" (see Appendix Section C for the derivation of the structural equation and the definition of the forecast residual η_i).[47,48] 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. I then test whether λ is equal to one, which it should be if the quality estimates α_j are asymptotically unbiased on average.[49] 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:*

$$\alpha_i^{\sim t(i)} = \delta_0 + Z_i' \delta_Z + X_i' \delta_X + e_i, \; \mathbb{E}\left[(Z_i', X_i')' \, e_i \right] = 0. \tag{3}$$

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

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

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.,*

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

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 potential sorting along these lines, I focus only on local variation in distance to

[47] I use leave-year-out quality estimates to avoid a mechanical relationship between resident outcomes and the quality estimate of residents' chosen nursing homes. Specifically, the outcome Y_i is used in the estimation of nursing home quality α_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 between Y_i and α_i. Due to the leave-year-out definition of the quality variable $\alpha_i^{\sim t(i)}$ used in this validation exercise, we can also interpret equation (2) as an out-of-sample test of the predictivity of my quality estimates.

[48] I set the quality estimate α_j of the first nursing home ($j = 1$) to zero since the notation of this section treats $j = 1$ as the "omitted category."

[49] 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 over time tends to infinity, in which case the empirical Bayes and OLS estimates will eventually coincide.

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,[50] 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 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).

2.2 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 percentage points 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. 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 in equation (2) 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.[51] By contrast, the lack of a clear relationship between the instrument and survival probability based on resident characteristics net of nursing home effects (which I use as a proxy for baseline health) lends confidence to the exclusion restriction.[52,53]

I provide a visualization of the main IV result in Figure 2 by plotting the second stage of the IV. If the quality estimates are asymptotically unbiased on average, the forecast coefficient is one so 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 respectively) 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

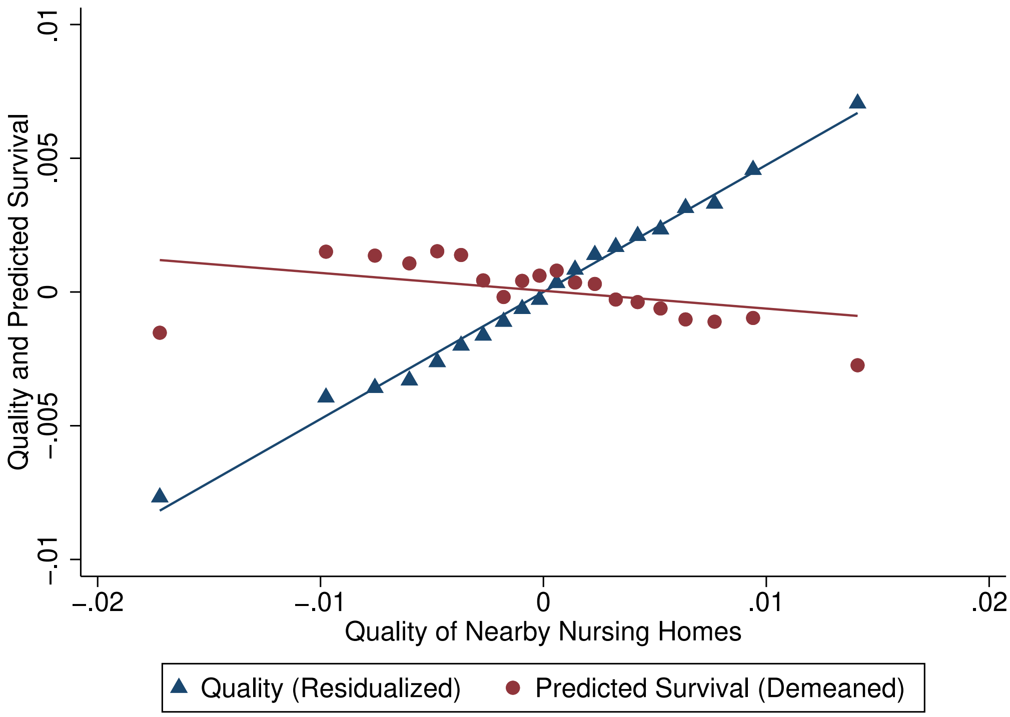
[50] In particular, η_i is not a structural parameter but rather arises from a complicated statistical relationship between the causal parameters β_j and the quality estimates α_j.

[51] 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 1 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.

[52] 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.

[53] The balance test in Figure 1 checks whether the residualized instrument is correlated with survival probability. A stronger version of this test that checks whether the *unresidualized* instrument is correlated with survival probability is shown in Appendix Figure A.3. These results reveal that the instrument and survival probability are uncorrelated as long as county fixed effects are accounted for.

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 plotted on the same scale).

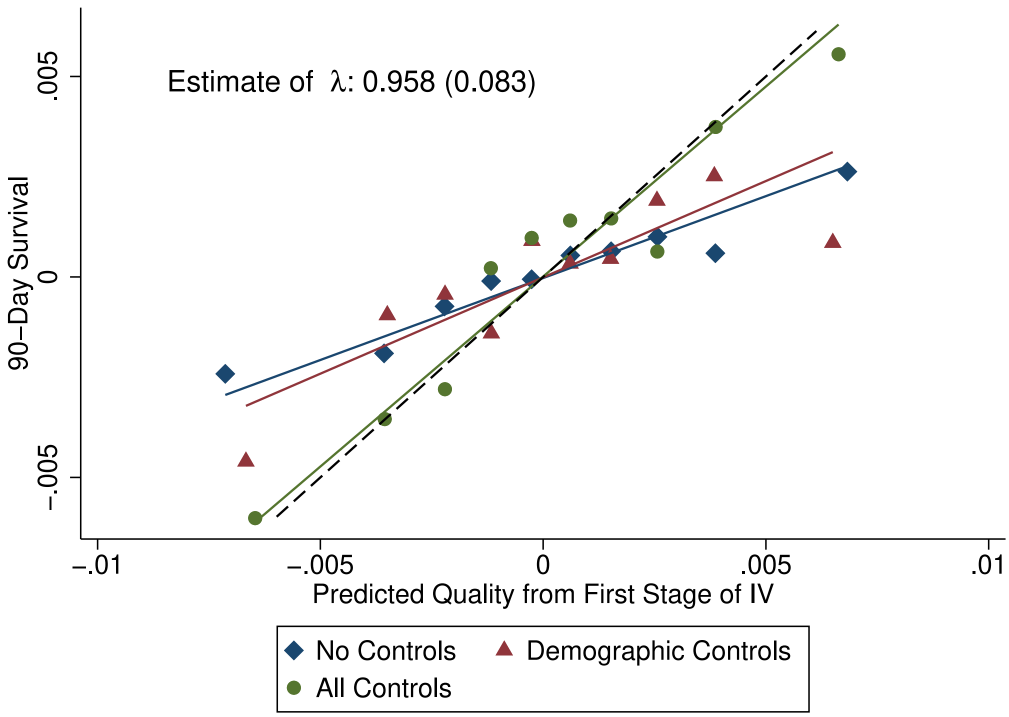
coefficient are not statistically different from one (at the 5 percent significance level).[54] In addition, the first stage and reduced form estimates for these models (shown in Appendix Figure A.4) demonstrate that the effects of nursing homes closer to a resident's prior address on both the quality 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 estimates to my main quality estimates that are based on 90-day survival.[55] Appendix Figure A.5

[54]The point estimates in most specifications are smaller than one, but are never statistically distinguishable from one at the 5 percent significance level. The point estimates being smaller than one may be an artifact of measurement error. In particular, if the measurement errors in the instrument and the endogenous variable are positively correlated (as they are likely to be in this setting), the IV estimates still suffer from a similar type of attenuation bias that OLS estimates do. See Appendix L for a more detailed explanation.

[55]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.

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.

shows that quality estimates based on mortality over different time horizons are all highly correlated, suggesting that truncation bias is unlikely to meaningfully affect my results.

Finally, in Appendix Section D, I consider several alternative identification strategies for estimating the forecast coefficient, such as using different instruments or conducting an event study based on entry by high-quality nursing homes. To summarize the results from these exercises briefly, I do not find any evidence that my quality estimates are biased; the identifying assumptions for these alternative specifications are not always met, but in the cases where they are, the results support the validity of my quality estimates. Given the evidence presented in this subsection on the validity of the quality estimates, for notational simplicity in subsequent text I will not distinguish between quality β_j and the quality estimates α_j (except when making a distinction is absolutely necessary, e.g., when discussing issues such as measurement error).[56]

[56]There is also a bit of slippage in the sense that there is a knife-edge case where some quality estimates α_j may be (asymptotically) biased for certain nursing homes but the forecast coefficient is still one. However, this only occurs if these biases have the opposite sign for different nursing nursing homes and exactly cancel out on average. Moreover, I present evidence in section 3 that my main result on residents' low demand for quality is robust to the potential for this type of bias.

2.3 Discussion of Nursing Home Quality Estimates

2.3.1 Relationship Between Different Dimensions of Quality

Previously, I conjectured that different dimensions of quality are likely to be positively correlated. Here, I provide evidence on this claim 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 — 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 (Grabowski et al. 2016; You et al. 2016).[57,58,59] However, only a small fraction of the variation in my quality measure can be explained by these quality indicators,[60] 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.[61]

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 survival-based quality estimates tend to be positively correlated with quality estimates based on these other outcomes.

2.3.2 Geographical 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-quality nursing homes. To shed light on whether this is the case, 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.[62] The figure shows that there is almost as much variation in nursing home quality within counties as there is overall,[63] implying that there tends to be

[57] 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.

[58] 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 estimates of residents' demand for quality.

[59] Appendix Table A.7 shows that the correlations between the *unshrunken* quality estimates (i.e., coefficients from a fixed effects regression, instead of empirical Bayes estimates) with these nursing home characteristics are very similar qualitatively.

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

[61] 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 coefficients and standard errors by 100 for better legibility. The coefficient on the leave-year-out quality estimate in these regressions is highly statistically significant and an order of magnitude larger than coefficients on other quality indicators (which are mostly statistically insignificant at the 5 percent significance level). Moreover, both the coefficient on my quality measure and the R-squared remain essentially unchanged when I include other quality indicators in the regression.

[62] 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.

[63] Moreover, within-county variation in quality accounts for far more of the overall variation than across-county variation in quality: while the Kolmogorov-Smirnov test for equality of distributions clearly rejects the null that the overall and

Table 2: Relationship Between Quality and Nursing Home Characteristics

	Quality Estimates (s.d.)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2009 Star Ratings (s.d.)	0.0783							0.0487
	(0.0310)							(0.0316)
RN hours per resident day (s.d.)		0.0921						0.199
		(0.0328)						(0.0459)
LPN hours per resident day (s.d.)			0.0977					0.107
			(0.0205)					(0.0309)
CNA hours per resident day (s.d.)				0.09				-0.0346
				(0.0159)				(0.0251)
Deficiencies (s.d.)					-0.0492			-0.0229
					(0.0158)			(0.0133)
For-Profit (s.d.)						-0.0846		-0.0662
						(0.0283)		(0.0255)
Chain (s.d.)							-0.0322	-0.0147
							(0.0277)	(0.0262)
N	10,103	10,118	10,119	10,112	10,121	10,121	10,121	10,089
R-squared	0.009	0.013	0.014	0.012	0.004	0.011	0.002	0.041

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 residents admitted to the nursing home for their first stay over the sample period. Standard errors are clustered by nursing home.

both high-quality and low-quality nursing homes close to each resident. Whether residents are able to take advantage of this variation is the main topic of the next section.

3 Residents' Demand for Nursing Home Quality

In this section, I use my survival-based quality estimates as inputs into a structural model to study residents' demand for quality. I start in section 3.1 by showing evidence that there are unobserved constraints on residents' choice sets due to selective admissions practices by nursing homes (Gandhi 2019), which poses a challenge for demand estimation. In section 3.2, 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 3.3, I present results from the structural demand estimation, and explore the plausibility of various explanations for the findings.

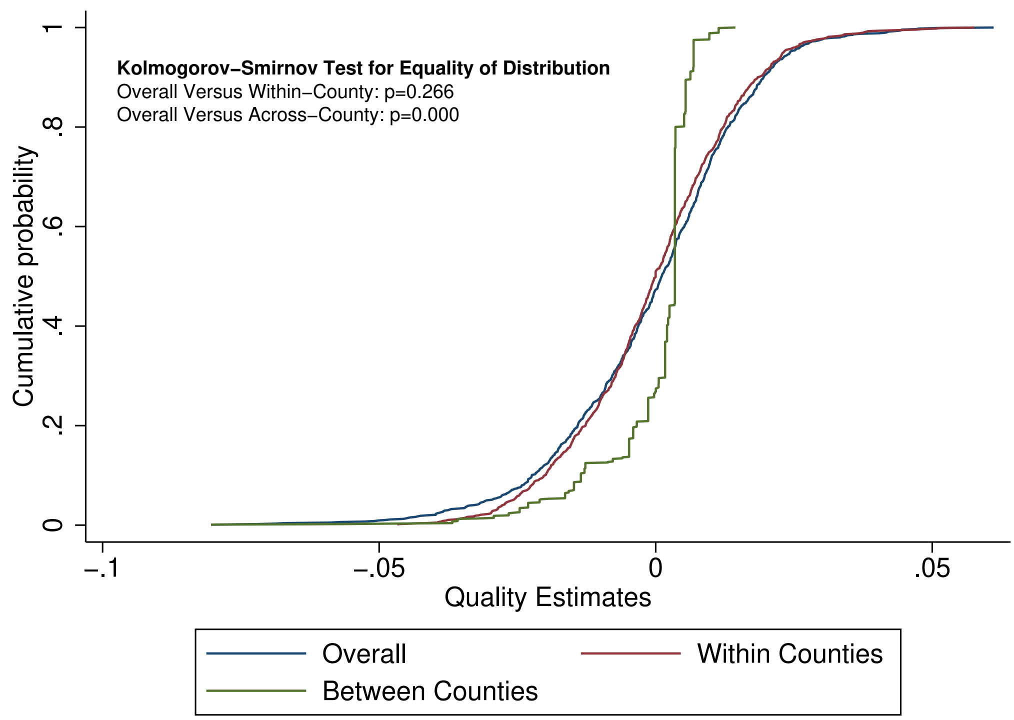
3.1 Selective Admissions by Nursing Homes

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.[64]

across-county distributions of quality are equal, the p-value for the null that the overall and within-county distributions of quality are equal is 0.266.

[64] It is worth noting that non-financial incentives such as taste-based discrimination or altruism may also give rise to selective admissions. For the purposes of this study, the precise sources of these selective admissions practices are less

Figure 3: ECDF of Quality Estimates (Overall, Within Counties, 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, and 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.

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 test two predictions from Gandhi's (2019) model of selective admissions, and briefly summarize the results (for a more detailed description of the tests and additional supporting evidence, see Appendix Section M).

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 based on regressions at the nursing home-day level: conditional on nursing home-month fixed effects, nursing homes admit fewer new residents on days

important, and instead, the focus is on preventing the resulting unobserved choice set constraints from contaminating the demand estimates.

when occupancy is higher than usual, and this finding is not sensitive to the precise definition of occupancy used.[65,66] Moreover, as Appendix Table A.9 shows, this result is robust to alternative measures of new admissions, such as a dummy for any new admissions, and the flow of residents (i.e., admissions minus discharges).

Figure 4 shows support for prediction 2: when nursing homes are closer to capacity, they are less likely to admit Medicaid residents, consistent with the low reimbursement rates associated with Medicaid, and this result holds whether or not we control for other resident characteristics. Appendix Table A.10 shows that this pattern is robust to different measures of occupancy, and Appendix Figure A.7 demonstrates that similar patterns hold for other resident characteristics such as race.

Table 3: Effect of Occupancy on Admissions

	Number of New Residents		
	(1)	(2)	(3)
Lagged 7-Day Avg. Log Occupancy	-0.48		
	(0.0616)		
Lagged 7-Day Avg. Occupancy		-0.00779	
		(0.000415)	
Lagged 7-Day Avg. Occ. Percentile			-0.00163
			(7.19e-05)
Nursing Home-Month Fixed Effects	X	X	X
N	1,103,528	1,103,528	1,103,528

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.

3.2 Structural Model for Demand Estimation

Having documented selective admissions practices by nursing homes, in this subsection I describe a structural demand model and estimation procedure that accounts for unobserved choice set constraints arising from selective admissions.

3.2.1 Identification of Two-Sided Matching Model

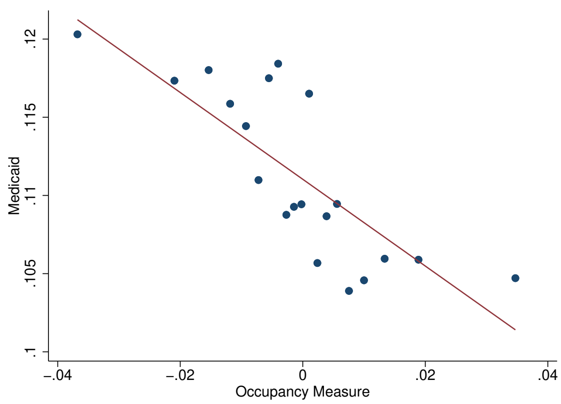
Given that nursing homes have rankings over residents and vice versa, the setting is well-approximated by a many-to-one two-sided matching market: each resident $i \in \mathcal{I}$ is matched to exactly one nursing home, whereas a nursing home $j \in \mathcal{J}$ can be matched with more than one resident.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

[65] I 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). Another advantage of using within-nursing home-month variation is that this measure does not require precise measurement of nursing homes' total capacity. In particular, 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.

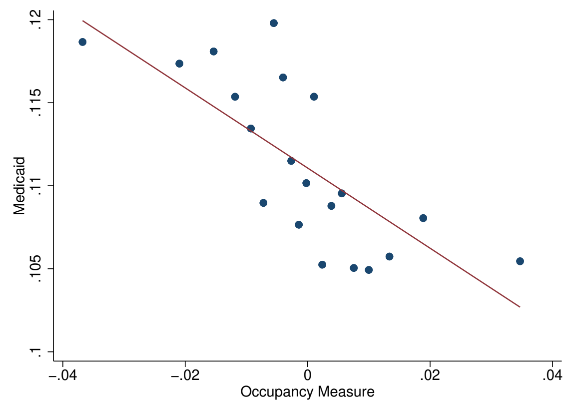
[66] Strictly speaking, prediction 1 only tests whether capacity constraints are relevant and does not directly test for selective admissions (unlike prediction 2). Nonetheless, we can still view prediction 1 as a kind of "first stage" or sanity check for the model of selective admissions.

Figure 4: Bin Scatters of Medicaid Status Against Occupancy

(a) Without Controlling for Other Characteristics



(b) Controlling for Other Characteristics



Notes: 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, and the unit of observation is a resident.

by:

$$\begin{aligned} v_{ij} &= v_j^1(x_i, \zeta_i) - v_j^2(x_i, dist_{ij}), \\ \pi_{ij} &= \pi_j(x_i, \zeta_i, occ_{ij}). \end{aligned} \tag{5}$$

Denote resident characteristics by x_i, and let ζ_i represent resident-specific heterogeneity. In addition, denote distance between resident i and nursing home j by $dist_{ij}$, and let occ_{ij} be a measure of short-term fluctuations in nursing home j's occupancy as 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,[67] which in my case are $dist_{ij}$ and occ_{ij}, respectively. Similar to standard IV, the validity of these instruments relies on a relevance (i.e., first stage) condition and an 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 supply instrument must be excluded from the demand side ($\partial v_{ij}/\partial occ_{ij} = 0$), and the demand instrument must be excluded from the supply side ($\partial \pi_{ij}/\partial dist_{ij} = 0$).*

Relevance of the demand instrument $dist_{ij}$ is demonstrated in the first stage of the distance IV, and similarly, relevance of the supply instrument occ_{ij} was shown when we tested prediction 1 in Table 3. 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

[67] See Appendix E for the full list of the technical assumptions required for identification, as well as a discussion on how they relate to my setting.

fluctuations as the supply instrument. Let us first consider why simply using occupancy as the supply 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 for this instrument). 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 several 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. Third, focusing on within-nursing home-month variation as opposed to simply within-nursing home variation also controls for changes in nursing home capacity over time (e.g., if nursing homes expand or contract).

3.2.2 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:

$$\begin{aligned} v_{ij} &= w_j'\kappa^1 + w_j'\kappa^2 x_i + dist_{ij}'\kappa^{dist} + \epsilon_{ij}, \\ \pi_{ij} &= x_i'\psi^1 + w_j'\psi^2 x_i + occ_{ij}'\psi^{occ} + \omega_{ij}, \end{aligned} \tag{6}$$

where x_i contains resident i's characteristics, w_j denotes nursing home j's characteristics (including quality β_j), and ϵ_{ij} and ω_{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,[68] 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 normalizations, I set the variances of ϵ_{ij} and ω_{ij} to one.

I assume that each resident i considers the set of all nursing homes within 15 miles of her prior address, $\mathcal{J}_i$. Some of these nursing homes may be unwilling to admit her but the identities of these nursing homes (if any) are unobserved by the econometrician. This gives rise to a curse of dimensionality, as the number of possibilities for the set of nursing homes that the resident is actually *able* to choose from is $2^{J_i} - 1$, where $J_i \equiv |\mathcal{J}_i|$.[69] This makes methods such as maximum likelihood that involve directly integrating over each possibility computationally infeasible, unless one is willing to make further assumptions to rule out some possibilities,[70] or restrict the number of nursing homes in

[68]Note that all residents in the sample choose a nursing home by construction, so we cannot estimate the value of the outside option of no nursing home.

[69]The average value of J_i is 50, and in this case there are $2^{50} - 1 > 10^{15}$ possibilities for the set of nursing homes that the resident is actually able to choose from. Moreover, J_i can be larger than 200, in which case there are $2^{200} - 1 > 10^{60}$ possibilities.

[70]For 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 restriction drastically reduces the possibilities for the set of nursing homes that are willing to admit the resident from 2^{J_i} to J_i.

each resident's choice set.[71]

To address the curse of dimensionality, I estimate the model using a Gibbs sampler with data augmentation on the latent variables v_{ij} and π_{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 (ϵ', ω') and the parameters $\theta \equiv (\kappa', \psi')'$,[72] in a way that respects the matching outcomes.[73] Under standard conditions, the draws of θ will eventually converge to the stationary distribution,[74] and I conduct inference using the draws of θ after an initial burn-in period.[75,76] Appendix F describes the full algorithm for the Gibbs sampler in detail.

3.3 Demand Estimation Results

3.3.1 Demand for Quality and Information Frictions

Table 4 presents estimates of the structural demand model described in the previous subsection, focusing on the demand side and leaving the full supply side estimates to Appendix Table A.12 given that they are less relevant for the main discussion.[77] Column 1 contains one of the main results of this paper: the estimate of residents' demand for quality is quantitatively very small. In particular, the demand coefficients on quality β_j and distance $dist_{ij}$ imply that the marginal rate of substitution (MRS) between quality (in percentage points) and distance is:

$$MRS = \frac{d(dist_{ij})}{d\beta_j} = -\frac{\partial v_{ij}/\partial \beta_j \cdot 0.01}{\partial v_{ij}/\partial dist_{ij}} \approx -\frac{0.00564}{-0.16} \approx 0.035.$$

Assuming full information on the part of residents, this MRS estimate implies that in exchange for a one-mile reduction in travel distance, residents are willing to tolerate a $1/0.035 = 28.6$ percentage point higher probability of dying within 90 days (more than 14 times the standard deviation in quality, which is 2 percentage points). Alternatively, if we convert the mortality effects into dollar terms, the

[71]Choosing a smaller radius than 15 miles will reduce the computational burden, but at the expense of excluding residents who choose nursing homes further away, and may thus result in an unrepresentative sample. The ECDF of residents' distances to their chosen nursing homes, shown in Appendix Figure A.9, indicates that more than 80 percent of residents choose a nursing home within 15 miles, but only less than 60 percent of residents choose a nursing home within 5 miles.

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

[73]Thus, denoting the nursing home that resident i is admitted to by $\mu(i)$, in each iteration, we must ensure that $\pi_{i,\mu(i)} \geq 0$, and that the resident prefers $\mu(i)$ to all other nursing homes that she is eligible for (i.e., $v_{i,\mu(i)} \geq v_{ij}$ for all j such that $\pi_{ij} \geq 0$). Similarly, for a nursing home j that is not chosen by resident i, we must ensure that either it is less desirable than her chosen nursing home ($v_{ij} \leq v_{i,\mu(i)}$) or that she is not eligible for it ($\pi_{ij} < 0$).

[74]Specifically, 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).

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

[76]Although the Gibbs sampler is derived following a Bayesian approach, the Bernstein-von Mises theorem implies that the distribution of θ 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.

[77]The supply side estimates should be interpreted as predictors of nursing homes' admission decisions rather than nursing homes' "preferences", due to the possibility of omitted variables bias. Nonetheless, the supply side estimates are generally quite intuitive: when capacity is strained, nursing homes are more likely to admit residents from acute care hospitals (who tend to be covered by Medicare, which has higher reimbursement rates than Medicaid) and less likely to admit black and hispanic residents. However, estimates on other resident characteristics are more difficult to interpret, given that these characteristics tend to be correlated with both expected costs and reimbursements in the same direction.

MRS implies that residents are willing to give up more than $100,000 in exchange for traveling one less mile, a very large amount given that most residents need to make only a single trip to their nursing homes.[78] 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).[79,80]

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).[81]

Moreover, this finding of low demand for quality is not limited to my survival-based quality measure: 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).[82,83] Furthermore, although estimates of residents' demand for skilled staffing are positive, they are quantitatively small and are similar to the estimate of residents' demand for quality in column 1 if we convert staffing levels into dollar terms based on wages.[84]

[78] This conversion requires a few assumptions. First, it is difficult to convert 90-day survival into life expectancy given that I cannot track non-Medicare residents after discharge, so I assume that conditional on surviving at least 90 days after admission, residents go on to live one year on average. This is a somewhat conservative assumption considering that most residents are admitted for rehabilitation, and Liu et al. (2021) find 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. Taking $400,000 as the value of a statistical life year, the implicit willingness-to-pay for a one-mile reduction in distance is given by:

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

[79] The MRS estimates from these other studies imply that residents are willing to tolerate between a $1/8 = 0.125$ to $1/1.8 \approx 0.556$ percentage point higher probability of death in exchange for a one-mile reduction in travel distance.

[80] 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 given that residents typically only need to travel to their nursing homes once. Distance may be more relevant for visiting family members, but most residents will only need to stay in their nursing homes for a relatively short period of time.

[81] I also test whether the demand estimates are sensitive to violations of the assumption that ϵ_{ij} and ω_{ij} are independent. In particular, we may worry that ϵ_{ij} and ω_{ij} are positively correlated, which might be the case if residents make an extra effort to convince nursing homes for which they have strong idiosyncratic preferences to admit them. I find in simulations that if anything, a positive correlation between ϵ_{ij} and ω_{ij} tends to result in an upward bias in the MRS between quality and distance, which would strengthen the main result in this section that residents are not very responsive to quality differences (and that even under a relatively high correlation, the bias is modest).

[82] For the structural estimation, I use deficiencies from inspections originating from complaints, which are likely even more salient to consumers than deficiencies from standard inspections. In addition, I also consider deficiencies per bed as a robustness check and obtain similar results.

[83] One possible explanation for this surprising pattern might be that for-profit nursing homes tend to spend more on advertising than not-for-profit nursing homes. Indeed, motivated by concerns that for-profit nursing homes spend too much of taxpayers' dollars they receive on advertising and executive pay, in 2021, Massachusetts, New Jersey, and New York set additional requirements related to how nursing homes may spend taxpayers' dollars (Jaffe 2021).

[84] Suppose that registered nurses (RNs) make $25 per hour, and that residents require 30 days of care on average (which is conservative relative to the estimated length of stay of 37 days in MedPac 2015). Then, the average cost associated with an additional hour of care per day for one resident is $\$25/\text{hour} \times 24 \text{ hours/day} \times 30 \text{ days} = \$180,000$. The coefficients for residents' demand for RN staffing and distance thus imply a willingness-to-pay for a one-mile reduction in travel distance of:

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

Table 4: Structural Demand Estimates

	(1)	(2)	(3)	(4)
Resident Preferences				
Distance to Nursing Home (in miles)	-0.159	-0.161	-0.160	-0.159
	(0.015)	(0.008)	(0.011)	(0.015)
Quality	0.564		0.257	5.678
	(0.326)		(0.145)	(0.755)
Deficiencies		0.008	0.008	
		(0.002)	(0.002)	
For-Profit Nursing Home		0.128	0.183	
		(0.110)	(0.134)	
RN Hours Per Resident-Day		0.294	0.283	
		(0.026)	(0.032)	
LPN Hours per Resident-Day		0.241	0.259	
		(0.036)	(0.042)	
CNA Hours Per Resident-Day		-0.122	-0.113	
		(0.019)	(0.023)	
Chain		0.152	0.165	
		(0.027)	(0.031)	
Quality x Dementia				-2.424
				(0.270)
Quality x Age				-0.059
				(0.009)
Quality x At Least Bachelor's Degree				1.326
				(0.299)
Quality x Post-Star Ratings				0.560
				(0.199)
Quality x Black				-1.369
				(0.513)
Quality x Hispanic				-2.325
				(0.528)
Quality x Female				-0.014
				(0.220)
Quality x Married				-0.323
				(0.246)
Quality x Lived Alone				-0.046
				(0.245)
Supply Side				
Temporary Fluctuations in log(occupancy)	-7.34	-7.97	-6.95	-7.38
	(1.07)	(1.24)	(1.37)	(1.10)
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. Resident characteristics are included in the supply side of the estimation, but their coefficient estimates are shown separately in the Appendix. A burn-in period corresponding to the first half of the chain was used.

The fact that residents do not seem to take full advantage of publicly available information suggests

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

that information frictions may play a role. As further evidence of this interpretation, column 4 shows that residents with characteristics associated with less information frictions display a greater demand for quality: the coefficients on the interaction terms indicate that residents who are free of dementia, are less advanced in age, or have at least a Bachelor's degree are more responsive to quality differences.[85] 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.[86,87] 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, 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.[88]

3.3.2 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

[85] It is likely that in cases where the resident is cognitively impaired, a surrogate (e.g., a family member or hospital discharge planner) chooses the nursing home in her place. My results on demand heterogeneity already incorporate this channel. One may also be interested in demand heterogeneity *holding the identity of the decision maker constant,* but to the extent that the cognitively impaired resident would have made worse choices than her surrogate, the patterns of demand heterogeneity will be even stronger. Effective design of information interventions also depends on the identity of the decision maker, but this is not the primary objective of the present paper.

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

[87] The magnitude of these differences in demand for quality are sizable relative to the average demand. 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 between residents choosing nursing homes before and after the introduction of star ratings are $12.991 \times 0.059/0.564 \approx 1.36$, $1.326/0.564 \approx 2.35$, and $0.56/0.564 \approx 0.99$ times the average demand for quality respectively.

[88] 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.

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. In particular, using methods from Cheng (2023) on selection on unobservables for discrete choice models, 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 (2019).[89]

A third possible explanation for the low demand estimate is that residents may not directly observe nursing home quality β_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 to make an optimal forecast about β_j.[90] I illustrate with a simple model in Appendix Section N that if this is the case, residents should value nursing home characteristics that are positively correlated with β_j.[91] 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 β_j, which provides evidence against this explanation.[92] 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 do not explain why we observe patterns of demand heterogeneity that are consistent with information frictions in column 3

[89] The proportional selection relationship refers to how much of the variation in the explanatory variable of interest (quality β_j in this case) is explained by the omitted variable relative to the included controls (i.e., observable nursing home characteristics $w_{\sim\beta,j}$ in this setting), as measured by the R-squared from a hypothetical regression of β_j on $w_{\sim\beta,j}$ and the omitted variable. Oster (2019) suggests that in many empirical applications, it is reasonable to assume that the omitted variable explains at most as much of the variation in the explanatory variable of interest as the omitted variable does.

[90] 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 from word-of-mouth recommendations.

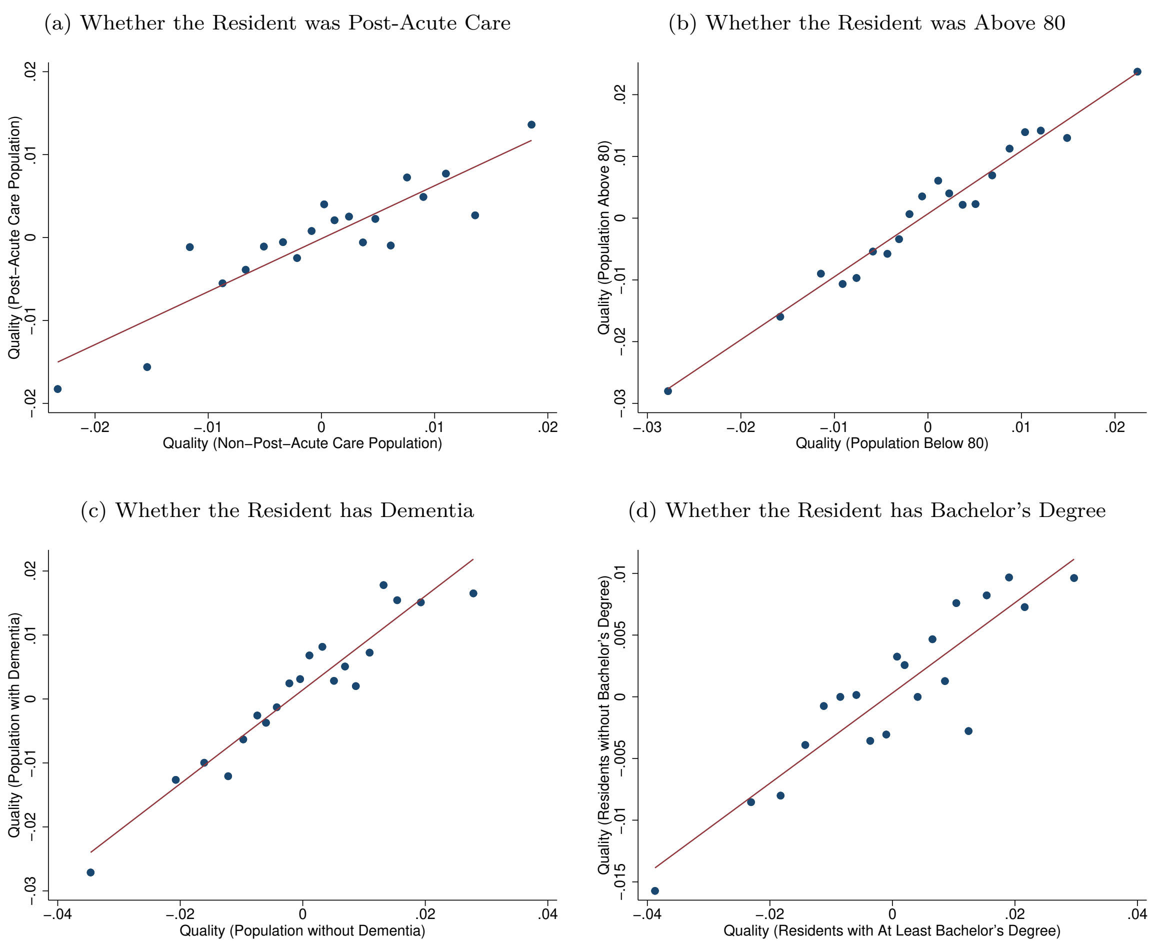
[91] Allowing for estimation noise in my quality measure makes no difference to the model's predictions.

[92] 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 components of nursing home characteristics not including quality $w_{\sim\beta,j}$ against the coefficients from a regression of my quality measure β_j on $w_{\sim\beta,j}$. 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.

of Table 4. So, I next explore potential explanations for demand heterogeneity.

First, there is a possibility that nursing homes specialize in treating different types of residents.[93] 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. This suggests 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.[94]

Figure 5: Bin Scatters of Quality Estimates for Different Subpopulations



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 then plot these two sets of coefficients against each other.

[93] For instance, it may be the case that some nursing homes specialize in providing care for long-stay residents, but are ill-equipped to care for short-stay residents.

[94] While 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.[95] Instead, the results in Appendix Table A.16 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.[96] While such an explanation is certainly plausible, 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.[97]

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.

4 Policy Counterfactuals

In this section, I quantify the costs of the information frictions documented previously via counterfactual simulations. In section 4.1, I describe my framework for modeling information interventions, and discuss the various treatment effect mechanisms at work in this setting. In section 4.2, I introduce a competing risks model for resident deaths and discharges that is necessary for my counterfactual simulations. In section 4.3, I present simulation results which quantify both the short- and long-run effects of eliminating information frictions on resident mortality as well as distributional consequences. Finally, in section 4.4, I discuss these results in the context of nursing home policies more generally.

4.1 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). Alternatively, we can view the costs of information frictions as the treatment effect of an intervention that eliminates information frictions, and for expositional purposes I sometimes refer to these costs as treatment effects. I remain largely agnostic about the precise ways by which a policy

[95] 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.

[96] 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 the quality adjustment (Gold, Stevenson, and Fryback 2002).

[97] 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. This is still far smaller than demand estimates from previous studies focusing on hospital settings ($0.035 \times 1.5 = 0.0525$ in my setting after the quality adjustment, compared to estimates ranging from 1.8–8 from hospital settings).

intervention may be able to achieve this (be it through information provision, or aid in interpreting and processing this information), given that this requires a detailed analysis of past policies that is beyond the scope of this paper.

4.1.1 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 brevity, 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, 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 resembling those of consumers in other healthcare settings.

For the short-run analysis, I simulate outcomes based on these counterfactual preferences, holding nursing home quality constant.[98] 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 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.[99] 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.[100] This provides an alternative way to interpret the magnitude of costs due to information frictions, and sheds light on the potential efficacy of other policies if effective information interventions are 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/exits). Because these restrictions on nursing home behavior become increasingly tenuous over

[98]For robustness, I run additional simulations under the assumption that the idiosyncratic utility shocks ϵ_{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.

[99]The structural demand estimation in the previous section 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 relative differences of nursing home quality at each point in time that matters for my analysis of residents' 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) for my demand estimation (which uses observations from 2008–2010) is unlikely to cause issues, and has the advantage of being more precisely estimated.

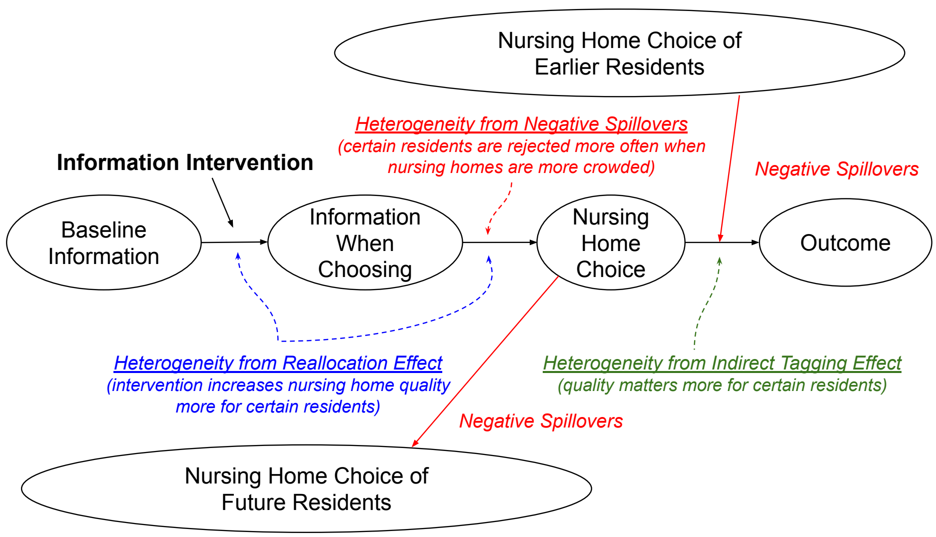
[100]Specifically, I assume that under the minimum standard mandate, the lowest-quality nursing homes are forced to improve their quality to a minimum acceptable level (corresponding to the 10th percentile of the quality distribution). To simulate the effect of a pay-for-performance scheme, I use estimates from Werner, Konetzka, and Polsky (2013) as a benchmark.

longer time horizons, I restrict my counterfactual simulations to the year 2009.

4.1.2 Treatment Effect Mechanisms

Given that there are multiple market imperfections (specifically, information frictions and selective admissions), the theory of the second best implies that eliminating information frictions alone may not necessarily improve resident outcomes. Distributional consequences are likewise ambiguous a priori. As a framework for understanding the various forces at work in this setting, I consider a simple model of an information intervention's effect on resident outcomes in Appendix Section H, but here I summarize the main insights from this model with the aid of Figure 6.[101]

Figure 6: Treatment Effect Mechanisms



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.[102] The magnitude of these negative spillovers relative to improvements in choice quality determines 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

[101] This model only considers the partial equilibrium effect of an information intervention (i.e., it does not account for potential endogenous quality adjustments by nursing homes).

[102] These negative spillover effects also imply that the Stable Unit Treatment Value Assumption (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.

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 *reallocation effect*, and in this instance it is likely to be progressive.[103]

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 tend to be the types of residents that nursing homes are less willing to admit. Therefore, negative spillovers from increased crowding at higher-quality nursing homes are likely to fall disproportionately on this group of residents.

4.2 Simulation Procedure

4.2.1 Competing Risks Model of Deaths and Discharges

While the structural model in section 3 provides a way to simulate nursing home admissions that accounts for selective admissions practices, 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,[104] I account for this codependency using a model of competing risks (specifically, using cause-specific hazards that depend on resident and nursing home characteristics), which I estimate semiparametrically by maximizing a modified partial likelihood.[105] 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 conservative treatment effect estimates.[106] For details about the competing risks model and the procedure for simulating death and discharge times, see Appendix Section G.

[103] I refer to an information intervention as progressive if its effect is increasing in baseline information frictions (assuming that the effect is positive).

[104] 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.

[105] 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.

[106] 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.

4.2.2 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.[107]

4.2.3 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 3: 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 times and causes 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 the simulation, I update nursing home occupancies for the start of the next day based on admissions and exits that occur on the current day of the simulation. For the full simulation algorithm, see Appendix Section G.4.

4.3 Simulation Results

4.3.1 Average Treatment Effects

Figure 7 shows the 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. Even under my most conservative assumption about the MRS,[108] 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, the dashed black line shows that the short-run effects will be even larger if we assume that the idiosyncratic utility shocks ϵ_{ij} are also mistakes due to information frictions. Additionally, in Appendix Table A.17, I explore the plausibility of my conservative assumption that discharge and mortality shocks are perfectly correlated within resident across potential nursing homes, and the results indicate that this assumption is reasonable as well as that the alternative assumption yields even larger

[107]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).

[108]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.

treatment effects.[109]

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 red and blue horizontal lines respectively. 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,[110] 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 demonstrates that the simulated long-run effect of eliminating information frictions is a 44 percent reduction in nursing home deaths under the most conservative MRS assumption, which is much larger than the other simulated effects. Appendix Table A.18 shows that the exact magnitude of the long-run effect depends on the functional forms for the demand and cost functions in the model of quality competition between nursing homes. Nonetheless, the simulated long-run effects are several times larger than the short-run effect in all functional forms considered, suggesting that supply side responses have the potential to substantially amplify the already sizable demand side effects of an information intervention.

4.3.2 Distributional Consequences

In Figure 8, I provide 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, but the fact that the surface mostly lies above the base (which corresponds to zero treatment effect) suggests that the intervention might be a Pareto improvement in the ex ante sense (i.e., based on expected treatment effects).

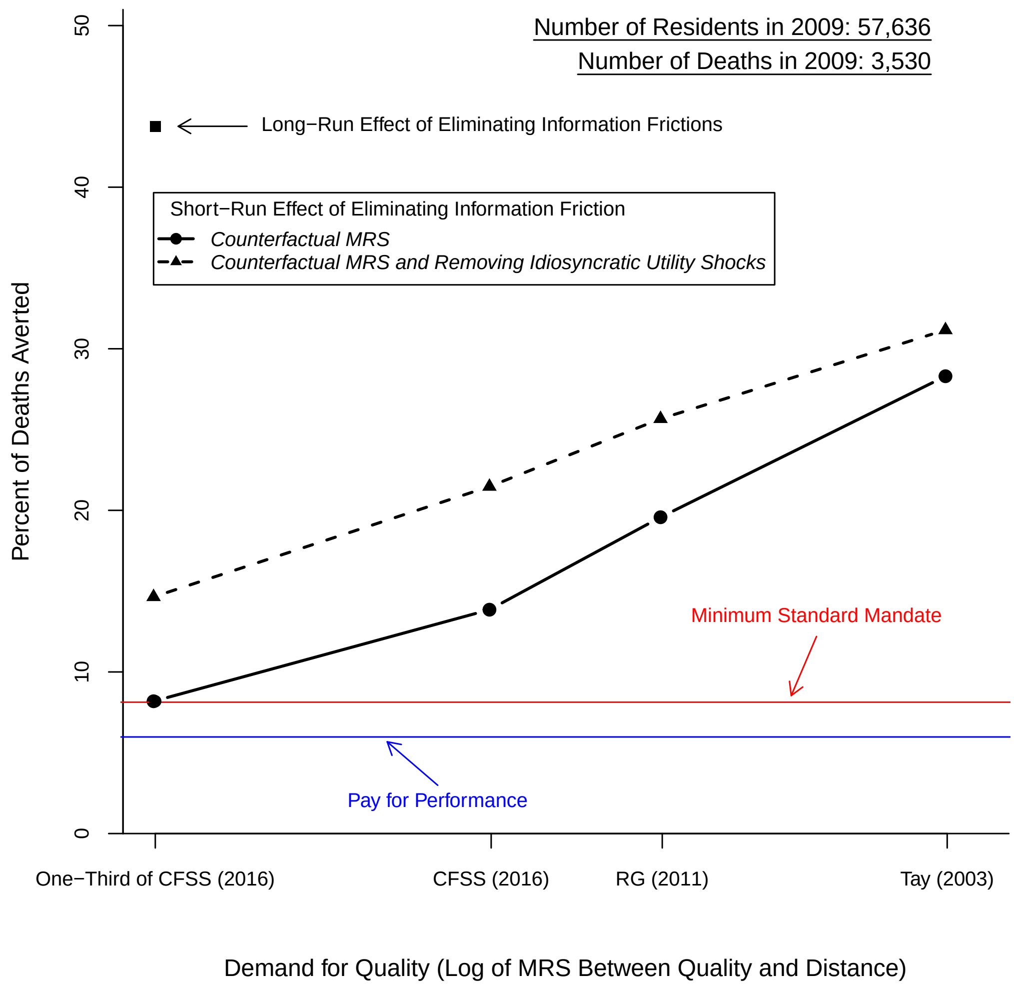
To gain a better understanding of the types of residents who benefit more from the intervention, I study treatment effect heterogeneity by resident characteristics in Appendix Figure A.18. The results indicate that the intervention has progressive distributional effects: we observe that residents from 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.[111] Consistent with these results, I find that one standard deviation greater information

[109]Specifically, in Appendix Table A.17 I compare my baseline assumption that mortality shocks are perfectly correlated within resident across nursing homes to an alternative assumption of independent shocks (and likewise for discharge shocks). The quality estimates are based on 90-day mortality, whereas the simulation considers mortality (or discharge) until the end of 2009. Given that residents arrive throughout the year 2009, average length of time from admission to the end of the simulation period for residents in the simulations is roughly 6 months. Hence, we would expect the effect of an information intervention on mortality to be somewhat larger than the effect on quality of residents' chosen nursing homes, but not an order of magnitude larger. This is indeed the case for the simulations under the baseline assumption of perfectly correlated shocks, whereas the simulations under the assumption of independent shocks results in much larger effects on mortality, which suggests that the latter assumption yields an overestimate of the effect of eliminating information frictions.

[110]Increasing the quality threshold for the minimum standard mandate even further yields much smaller improvements in resident outcomes in the simulations, given that setting the minimum threshold at the 10th percentile already eliminates the long left tail of the bell-shaped distribution of quality estimates.

[111]Appendix Figure A.18 also explores treatment effect heterogeneity by resident characteristics for the simulated long-run effects of eliminating information frictions, as well as the effects of a minimum standard mandate and pay-for-

Figure 7: Main Simulation Results



Notes: The figure shows the percent of nursing home deaths averted in the simulations under various policy counterfactuals. The labels for the (log of) MRS between quality and distance corresponds to different estimates from the literature. 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 MRS estimate from CFSS by three because the quality measure in CFSS was based on 30-day mortality, whereas the quality measure in the present paper is based on 90-day mortality.

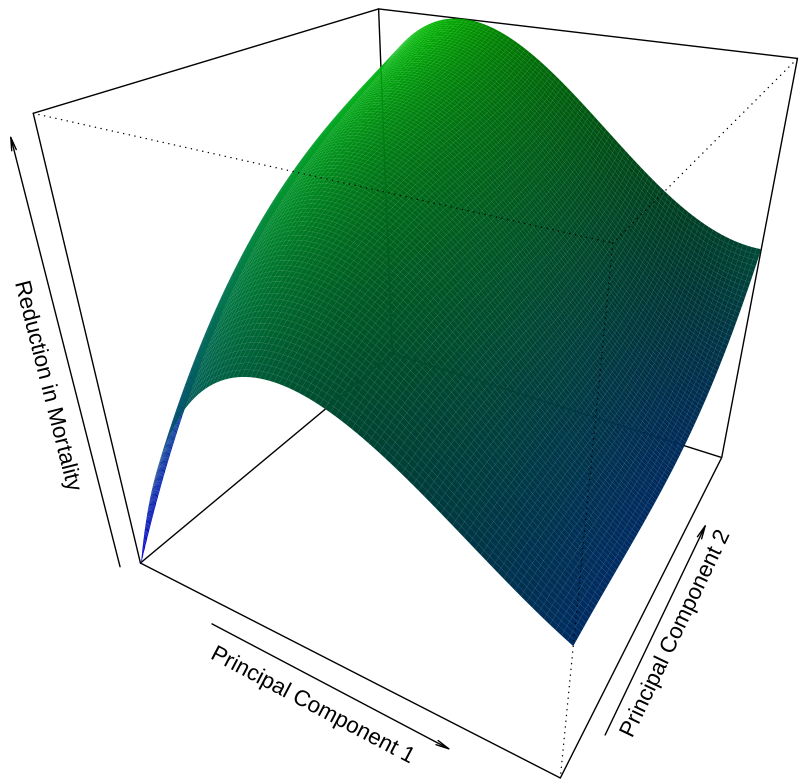
frictions at baseline is associated with 12.5 percent greater treatment effect relative to the average treatment effect.[112]

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

performance policy.

[112]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).

Figure 8: Reduction in Mortality as a Function of the Principal Components of Resident Characteristics



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.

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).[113] Interestingly, the decomposition also reveals that the indirect tagging effect is more important than the reallocation effect in explaining the treatment effect heterogeneity.[114,115]

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-

[113]Specifically, negative spillovers explain less than 5 percent of the treatment effect and its heterogeneity compared to the reallocation effect.

[114]In particular, the indirect tagging effect is roughly 1.6 times as important as the reallocation effect.

[115]A potential reason for this finding of limited externalities is that these simulations cover only a relatively short time frame of a single year: 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. On the other hand, unmodeled expansion and entry of high-quality nursing homes (and contraction and exit of low-quality nursing homes) are likely to mitigate these capacity strains over longer time horizons. Developing a fully dynamic model that incorporates these forces is beyond the scope of this paper.

quality nursing homes located in the vicinity of residents' prior addresses. Thus, it is rare that a resident is rejected by all of them and forced to go to a low-quality nursing home.

4.4 Discussion

The simulated short-run effects of information frictions are likely a conservative lower bound for a number of reasons. First, we observed that allowing endogenous quality adjustments by nursing homes may lead to larger reductions in mortality, and there is reason to believe that other unmodeled dynamics such as entry or expansion (respectively, exit or contraction) of high-quality (low-quality) nursing homes may lead to further improvements in resident outcomes. 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, under alternative modeling assumptions (e.g., about the correlation between mortality and discharge shocks, or assuming 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,[116] 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.[117]

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 suggest that it may also have increased disparities.[118,119] 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 Gopal 2016, 2018; Ryskina, Konetzka, and Werner 2018).[120] Finally, the five-star ratings system is a rather crude measure of quality: although Table 2 shows that my quality estimates are positively 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.[121]

[116] For example, there is work studying the effects of information interventions on enrollment in various public programs such as the Earned Income Tax Credit (Bhargava and Manoli 2015; Guyton et al. 2017), Social Security Disability Insurance (Armour 2018), and SNAP (Finkelstein and Notowidigdo 2019), as well as in education settings (Hastings and Weinstein 2008; Bettinger, Long, Oreopoulos, and Sanbonmatsu 2012).

[117] 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).

[118] 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.

[119] Similarly, Konetzka et al. (2015) find that the five-star ratings system exacerbated disparities in quality by payer source.

[120] 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.

[121] 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."

Quality measures such as the risk-adjusted survival rate estimated in this paper (if made public) can potentially provide residents with a better way to evaluate nursing home quality than the five-star ratings system for at least two reasons.[122] First, it provides more detailed information about nursing home quality than measures such as the five-star ratings system, and can be validated using quasi-experimental methods (e.g., the distance-based IV strategy described in section 2). Second, this type of quality measure may be harder for nursing homes to game, given that mortality is rarely misreported, and the complexity of the risk adjustment formula makes it difficult for nursing homes to determine how to manipulate baseline characteristics used in the risk adjustment to obtain a higher quality score.[123,124]

The specifics of designing an effective information intervention depend on the nature of information frictions, which for example, may be due to lack of relevant information or an inability to process this information (Handel and Schwartzstein 2018). The fact that residents do not consistently value publicly observable nursing home characteristics that are associated with higher quality suggests that the first explanation plays a role,[125] 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 the cognitive impairments that many residents suffer from, and uncertainty over the identity of the individual who initially chose the nursing home.

In addition, information interventions can be targeted at different stakeholders, and take very different forms. In many cases, residents and/or their family members have the ultimate say in nursing home choice, so it seems natural to target them through information outreach interventions. On the other hand, (in the case of post-acute care residents) hospital discharge planners are well-placed to provide (future) residents (and their families) with information on nursing home quality, and planners are also likely to be more informed about nursing home quality than residents given the number of discharges they have overseen. Yet, they often do not provide residents with this information due to legal risks associated with choice requirement laws (Tyler et al. 2017).[126] Hence, an intervention targeted at hospital discharge planners may instead involve reforms to choice requirement laws that make it easier for planners to share quality information with residents.

Finally, instead of information interventions, policymakers have largely used regulations in their

[122] In practice, the CMS is unlikely to risk-adjust based on protected characteristics such as race and gender. Nevertheless, the quality estimates I obtain with or without controlling for protected characteristics are highly correlated: a bivariate regression between the two sets of quality estimates yields an R-squared of 0.98.

[123] Interestingly, if publishing risk-adjusted quality measures does in fact improve resident choice, accompanying changes in selection patterns may lead to selection bias in future quality estimates (Oster 2019). To avoid this, quality estimates should periodically be validated to ensure they remain informative (e.g., using quasi-experimental methods).

[124] A potential concern however 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).

[125] 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.

[126] For example, Tyler et al. (2017) 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."

attempts to improve outcomes for nursing home residents.[127] However, despite substantial increases in nursing home regulations over the past few decades,[128] 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.[129,130]

5 Conclusion

In this paper, I study demand for quality in the Californian 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, are not responsive to quality differences. The costs of these information frictions are substantial, accounting for at least 8–28 percent of nursing home deaths, even before considering potential supply side responses by nursing homes.

The finding that information frictions play an important role naturally leads to a number of follow-up questions. For example, 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 instance, in ongoing work, I find that other than in-group preferences, discrimination, and location, information frictions are another contributory factor to the disproportionate concentration of black residents in low-quality nursing homes.

[127] In fact, 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.

[128] Examples include federal regulations over reporting requirements and the use of physical or chemical restraints, as well as state regulations over minimum staffing standards.

[129] An example of this is the introduction of payroll-based journal reporting of nursing home staffing in 2016. This 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).

[130] Nursing homes often argue that they are already operating on thin margins, and that additional compliance costs may force them to shut down. Such claims are difficult to verify, considering the opacity of nursing homes' finances due to practices such as related party transactions (Cenziper, Jacobs, Crites, and Englund 2020).

References

▶**Abaluck, Jason, Mauricio M. Caceres Bravo, Peter Hull, and Amanda Starc.** 2021. "Mortality effects and choice across private health insurance plans." *The Quarterly Journal of Economics* 136 (3): 1557–1610.

Abaluck, Jason, and Giovanni Compiani. 2020. "A Method to Estimate Discrete Choice Models that is Robust to Consumer Search." Working paper.

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

▶**Abaluck, Jason, and Jonathan Gruber.** 2016. "Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance." *American Economic Review* 106 (8): 2145–2184.

▶**Agarwal, Nikhil, and Paulo Somaini.** 2022. "Demand Analysis under Latent Choice Constraints." Working Paper.

▶**Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber.** 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1): 151–184.

▶**Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters.** 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation." *The Quarterly Journal of Economics* 132 (2): 871–919.

Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. 2021. "Race and the Mismeasure of School Quality." National Bureau of Economic Research Working Paper 29608.

▶**Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin.** 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–455.

▶**Armour, Philip.** 2018. "The Role of Information in Disability Insurance Application: An Analysis of the Social Security Statement Phase-In." *American Economic Journal: Economic Policy* 10 (3): 1–41.

▶**Aron-Dine, Liran Einav, Amy Finkelstein, and Mark Cullen.** 2015. "Moral Hazard in Health Insurance: Do Dynamic Incentives Matter?" *The Review of Economics and Statistics* 97 (4): 725–741.

▶**Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen.** 2014. "Inference on Treatment Effects after Selection among High-Dimensional Controls." *Review of Economic Studies* 81 (2): 608–650.

▶**Bettinger, Eric P., Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu.** 2012. "The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment." *The Quarterly Journal of Economics* 127 (3): 1205–1242.

▶**Bhargava, Saurabh, and Dayanand Manoli.** 2015. "Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment." *American Economic Review* 105 (11): 3489–3529.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2022. "Revisiting Event Study Designs: Robust and Efficient Estimation." Working paper.

▶**Breslow, Norman E.** 1972. "Contribution to the discussion on the paper by DR Cox. Regression and life tables." *JR Statist. Soc. B* 34: 216–217.

▶**Brooks, Stephen P., and Andrew Gelman.** 1998. "General methods for monitoring convergence of iterative simulations." *Journal of computational and graphical statistics* 7 (4): 434–455.

▶**Brown, Jeffrey R., and Amy Finkelstein.** 2007. "Why is the market for long-term care insurance so small?" *Journal of Public Economics* 91 (10): 1967–1991.

▶**Brown, Jeffrey R., and Amy Finkelstein.** 2008. "The Interaction of Public and Private Insurance: Medicaid and the Long-Term Care Insurance Market." *American Economic Review* 98 (3): 1083–1102.

▶**Brown, Jeffrey R., and Amy Finkelstein.** 2011. "Insuring Long-Term Care in the United States." *Journal of Economic Perspectives* 25 (4): 119–142.

▶**Callaway, Brantly, and Pedro HC Sant'Anna.** 2021. "Difference-in-differences with multiple time periods." *Journal of Econometrics* 225 (2): 200–230.

▶**Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik.** 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs". *Econometrica* 82 (6): 2295–2326.

Card, David. 1993. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." National Bureau of Economic Research Working Paper 4483.

▶**Card, David, and Stefano DellaVigna.** 2020. "What do Editors Maximize? Evidence from Four Economics Journals." *The Review of Economics and Statistics* 102 (1): 195–217.

Card, David, Alessandra Fenizia, and David Silver. 2019. "The Health Impacts of Hospital Delivery Practices." National Bureau of Economic Research Working Paper 25986.

▶**Castle, Nicholas G., and Jamie C. Ferguson.** 2010. "What Is Nursing Home Quality and How Is It Measured?" *The Gerontologist* 50 (4): 426–442.

Centers for Medicare & Medicaid Services. 2000–2010. "Cost Reports." Department of Health and Human Services. https://data.cms.gov/provider-compliance/cost-report/skilled-nursing-facility-cost-report.

Centers for Medicare & Medicaid Services. 2000–2010. "Health Deficiencies." Department of Health and Human Services. https://data.cms.gov/provider-data/dataset/r5ix-sfxw.

Centers for Medicare & Medicaid Services. 2000–2010. "Provider Information." Department of Health and

Human Services. https://data.cms.gov/provider-data/dataset/4pq5-n9py.

Cenziper, Debbie, Joel Jacobs, Alice Crites, and Will Englund. 2020. "Profit and pain: How California's largest nursing home chain amassed millions as scrutiny mounted." *The Washington Post*, December 31. https://www.washingtonpost.com/business/2020/12/31/brius-nursing-home/ (accessed January 1, 2021).

▶**Chan, David C., Matthew Gentzkow, and Chuan Yu.** 2022. "Selection with variation in diagnostic skill: Evidence from radiologists." *The Quarterly Journal of Economics* 137 (2): 729–783

Chen, Yiwei. 2018. "User-Generated Physician Ratings — Evidence from Yelp." Working paper.

Cheng, Alden. 2023. "Selection on Unobservables in Discrete Choice Models." Working paper.

▶**Comondore, Vikram R, P J Devereaux, Qi Zhou, Samuel B Stone, Jason W Busse, Nikila C Ravindran, Karen E Burns, Ted Haines, Bernadette Stringer, Deborah J Cook, Stephen D Walter, Terrence Sullivan, Otavio Berwanger, Mohit Bhandari, Sarfaraz Banglawala, John N Lavis, Brad Petrisor, Holger Schünemann, Katie Walsh, Neera Bhatnagar, Gordon H Guyatt.** 2009. "Quality of care in for-profit and not-for-profit nursing homes: systematic review and meta-analysis." *Bmj* 339.

Cooper, Zack, Joseph J. Doyle Jr., John A. Graves, and Jonathan Gruber. 2022. "Do Higher-Priced Hospitals Deliver Higher-Quality Care?" National Bureau of Economic Research Working Paper 29809.

▶**De Chaisemartin, Clément, and Xavier d'Haultfoeuille.** 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review* 110 (9): 2964–96.

▶**Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins.** 2018. "Double/Debiased Machine Learning for Treatment and Structural Parameters." *The Econometrics Journal* 21 (1): C1–C68.

▶**Chetty, Raj, John N. Friedman, and Jonah E. Rockoff.** 2014. "Measuring the Impacts of Teachers I: Evaluating bias in Teacher Value-Added Estimates." *American Economic Review* 104 (9): 2593–2632.

▶**Deryugina, Tatyana, and David Molitor.** 2020. "Does When You Die Depend on Where You Live? Evidence from Hurricane Katrina. *American Economic Review* 110 (11): 3602–3633.

▶**Donabedian, Avedis.** 1985. "Twenty years of research on the quality of medical care: 1964-1984." *Evaluation and the Health Professions* 8: 243–265.

▶**Doyle, Joseph, John Graves, Jonathan Gruber, and Samuel Kleiner.** 2015. "Measuring Returns to Hospital Care: Evidence from Ambulance Referral Patterns." *Journal of Political Economy* 123 (1): 170–214.

▶**Dranove, David, Daniel Kessler, Mark McClellan, and Mark Satterthwaite.** 2003. "Is More Information Better? The Effects of "Report Cards" on Health Care Providers." *Journal of Political Economy* 111 (3): 555–588.

▶**Einav, Liran, Amy Finkelstein, and Neale Mahoney.** 2018. "Provider incentives and healthcare costs: Evidence from long-term care hospitals." *Econometrica* 86 (6): 2161–2219.

Einav, Liran, Amy Finkelstein, and Neale Mahoney. 2022. "Producing Health: Measuring Value Added of Nursing Homes." Working Paper.

▶**Einav, Liran, Amy Finkelstein, Sendhil Mullainathan, and Ziad Obermeyer.** 2018. "Predictive modeling of US health care spending in late life." *Science* 360 (6396): 1462–1465.

▶**Eliason, Paul J., Paul L.E. Grieco, Ryan C. McDevitt, and James W. Roberts.** 2018. "Strategic patient discharge: The case of long-term care hospitals." *American Economic Review* 108 (11): 3232–65.

▶**Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams.** 2021. "Place-Based Drivers of Mortality: Evidence from Migration." *American Economic Review* 111 (8): 2697–2735.

▶**Finkelstein, Amy, and Matthew J. Notowidigdo.** 2019. "Take-up and targeting: Experimental evidence from SNAP." *The Quarterly Journal of Economics* 134 (3): 1505–1556.

Fletcher, Jason M., Leora I. Horwitz, and Elizabeth Bradley. 2014. "Estimating the Value Added of Attending Physicians on Patient Outcomes." National Bureau of Economic Research Working Paper 20534.

Gandhi, Ashvin. 2019. "Picking your patients: Selective admissions in the nursing home industry." Working Paper.

Gandhi, Ashvin, YoungJun Song, and Prabhava Updrashta. 2022. "Private equity, consumers, and competition: Evidence from the nursing home industry." Working Paper.

▶**Geng, Fangli, David G. Stevenson, and David C. Grabowksi.** 2019. "Daily Nursing Home Staffing Levels Highly Variable, Often Below CMS Expectations." *Health Affairs* 38 (7): 1095–1100.

▶**Gold, Marthe R., David Stevenson, and Dennis G. Fryback.** 2002. "HALYS AND QALYS AND DALYS, OH MY: Similarities and Differences in Summary Measures of Population Health." *Annual Review of Public Health* 23: 115–34.

▶**Grabowski, David C., Jonathan Gruber, and Joseph J. Angelelli.** 2008. "Nursing Home Quality as a Common Good." *The Review of Economics and Statistics* 90 (4): 754–764.

▶**Guo, Audrey, and Jonathan Zhang.** 2019. "What to expect when you are expecting: Are health care consumers forward-looking?" *Journal of Health Economics* 67: 1–16.

Gupta, Atul, Sabrina T. Howell, Constantine Yannelis, and Abhinav Gupta. 2021. "Does Private Equity Investment in Healthcare Benefit Patients? Evidence from Nursing Homes." National Bureau of Economic Research Working Paper 28474.

▶**Guyton, John, Pat Langetieg, Day Manoli, Mark Payne, Brenda Schafer and Michael Sebastiani.** 2017. "Reminders and recidivism: using administrative data to characterize nonfilers and conduct EITC outreach."

American Economic Review 107 (5): 471–75.

▸**Hackmann, Martin B.** 2019. "Incentivizing Better Quality of Care: The Role of Medicaid and Competition in the Nursing Home Industry." *American Economic Review* 109 (5): 1684–1716.

Hackmann, Martin B., R. Vincent Pohl, and Nicolas R. Ziebarth. 2020. "Patient Versus Provider Incentives in Long-Term Care." Working Paper.

▸**Han, Xu, Niam Yaraghi, and Ram Gopal.** 2016. "Five-star ratings for sub-par service: Evidence of inflation in nursing home ratings." *Governance Studies at Brookings*: 1–22.

▸**Han, Xu, Niam Yaraghi, and Ram Gopal.** 2018. "Winning at all costs: Analysis of inflation in nursing homes' rating system." *Production and Operations Management* 27 (2): 215–233.

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

▸**Handel, Benjamin R., and Jonathan T. Kolstad.** 2015. "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare" *American Economic Review* 105 (8): 2449–2500.

Handel, Benjamin R., Jonathan Kolstad, Thomas Minten, and Johannes Spinnewijn. 2021. "The Social Determinants of Choice Quality: Evidence from Health Insurance in the Netherlands." Working Paper.

▸**Handel, Benjamin R., and Joshua Schwartzstein.** 2018. "Frictions or Mental Gaps: What's Behind the Information We (Don't) Use and When Do We Care?" *Journal of Economic Perspectives* 32 (1): 155–178.

▸**Hastings, Justine S., and Jeffrey M. Weinstein.** 2008. "Information, school choice, and academic achievement: Evidence from two experiments." *The Quarterly Journal of Economics* 123 (4): 1373–1414.

▸**Hull, Peter.** 2018. "Estimating Hospital Quality with Quasi-experimental Data." Working Paper.

▸**Hurd, Michael D., Pierre-Carl Michaud, and Susann Rohwedder.** 2017. "Distribution of lifetime nursing home use and of out-of-pocket spending." *PNAS* 104 (37): 9838–9842.

▸**Imbens, Guido, and Karthik Kalyanaraman.** 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *The Review of Economic Studies* 79 (3): 933–959.

LTCFocus. 2000–2010. "LTCFocus Public Use Data." Sponsored by the National Institute on Aging (P01 AG027296) through a cooperative agreement with the Brown University School of Public Health. https://doi.org/10.26300/h9a2-2c26

Jaffe, Susan. 2021. "3 States to Limit Nursing Home Profits in Bid to Improve Care." *Kaiser Health News*, October 25. https://susanjaffe.org/2021/10/25/https-khn-org-news-article-3-states-limit-nursing-home-profits-in-bid-to-improve-care/ (accessed July 6, 2022).

▸**Kahneman, Daniel, and Amos Tversky.** 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–292.

▸**Kolesár, Michael, and Christophe Rothe.** 2018. "Inference in Regression Discontinuity Designs with a Discrete Running Variable." *American Economic Review* 108 (8): 2277–2304.

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

▸**Konetzka, R. Tamara, David C. Grabowski, Marcelo Coca Perraillon, and Rachel M. Werner.** 2015. "Nursing Home 5-Star Rating System Exacerbates Disparities In Quality, By Payer Source." *Health Affairs* 34 (5): 819–827.

▸**Konetzka, R. Tamara, and Marcelo Coca Perraillon.** 2016. "Use Of Nursing Home Compare Website Appears Limited By Lack Of Awareness And Initial Mistrust Of The Data". *Health Affairs* 35 (4): 706–713.

▸**Konetzka, Tamara R., Daniel Polsky, and Rachel M. Werner.** 2013. "Shipping out instead of shaping up: Rehospitalization from nursing homes as an unintended effect of public reporting." *Journal of Health Economics* 32 (2): 341–352.

▸**Lee, David S., and David Card.** 2008. "Regression discontinuity inference with specification error." *Journal of Econometrics* 142 : 655–674.

▸**Liu, Enwu, Maggie Killington, Ian D. Cameron, Raymond Li, Susan Kurrle, and Maria Crotty.** 2021. "Life expectancy of older people living in aged care facilities after a hip fracture." *Scientific Reports* 11 (1): 1–9.

▸**Mullainathan, Sendhil, and Ziad Obermeyer.** 2022. "Diagnosing Physician Error: A Machine Learning Approach to Low-Value Health Care." *The Quarterly Journal of Economics* 137 (2): 679–727.

National Center for Health Statistics. 2017. "Nursing Home Care." Centers for Disease Control and Prevention. https://www.cdc.gov/nchs/fastats/nursing-home-care.htm (accessed: July 16, 2023).

Novotney, Amy. 2020. "More than 6 in 10 Americans now say they would rather die than live in nursing home: survey." *McKnight's Business Daily*, December 9. https://www.mcknightsseniorliving.com/home/news/business-daily-news/more-than-6-in-10-americans-now-say-they-would-rather-die-than-live-in-nursing-home-survey/ (accessed July 13, 2023).

▸**Oster, Emily.** 2019. "Unobservables Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37 (2): 187–204.

Rafiei, Yasmin. 2022. "When Private Equity Takes Over a Nursing Home." *The New Yorker*, August 25. https://www.newyorker.com/news/dispatch/when-private-equity-takes-over-a-nursing-home (accessed August 26, 2022).

▶**Romley, John A., and Dana P. Goldman.** 2011. "How Costly Is Hospital Quality? A Revealed-Preference Approach." *Journal of Industrial Economics* 59 (4): 578–608.

▶**Ryskina, Kira L., R. Tamara Konetzka, and Rachel M. Werner.** 2018. "Association between 5-star nursing home report card ratings and potentially preventable hospitalizations." *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* 55: 1–14.

▶**Schelling, Thomas C.** 1971. "Dynamic models of segregation." *Journal of Mathematical Sociology* 1 (2): 143–186.

▶**Shin, Juh Hyun, and Yvonne Scherer.** 2009. "Advantages and Disadvantages of Using MDS Data in Nursing Research." *Journal of Gerontological Nursing* 35 (1): 7–17.

Silver-Greenberg, and Robert Gebeloff. 2021. "Maggots, Rape and Yet Five Stars: How U.S. Ratings of Nursing Homes Mislead the Public." *The New York Times*, March 13. https://www.nytimes.com/2021/03/13/business/nursing-homes-ratings-medicare-covid.html (accessed March 13, 2021).

▶**Sun, Liyang, and Sarah Abraham.** 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics* 225 (2): 175–199.

Thomas, Katie, Robert Gebeloff, and Jessica Silver-Greenberg. 2021. "Phony Diagnoses Hide High Rates of Drugging at Nursing Homes." *The New York Times*, September 11. https://www.nytimes.com/2021/09/11/health/nursing-homes-schizophrenia-antipsychotics.html (accessed June 22, 2022).

▶**Tyler, Denise A., Emily A. Gadbois, John P. McHugh, Renée R. Shield, Ulrika Winblad, and Vincent Mor.** 2017. "Patients Are Not Given Quality-Of-Care Data About Skilled Nursing Facilities When Discharged From Hospitals." *Health Affairs* 36 (8): 1385–1391.

▶**Vats, Dootika, James M. Flegal, and Galin L. Jones.** 2019. "Multivariate output analysis for Markov chain Monte Carlo." *Biometrika* 106 (2): 321–337.

Walters, Christopher. 2012. "A Structural Model of Charter School Choice and Academic Achievement." Working Paper.

▶**Werner, Rachel M., R. Tamara Konetzka, and Daniel Polsky.** 2013. "Nursing Homes: Evidence from State Medicaid Programs." *Health Services Research* 48 (4): 1393–1414.

▶**World Health Organization.** 2013. "WHO methods and data sources for global burden of disease estimates 2000–2011." *Geneva: Department of Health Statistics and Information Systems.*

For Online Publication

Appendix

A Additional Background on Nursing Homes

A.1 Brief History of the Nursing Home Industry

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 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 superseded the original MDS in 1995, and was in turn 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 reports 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.[131] 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 that were already operating on thin profit margins before the pandemic.

A.2 More Details on Various Nursing Home Quality Measures

Donabedian (1985) provides a useful classification of the numerous quality measures that have been used for nursing homes. 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 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 previously already able to access data on quality measures such as staffing levels, ownership status, and deficiency citations through the Nursing Home Compare website, 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).

[131] 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.

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 (Thomas, Gebeloff, and Silver-Greenberg 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.[132]

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.

I do not include expected out-of-pocket price in 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), and these are often uncertain to residents at the time of their nursing home choice. It is therefore unclear the extent to which residents react to the spot price or highly uncertain future prices — previous studies have rejected the hypotheses of both fully myopic and fully forward-looking individuals (Aron-Dine, Einav, Finkelstein, and Cullen 2015; Guo and Zhang 2019), and given the

[132]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" (Rafiei 2022).

heterogeneity in information frictions faced by consumers, we may also expect corresponding variation in their degree of foresight.

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 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, I use payer source at admission in the MDS for my analysis for two reasons. First, using the MDS data instead of claims data allows me to study residents from all payer sources.[133] Second, 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 initial MDS assessment form) rather than what the ultimate payer source ends up being (which is recorded in the claims data).[134]

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' weights, 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 heights 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 five ICD-9 codes entered for the resident has the corresponding first three digits, and zero otherwise.

[133] 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.

[134] 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 MDS data is largely irrelevant to my analysis because I use the initial assessment form for most of my analysis, only using subsequent assessment forms to determine future outcomes of residents.

C Model for Quality Estimation

Equation (1) in the main text can be derived from a simple additive causal model such as the one in Abaluck, Bravo, Hull, and Starc (2021). Suppose that:

$$Y_{ij} = \mu_j + a_i, \tag{7}$$

where Y_{ij} is the potential outcome for resident i in nursing home j, μ_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:

$$\begin{aligned} Y_i &= Y_{i1} + \sum_{j=2}^{J} (Y_{ij} - Y_{i1}) D_{ij} \\ &= \mu_1 + \sum_{j=2}^{J} \beta_j D_{ij} + a_i, \end{aligned} \tag{8}$$

where D_{ij} is a dummy variable for whether resident i chooses nursing home j, and β_j is the quality of nursing home j relative to the omitted nursing home, indexed by $j = 1$. Finally, I decompose the resident-specific 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:

$$a_i = X_i'\gamma + u_i, \; \mathbb{E}[X_i u_i] = 0, \tag{9}$$

and substitute this into equation (8) to obtain 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 β_j on the quality estimates α_j, normalizing η_j to have mean zero:

$$\beta_j = \lambda\alpha_j + \eta_j, \; \mathbb{E}[\alpha_j \eta_j] = 0. \tag{10}$$

Substituting equation (10) into equation (1), I obtain equation (2) as desired.

D Alternative Identification Strategies for Estimation of the Forecast Coefficient

In this section, I describe three 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 λ. 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 $\{\alpha^{\sim t(i),k}, occ^{t(i),k}, \alpha_i^{\sim t(i),k} \times occ^{t(i),k}\}_{k=1}^{K}$ as my instruments, where the superscript k denotes the kth closest nursing home to the resident.[135] 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).[136]

An alternative distance IV strategy is to use the quality of nearby nursing homes *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 in section 3 that average demand for quality is low and that substantial information frictions are likely to be present.

Finally, one might consider estimating λ using an event study analysis based on nursing home entry (or exit). 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 study 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 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.[137]

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

$$\begin{aligned} v_{ij} &= v_j^1(x_i, \zeta_i) - v_j^2(x_i, dist_{ij}), \\ \pi_{ij} &= \pi(x_i, \zeta_i, occ_{ij}), \end{aligned}$$

where I denote resident-specific preference heterogeneity by ζ_i. We can set $|v_j^2(x_i, dist_{ij})| = 1$ 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

[135] The definition of the occupancy measure *occ* corresponds to the definition used for the structural demand estimation in section 3: lagged 7-day average of log occupancy, residualized of nursing home-month fixed effects.

[136] I do not include results with $K > 3$ instruments in Appendix Table A.6 because the F-statistics suggest that these specifications may suffer from many weak instruments bias.

[137] The few differences in the description below from my setup in section 3.2 (e.g., scale normalizations) make no difference to the identification argument and avoids the need to introduce additional notation.

decision can be written as:

$$\begin{aligned}\sigma_{ij} &= \sigma_j(x_{i,}\zeta_i, occ_{ij}) \\ &= \mathbb{I}[\pi(x_i, \zeta_i, occ_{ij}) \geq 0].\end{aligned}$$

Assumption I1*. Unobserved consumer-specific heterogeneity ζ_i is conditionally independent of $(dist'_i, occ'_i)$ given x_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 $\sigma_j(x_i, \zeta_i, occ_{ij})$ is weakly decreasing in occ_{ij}. In addition, for all nursing homes j and resident characteristics x_i, and unobserved preference heterogeneity ζ_i, we have* $\lim_{occ \to -\infty} \sigma_j(x_i, \zeta_j, occ) = 1$, *and* $\lim_{occ \to \infty} \sigma_j(x_i, \zeta_j, occ) = 0$.

Recall that my occupancy measure represents temporary fluctuations (so negative values of *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 $\lim_{occ \to \infty} \sigma_j(x_i, \zeta_j, occ) = 0$. Finally, for $\lim_{occ \to -\infty} \sigma_j(x_i, \zeta_j, occ) = 1$ 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.

For a more succinct statement of the next assumption, I introduce some terminology. We will say that nursing homes j and k are strict substitutes in $dist_{ij}$ at $(x_i, dist_i, occ_i)$ if $\partial s_j(x_i, dist_i, occ_i)/\partial dist_{ik}$ and $\partial s_k(x_i, dist_i, occ_i)/\partial dist_{ij}$ both exist and are strictly positive, where $s_j(x_{i,} dist_i, occ_i)$ is the share of residents with $(x_{i,} dist_i, occ_i)$ who are matched with nursing home j. Define $\Sigma(x_i, dist_i, occ_i)$ to be the matrix with (j, k)th entry equal to one if j and k are strict substitutes in $dist$ at $(x_i, dist_i, occ_i)$ and zero otherwise. In addition, let $\Sigma(x_i, dist_i) \equiv \vee_{occ \in supp(occ)} \Sigma(x_i, dist_i, occ)$ so that the (j, k)th entry is one if j and k are strict substitutes at some occupancy.

Assumption I3*. For every x_i and all but a finite set of $dist_{ij}$ in its support, the graph of $\Sigma(x_i, dist_i)$ 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 $dist_i$ is rectangular with non-empty interior. (ii) For each x_i and j, the function $v_j^2(x_i, dist_j)$ is continuously differentiable in $dist_j$ with $v_j^{2\prime}(x_i, dist_j) \neq 0$ for all $dist_j$.*

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* $|\mathcal{J}| > 1$*, then for every* w*, (i) the function* $v_j^2(x, \cdot)$ *is identified for every* $j \in \mathcal{J}$ *and* $dist_j \in supp(dist_j)$*, and (ii) the joint distribution of* (u_i, π_i^{cutoff}) *is identified for every value* (u, π) *in the interior of* $v^2(x, supp(dist)) \times supp(occ) = \Pi_{j=1}^{J} v_j^2(x, supp(dist_j)) \times supp(occ)$*, where* $\pi_i^{cutoff}(x_i, \zeta_i) \equiv sup\{occ : \pi(x_i, \zeta_i, occ) \geq 0\}$*.*

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 residents who choose a nursing home further away. Denote the potential set of nursing homes in resident i's choice set by $\mathcal{J}_i \equiv \{j \in J | dist_{ij} \leq 15 \text{ miles}\}$. Even after restricting to nursing homes within 15 miles, the size of residents' choice sets (before any supply side constraints) $J_i \equiv |\mathcal{J}_i|$ tends to be quite large: it has a median of roughly 30, a mean of 50, and can take values greater than 200.

In the following description for the Gibbs sampler, when drawing structural error terms in sequence for $j \in \mathcal{J}_i$, I assume an increasing order (although obviously any other order works as well). In addition, to simplify notation, I denote variables in residents' utility and nursing homes' admissions equations by $\mathbf{X}_{ij}$ and W_{ij} respectively,[138] and refer to the nursing home that resident i ends up in by $\mu(i)$.

Denoting iterations of the Gibbs sampler by k and indicating the values of various parameters in the kth iteration of the Gibbs sampler using a superscript k, the steps for implementing the Gibbs sampler are as follows.

1. Initialization ($k = 0$): I assume that $(\epsilon_{ij}, \omega_{ij}) \sim^{i.i.d.} N(0, I_2)$ and set the following conjugate priors for the parameters: $(\kappa', \psi')' \sim N(0, 100I)$.
    (a) Set the initial values of the parameters $\theta^0 = \left(\kappa^{0\prime}, \psi^{0\prime}\right)$ at their prior mean.
    (b) Initial data augmentation: For each resident i, draw the vector ϵ_i^0 such that $v_{i,\mu(i)}^0 \geq v_{ij}^0$ for all $j \in \mathcal{J}_i$.
        i. Draw $\omega_{i,\mu(i)}^0$ such that $\omega_{i,\mu(i)}^0 \geq -W_{ij}'\psi^0$ and for $j \neq \mu(i)$ draw ω_{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 ϵ_{ij}^0 such that $\epsilon_{ij}^0 \leq \left(\mathbf{X}_{i,\mu(i)} - \mathbf{X}_{ij}\right)' \kappa^0 + \epsilon_{i,\mu(i)}^0$ if $\pi_{ij}^0 \geq 0$ or draw ϵ_{ij}^0 unconditionally otherwise.
2. For $k + 1 = 1, ..., K$:
    (a) Draw the profit shocks $\omega_i^{k+1} | v_i^k; \psi^k$ in sequence for $j \in \mathcal{J}_i$.
        i. If $v_{ij}^k < v_{i,\mu(i)}^k$, draw ω_{ij}^{k+1} unconditional on assignment (given that even if i is eligible for j, i would not choose j).

[138] These include resident characteristics x_i, nursing home characteristics w_j, distance between residents and nursing homes $dist_{ij}$, occupancy fluctuations at nursing homes occ_{ij}, and interactions between these variables.

ii. If $v_{ij}^k > v_{i,\mu(i)}^k$, draw ω_{ij}^{k+1} from a truncated normal with mean and variance given by the conditional distribution and truncation point $\omega_{ij}^{k+1} < -W_{ij}\psi^k$ (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 $\omega_{ij}^{k+1} \geq -W_{ij}'\psi^k$ (given that i must always be eligible for the facility she was ultimately assigned to).

(b) Update π_i^{k+1} according to $\pi_{ij}^{k+1} = W_{ij}'\psi^k + \omega_{ij}^{k+1}$.

(c) Draw the utility shocks $\epsilon_i^{k+1}|\pi_i^{k+1}; \kappa^k$ in sequence, for $j \in \mathcal{J}_i$.

i. If $\pi_{ij}^{k+1} < 0$, draw ϵ_{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 ϵ_{ij}^{k+1} from the conditional distribution with truncation point given by $v_{ij}^{k+1} < \mathbf{X}_{ij}'\kappa^k$.

iii. For $j = \mu(i)$, draw $\epsilon_{i,\mu(i)}^{k+1}$ such that $v_{i,\mu(i)}^{k+1}$ 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 $v_{ij}^{k+1} = \mathbf{X}_{ij}'\kappa^k + \epsilon_{ij}^{k+1}$.

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

i. First, we update κ. Denote the design matrix in the equation for indirect utilities by $\mathbb{X}$. In matrix notation, we have:

$$v = \mathbb{X}\kappa + \epsilon,\ \epsilon \sim N(0, I).$$

We have a normal conjugate prior for κ, with mean μ_κ^0 and covariance matrix Σ_κ^0. The posterior distribution of κ conditional on v and W is:

$$\kappa|(v, \mathbb{X}) \sim N(\tilde{\mu}_\kappa, \tilde{\Sigma}_\kappa),$$

where the posterior mean and covariance matrix are given by:

$$\begin{aligned}\tilde{\mu}_\kappa &= \left(\frac{\mathbb{X}'\mathbb{X}}{\sigma_\epsilon^2} + \left(\Sigma_\kappa^0\right)^{-1}\right)^{-1}\left(\left(\Sigma_\kappa^0\right)^{-1}\mu_\kappa^0 + \frac{\mathbb{X}'\kappa}{\sigma_\epsilon^2}\right) \\ &= \left(\mathbb{X}'\mathbb{X} + \left(\Sigma_\kappa^0\right)^{-1}\right)^{-1}\left(\left(\Sigma_\kappa^0\right)^{-1}\mu_\kappa^0 + \mathbb{X}'\kappa\right),\end{aligned}$$

$$\begin{aligned}\tilde{\Sigma}_{\theta_v} &= \left(\frac{\mathbb{X}'\mathbb{X}}{\sigma_\epsilon^2} + \left(\Sigma_\kappa^0\right)^{-1}\right)^{-1} \\ &= \left(\mathbb{X}'\mathbb{X} + \left(\Sigma_\kappa^0\right)^{-1}\right)^{-1}.\end{aligned}$$

We then set κ^{k+1} by drawing from this posterior distribution.

A. Next, we will update ψ. Denote the design matrix in the equation for the admissions rule by W. In matrix notation, we have:

$$\pi = W\psi + \omega,\ \omega \sim N(0, I).$$

We have a normal prior for ψ, with mean μ_ψ^0 and covariance matrix Σ_ψ^0, so the posterior distribution of θ_π conditional on π and W is:

$$\psi|(\pi, W) \sim N(\tilde{\mu}_\psi, \tilde{\Sigma}_\psi),$$

with posterior mean and covariance matrices given by:

$$\begin{aligned}
\tilde{\mu}_\psi &= \left(\frac{W'W}{\sigma_\omega^2} + \left(\Sigma_\psi^0\right)^{-1}\right)^{-1} \left(\left(\Sigma_\psi^0\right)^{-1} \mu_\psi^0 + \frac{W'\psi}{\sigma_\omega^2}\right) \\
&= \left(W'W + \left(\Sigma_\psi^0\right)^{-1}\right)^{-1} \left(\left(\Sigma_\psi^0\right)^{-1} \mu_\psi^0 + W'\psi\right),
\end{aligned}$$

$$\begin{aligned}
\tilde{\Sigma}_\psi &= \left(\frac{W'W}{\sigma_\omega^2} + \left(\Sigma_\psi^0\right)^{-1}\right)^{-1} \\
&= \left(W'W + \left(\Sigma_\psi^0\right)^{-1}\right)^{-1}.
\end{aligned}$$

We then set ψ^{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 are relatively rare events. 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 bin scatter 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.[139] 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.[140] 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,[141] and nursing homes face legal liabilities if they try to forcefully evict residents who are not ready to be discharged (§483.10, §483.21).[142] Hence, 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.

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

[139] Long-term care hospitals also admit patients from acute care hospitals, but these are typically more clinically intensive cases.

[140] According to a Nationwide Retirement Institute survey conducted by the Harris Poll, more than half of the almost 1,300 surveyed US adults aged 24 or above responded that they would rather die than live in a nursing home (Novotney 2020).

[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.

have died if she had stayed in the nursing home instead.

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, d_i is the number of individuals experiencing the event at time t_i, and n_i is the number of individuals that have not experienced the event until at least time t. Initially, it may seem natural to simply use the Kaplan–Meier estimator separately for each event type, i.e.,

$$\hat{S}_{death}(t) = \prod_{i:t_i^{death} \leq t} \left(1 - \frac{d_i^{death}}{n_i^{death}}\right), \ \hat{S}_{discharge}(t) = \prod_{i:t_i^{discharge} \leq t} \left(1 - \frac{d_i^{discharge}}{n_i^{discharge}}\right).$$

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:

$$\hat{S}_{death}(t) + \hat{S}_{discharge}(t) \geq \hat{S}_{death\ or\ discharge}(t),$$

which 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 models is the cause-specific cumulative incidence function (CIF), defined as $CIF_c(t) \equiv Pr(T \leq t, C = c)$, where T is the time of the first event, and C is the cause (i.e., the type of event). In words, the cause-specific CIF tells us the probability that an event occurs before time t *and* that the cause is c. An attractive feature of CIFs is that they have the property that:

$$CIF_{death}(t) + CIF_{discharge}(t) = CIF_{death\ or\ discharge}(t),$$

in contrast to the Kaplan–Meier estimates in the presence of competing risks. If we introduce covariates $\mathbf{X}$ into the model, we can define the CIF analogously as being conditional on $\mathbf{X}$, i.e., $CIF(t|\mathbf{X})$.

For my analysis, I use the cause-specific hazard model, which is commonly used model of competing risks. This involves estimating cause-specific hazard functions, which are defined as:

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

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

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

[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., individuals in Boston receiving their first dose of caffeine in the morning, and kangaroos in Australia waking up in the morning. Even in this case, the above property holds.

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

$$H_c(t|\mathbf{X}) \equiv \int_0^t h_c(u|\mathbf{X})du,$$

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

$$CIF_c(t|\mathbf{X}) = \int_0^t S(u|\mathbf{X})h_c(t|\mathbf{X})du, \tag{13}$$

where $S(t|\mathbf{X}) = exp\left(-\sum_c H_c(t|\mathbf{X})\right)$ is the survivor function for the composite event.

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

$$L(\beta^k_{c,haz}) = \prod_i \left(\frac{exp(\mathbf{X}'_i\beta_{c,haz})}{\sum_{i'\in\mathcal{R}_i} exp(\mathbf{X}'_{i'}\beta_{c,haz})} \right)^{\mathbb{I}[C_i=c]},$$

where $\mathbb{I}[C_i = c]$ 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, $\mathbf{X}_i$ is the vector of covariates associated with resident i and the nursing home she went to, and $\mathcal{R}_i$ contains residents who do not exit their nursing home and are not censored before the minimum of resident i's censoring and exit times.

It should be stressed that the cause-specific hazard model is used here only to predict (counterfactual) death and discharge times, and that one should not attempt to interpret the coefficients $\beta^k_{c,haz}$. In fact, equation (13) reveals that the sign of the hazard ratio $\beta^k_{c,haz}$ associated with the kth covariate in $\mathbf{X}$ does not tell us whether $CIF_c(t|\mathbf{X})$ is increasing or decreasing in this covariate. This is because the $CIF_c(t|\mathbf{X})$ depends on the survivor function $S(u|\mathbf{X})$, which itself depends on the hazard functions for other causes. Hence, $CIF_c(t|\mathbf{X})$ depends on the kth covariate in $\mathbf{X}$ not only through $\beta^k_{c,haz}$ but also through the hazard ratios associated with this covariate in the *other* cause-specific hazards, $\beta^k_{c',haz}$ (which of course may have a different sign).

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

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

$$T^c_{ij} = H^{-1}_{c,0}\left(-log(U^c_{ij})exp(-\mathbf{X}'_{ij}\beta_{c,haz})\right), \tag{14}$$

where U^c_{ij} is a uniformly distributed random variable, and $c \in \{death, discharge\}$.

A practical difficulty in implementing this formula is that $H_{c,0}(\cdot)$ is non-parametric in my model, and thus there is no closed form formula for $H^{-1}_{c,0}(\cdot)$. Hence, I approximate $H^{-1}_{c,0}(\cdot)$ as follows. Examining the estimates of the cumulative baseline hazard functions $H_{c,0}(\cdot)$ in Appendix Figure A.10, I 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 $H^{-1}_{c,0}(\cdot)$, i.e., $\hat{H}^{-1}_{c,0}(x) = \frac{1}{b^c} \cdot \left[exp\left(\frac{x}{a^c}\right) - 1\right]$.

As was mentioned briefly in the main text, I consider different assumptions about the correlation between U^c_{ij} and $U^c_{ij'}$ in my simulations. Exactly how I implement this is described in the next subsection.

G.4 Simulation Algorithm

In my simulations, I consider various counterfactuals by modifying parameters in the structural demand model, or by changing various quantities such as nursing homes' quality. In particular, to mimic the effect of an information intervention, I modify residents' demand for quality from κ_i^β (where κ_i^β is resident i's demand for quality β taking into account her characteristics from the structural demand model with heterogeneous preferences) to a counterfactual value $\tilde{\kappa_i}^\beta$ (where $\tilde{\kappa_i}^\beta$ is an MRS that is consistent with previous demand estimates in the literature on hospital settings). To model the effect of quality changes arising from minimum standard mandates, a pay-for-performance scheme, or endogenous quality adjustments by nursing homes, I simply change nursing home quality from β to $\tilde{\beta}$ in my simulations.

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 (Kth) 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 reveal 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 β_j of nursing homes with quality below the 10th percentile to be equal to the 10th percentile of quality, and I model the pay-for-performance scheme by increasing β 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 I.

Next, I introduce some notation that will be necessary for a detailed description of the simulation algorithm. 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

[144] In 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.

first day of the sample, and $t = T^{sim}$ be the last day. We can order residents who are admitted to a nursing home on day t arbitrarily, from $i_t = 1, ..., I_t$. Also, let T_{i_t} by the death or discharge date of resident i. The simulation proceeds as follows:

1. Initialize the counterfactual flows to be the same as the original flows: $\tilde{flow}_{jt} = flow_{jt}$.
2. For $t = 1, ..., T^{sim}$:
   (a) For resident $i_t = 1, ..., I_t$:
      i. Simulate residents' and nursing homes' counterfactual preferences, as given by:
$$\tilde{v}_{ij} = v_{ij}^{(K)} - \kappa_i^{\beta}\beta_j + \tilde{\kappa}_i^{\beta}\tilde{\beta}_j,$$
$$\tilde{\pi}_{ij} = \pi_{ij}^{(K)} + (\tilde{occ}_{ij} - occ_{ij})'\psi^{occ}.$$
      ii. Find the nursing home the resident is assigned to given these counterfactual values of the latent variables, which I denote by $\tilde{\mu}(i_t)$.[145] This is given by:
$$\tilde{\mu}(i_t) = argmax\{j : \pi_{ij} \geq 0, \text{ and } \tilde{v}_{ij} \geq v_{ij'} \forall j' \text{ s.t. } \pi_{ij'} \geq 0\}.$$
      iii. If $\tilde{\mu}(i_t) = \mu(i_t)$, and $\beta_{\mu(i_t)} = \tilde{\beta}_{\mu(i_t)}$, 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 $\beta_{\mu(i_t)} \neq \tilde{\beta}_{\mu(i_t)}$, and we are assuming that unobserved risks are independent:
         A. Simulate her event time $\tilde{T}^c_{i,\tilde{\mu}(i_t)}$ for cause $c \in \{death, discharge\}$ by drawing $U_{i,\tilde{\mu}(i_t)}$ from a uniform distribution on $[0, 1]$ and calculating:
$$\tilde{T}^c_{i,\tilde{\mu}(i_t)} = \hat{H}^{-1}_{c,0}\left(-log(U^c_{i,\tilde{\mu}(i_t)})exp(-\tilde{\mathbf{X}}'\beta_{c,haz})\right).$$
         Set $\tilde{T}_{i_t} = \lceil \min\{\tilde{T}^{death}_{i,\tilde{\mu}(i_t)}, \tilde{T}^{discharge}_{i,\tilde{\mu}(i_t)}\}\rceil$.
      v. If $\tilde{\mu}(i_t) \neq \mu(i_t)$, or $\tilde{\mu}(i_t) = \mu(i_t)$ but $\beta_{\mu(i_t)} \neq \tilde{\beta}_{\mu(i_t)}$ and we are assuming that unobserved risks are perfectly correlated:
         A. If she exits the nursing home due to cause c at event time $T^c_{i,\mu(i_t)}$, then compute her unobserved risk for cause c by inverting the equation $T^c_{i,\mu(i_t)} = \hat{H}^{-1}_{c,0}\left(-log(U^c_i)exp(-\mathbf{X}'\beta_{c,haz})\right)$, i.e., compute:
$$U^c_i = exp\left(-(\hat{H}_{c,0}(T_{i,\mu(i_t)})exp(\mathbf{X}'\beta_{c,haz})\right).$$
         Next, because we know that $T^{c'}_{i,\mu(i_t)} \geq T^c_{i,\mu(i_t)}$, it must be the case that:
$$U^{c'}_i \leq g(U^c_i, \mathbf{X}, \beta_{c,haz}, \beta_{c',haz})$$
$$\equiv exp\left(-\hat{H}_{c',0}\left(\hat{H}^{-1}_{c,0}\left(-log(U^c_i)exp(-\mathbf{X}'\beta_{c,haz})\right)\right)exp(\mathbf{X}'\beta_{c',haz})\right).$$

[145] I drop the small number of residents who are ineligible for any nursing home in the counterfactual.

Thus, I draw $U_i^{c'}$ from a uniform distribution on $[0, g(U_i^c, \mathbf{X}, \beta_{c,haz}, \beta_{c',haz})]$.

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 $\min\left\{T_{i,\mu(i_t)}^{death}, T_{i,\mu(i_t)}^{discharge}\right\} > T^{sim} - t$. This implies that:

$$U_i^c \leq h_c(\mathbf{X}_{ij}, \beta_{c,haz}, T^{sim} - t) \equiv exp\left(-\hat{H}_{c,0}(T^{sim} - t)exp(\mathbf{X}'\beta_{c,haz})\right),$$

for $c \in \{death, discharge\}$. So, I draw U_i^c from uniform distributions on $[0, h_c(\mathbf{X}_{ij}, \beta_{c,haz}, T^{sim} - t)]$ for $c \in \{death, discharge\}$.

C. Compute

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

for $c \in \{death, discharge\}$, and set $\tilde{T}_{i_t} = \lceil \min\{\tilde{T}_{i,\tilde{\mu}(i_t)}^{death}, \tilde{T}_{i,\tilde{\mu}(i_t)}^{discharge}\}\rceil$.

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

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

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

C. If $\tilde{T}_{death} < \tilde{T}_{discharge}$ and $\tilde{T}_{death} \leq T^{sim}$, record her counterfactual outcome as *death*.

D. If $\tilde{T}_{death} \geq \tilde{T}_{discharge}$ and $\tilde{T}_{discharge} \leq T^{sim}$, record her counterfactual outcome as *discharge*.

E. If $\tilde{T}_{i_t} > T^{sim}$, 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 $\tilde{flow}_{\tilde{\mu}(i_t),t+\tilde{T}_{i_t}+1}$ because she exits her counterfactual nursing home at $t + \tilde{T}_{i_t}$.

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

$$\tilde{o}_{j,t+1} = \tilde{o}_{j,t} + \tilde{flow}_{j,t+1},$$

and compute the counterfactual occupancy measure using the formula:

$$\tilde{occ}_{j,t+1} = \left(\frac{1}{7}\sum_{s=t-6}^{t} log(\tilde{o}_{j,s})\right) - \bar{\tilde{occ}}_{j,m(t)},$$

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

H Simple Model of an Information Intervention's Mechanisms

H.1 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_i^0 and the quality of the nursing home she goes to $\beta_i \equiv \beta_{\mu(i)}$. Her nursing home quality β_i depends on how informed she is I_i, which is a function $I(I_i^0, D_i^{info})$ of her pre-intervention knowledge I_i^0 and whether she is treated by the information intervention D_i^{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:

$$Y\left(H(I_i^0), Q_i(I(I_i^0, D_i^{info}))\right),$$

where $\beta_i = Q_i(I(I_i^0, D_i^{info}))$.

Additionally, I make the following assumptions:

Assumption M1. *Outcomes improve as H_i^0 and β_i increase: $Y_H > 0$ and $Y_\beta > 0$.*

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 holding treatment constant ($\partial I/\partial I^0 > 0$), 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, ($I(I_i^0,1) - I(I_i^0,0) > 0$), 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, $Y_{\beta\beta} \leq 0$.*[146]

Assumption M5. *Quality matters more for the outcomes of sicker residents: $Y_{H\beta} \leq 0$.*

Assumption M6. *Baseline health is positively related with baseline information: $H'(I_i^0) > 0$.*

Proposition 1. *Under Assumptions M1–M6, the treatment effect is positive and decreasing in baseline information.*

Proof. The treatment effect for resident i is:

$$\tau_i = \tau(I_i^0) = Y\left(H(I_i^0), Q(I(I_i^0,1))\right) - Y\left(H(I_i^0), Q(I(I_i^0,0))\right).$$

The fact that $\tau_i > 0$ is a straightforward consequence of Assumptions M1 and M2. Taking the derivative

[146] According to the linear model in the paper, average outcomes should typically increase one-for-one with quality, in which case $Y_{\beta\beta} = 0$. I allow for concavity here to accommodate a potential ceiling effect (e.g., if Y represents survival probability, it cannot increase beyond 1).

with respect to baseline information I_i^0, I obtain:

$$\tau'(I_i^0) = \overbrace{\left(\frac{\partial Y\left(H(I_i^0), Q(I(I_i^0,1))\right)}{\partial H} - \frac{\partial Y\left(H(I_i^0), Q(I(I_i^0,0))\right)}{\partial H}\right)\cdot H'(I_i^0)}^{\text{“Indirect Tagging” Effect}}$$
$$+ \underbrace{\frac{\partial Y\left(H_i, Q(I(I_i^0,1))\right)}{\partial Q} Q'(I(I_i^0,1))\frac{\partial I(I_i^0,1)}{\partial I^0} - \frac{\partial Y\left(H, Q(I(I_i^0,0))\right)}{\partial Q} Q'(I(I_i^0,0))\frac{\partial I(I_i^0,0)}{\partial I^0}}_{\text{“Reallocation” Effect}}. \quad (15)$$

In the first line, the term in parentheses is negative under Assumption M5, and $H'(I_i^0)$ is positive under Assumption M6. Each of the terms in the second line $\partial Y\left(H(I_i^0), Q(I(I_i^0,1))\right)/\partial Q$, $Q'(I(I_i^0,1))$, $\partial I(I_i^0,1)/\partial I^0$, $\partial Y\left(H(I_i^0), Q(I(I_i^0,0))\right)/\partial Q$, $Q'(I(I_i^0,0))$, and $\partial I(I_i^0,0)/\partial I^0$ is positive due to Assumptions M1 and M2. Combined with the fact that $\partial Y\left(H(I_i^0), Q(I(I_i^0, D_i^{info}))\right)/\partial Q$ is decreasing in D_i^{info} due to Assumptions M2 and M4, and $Q'(I(I_i^0, D_i^{info}))$ 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 the nursing home resident i goes to β_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_{-i}^0 as well as whether they are treated D_{-i}^{info}. For simplicity, I abstract from dynamics.[147] Under this setup, the outcome for resident i under the policy D_i^{info} can be written as:

$$Y_i = Y\left(H_i, Q_i\left(I_i(I_i^0, D_i^{info}), I_{-i}(I_{-i}^0, D_{-i}^{info})\right)\right).$$

I also 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^{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 $\tau_i'(I_i^0) < 0$.

[147] For example, if we consider dynamic effects, there will be little spillover effects for the first residents to receive the intervention, but spillovers will tend to affect subsequent cohorts more.

Now, suppose instead that $Y_i = MH_i + H_iQ_i$, where M is a large number such that $\partial Y_i/\partial H_i > 0$, and suppose $H_i(I_i^0) = -I_0^1$. 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 to compute 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:

$$\tau_i^{D_{-i}^{info}} = Y\left(H_i(I_0^i), Q_i\left(I_i(I_i^0, 1), I_{-i}(I_{-i}^0, D_{-i}^{info})\right)\right) - Y\left(H_i, Q_i\left(I_i, I_{-i}^0\right)\right).$$

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

$$\begin{aligned}\tau_i^{D^{info}} = & \overbrace{Y\left(H_i(I_0^i), Q\left(I_i(I_i^0, 1), I_{-i}^{D^{info}}\right)\right) - Y\left(H_i(I_i^0), Q_i\left(I_i^0, I_{-i}^{D^{info}}\right)\right)}^{\text{Standard Individual-Level Treatment Effect}} \\ & + \underbrace{Y\left(H_i(I_0^i), Q_i\left(I_i^0, I_{-i}(I_{-i}^0, D_{-i}^{info})\right)\right) - Y\left(H_i(I_i^0), Q\left(I_i(I_i^0, 1), I_{-i}^0\right)\right)}_{\text{Spillover of Policy on } i \text{ Due to Changes in Occupancy}}.\end{aligned}$$

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 7, 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

$$\bar{\tau}^{D^{info}} = \int_{i\in\mathcal{I}} \tau_i^{D_{-i}^{info}} dF_i(i)$$

is ambiguous as well.

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

$$\begin{aligned}\frac{d\tau_i^{D_{-i}^{info}}}{dI_i^0} = & \overbrace{\frac{d\left[Y\left(H_i(I_0^i), Q\left(I_i(I_i^0, 1), I_{-i}^{D^{info}}\right)\right) - Y\left(H_i(I_i^0), Q_i\left(I_i^0, I_{-i}^{D^{info}}\right)\right)\right]}{dI_i^0}}^{\text{Standard Individual-Level Tratment Effect Heterogeneity}} \\ & + \underbrace{\frac{d\left[Y\left(H_i(I_0^i), Q_i\left(I_i^0, I_{-i}(I_{-i}^0, D_{-i}^{info})\right)\right) - Y\left(H_i(I_i^0), Q\left(I_i(I_i^0, 1), I_{-i}^0\right)\right)\right]}{dI_i^0}}_{\text{Heterogeneity Due to Negative Spillovers}} \qquad (16)\end{aligned}$$

The first line is the standard individual-level treatment effect heterogeneity as in equation (15) 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 $d\tau_i^{D_{-i}^{info}}/dI_i^0$ 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 as 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 $\varphi_{\tau\beta}$ and $\varphi_{\tau H}$ respectively. The indirect tagging effect is given by the expression:

$$\tau'_{tag}(I_i^0) \equiv \underbrace{\left(\frac{\partial Y\left(H(I_i^0), Q(I(I_i^0,1), I_{-i}^{D^{info}})\right)}{\partial H} - \frac{\partial Y\left(H(I_i^0), Q(I_i^0, I_{-i}^{D^{info}})\right)}{\partial H} \right)}_{\approx \varphi_{\tau H}} \cdot \overbrace{H'(I_i^0)}^{\approx \varphi_{HI^0}},$$

and hence, to approximate this I use the product of $\varphi_{\tau H}$ and φ_{HI^0}, where the latter term is the coefficient from a regression of baseline health on baseline information. The reallocation effect is given by:

$$\begin{aligned} \tau'_{realloc}(I_i^0) \equiv & \frac{\partial Y\left(H(I_i^0), Q(I(I_i^0,1), I_{-i}^{D^{info}})\right)}{\partial Q} \frac{\partial Q(I(I_i^0,1), I_{-i}^{D^{info}})}{\partial I_i^0} \\ & - \frac{\partial Y\left(H(I_i^0), Q(I(I_i^0,0), I_{-i}^{D^{info}})\right)}{\partial Q} \frac{\partial Q(I(I_i^0,0), I_{-i}^{D^{info}})}{\partial I_i^0}, \end{aligned}$$

and I approximate this using the product of $\varphi_{\tau\beta}$ and $\varphi_{\beta I^0}$, 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 $\tau'_{spill}(I_i^0)$, 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:

$$
\begin{aligned}
\bar{\tau} &= \int_{I^0_{min}}^{I^0_{max}} \tau(I^0) dF(I^0) \\
&= \int_{I^0_{min}}^{I^0_{max}} \left[\left(\int_{I^0_{max}}^{I^0} \tau'(I) dI \right) + \tau(I^0_{max}) \right] dF_{I^0}(I^0) \\
&= \underbrace{\int_{I^0_{min}}^{I^0_{max}} \left(\int_{I^0_{max}}^{I^0} \tau'_{tag}(I) dI \right) dF_{I^0}(I^0)}_{\text{Indirect Tagging Effect}} + \underbrace{\int_{I^0_{min}}^{I^0_{max}} \left(\int_{I^0_{max}}^{I^0} \tau'_{realloc}(I) dI \right) dF_{I^0}(I^0)}_{\text{Reallocation Effect}} \\
&+ \underbrace{\int_{I^0_{min}}^{I^0_{max}} \left(\int_{I^0_{max}}^{I^0} \tau'_{spill}(I) dI \right) dF_{I^0}(I^0)}_{\text{Supply Side Spillovers}} + \tau(I^0_{max}),
\end{aligned}
$$

where I also reverse the bounds of the integration so that the effect of the constant term at the end is minimized.[148] I use the same linear approximation for the derivative of the treatment effect function as when analyzing treatment effect heterogeneity.[149]

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):

$$\max_{\beta_j} \Pi_j = \bar{p} N_j - c(N_j, \beta_j),$$

where Π_j is the profit of nursing home j, $\bar{p}$ is the reimbursement rate for each resident, and $c(N_j, \beta_j)$ 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, β_j.[150] The number of residents who choose nursing home j is given by the demand function $N_j(\bar{p}_{res}, \beta_j, \beta_{-j})$ which depends on the price the resident has to pay, $\bar{p}_{res}$ (e.g. a copay) and the quality of all nursing homes, $\{\beta_j\}$. In addition, I assume that the cost of quality is given by $c(N_j, \beta_j) = \left(\frac{c^\beta_j}{2} \beta_j^2 + c^N_j \right) N_j$.

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 $\kappa^{quality}_i \beta_j$, a nursing home fixed effect ξ_j, and an i.i.d. error term with a type-II extreme value distribution. Although capacity constraints are not explicitly

[148]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 $\bar{\tau}$, 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 $\tau(I^0_{max})$.

[149]In principle, 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", so I do not take this additional step.

[150]I ignore fixed costs considering that I hold entry/exit constant in my simulations.

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:

$$s_j(\kappa^\beta, \beta) = \frac{exp(\kappa_j^\beta \beta_j + \xi_j)}{\sum_l exp(\kappa_l^\beta \beta_l + \xi_l)},$$

where I define the firm-specific demand parameters κ_j^β as the average demand for quality among its potential residents: $\kappa_j^\beta \equiv \left(\sum_i \kappa_i^\beta \cdot \mathbb{I}[j \in \mathcal{J}_i]\right) / \left(\sum_i \mathbb{I}[j \in \mathcal{J}_i]\right)$.

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

$$\overbrace{\kappa^\beta \bar{p} s_j (1 - s_j) \bar{N}}^{\text{MR}} = \overbrace{c_j^\beta \beta_j s_j \bar{N}}^{\text{MC for Inframarginal Residents}} \quad \overbrace{+ \kappa^\beta \left(\frac{1}{2} c_j^\beta \beta_j^2 + c^N\right) s_j (1 - s_j) \bar{N}}^{\text{MC for Marginal Resident}},$$

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 first term in the marginal cost expression corresponds 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:

$$(\gamma_j)(\Delta^* \kappa_j^\beta) \approx \Delta^* \beta_j,$$

where $\gamma_j \equiv (\bar{p} - c^N)/c_j^\beta$ 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.[151] 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 Δ^* 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 $\Delta^* \kappa_j^\beta$ in the structural demand estimation, and I will describe my estimation of the change in quality $\Delta^* \beta_j$ in the next subsection.

After computing γ_j using the changes induced by the star ratings system (specifically, using estimated values of $\Delta^* \kappa_j^\beta$ and $\Delta^* \beta_j$), I can then simulate nursing homes' quality choice $\tilde{\beta}_j$ for the counterfactual demand for quality $\tilde{\kappa}^{quality}$ using the same equation, by replacing the pre- and post-

[151] I also make use of the approximations that $\beta^2 \approx 0$ and $\Delta\beta^2 \approx 0$ 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, γ_j is an underestimate.

star ratings demand and quality with the observed and counterfactual values:

$$(\gamma_j)(\tilde{\Delta}\kappa_j^{\beta}) \approx \tilde{\Delta}\beta_j,$$

the difference being that γ_j is now known, and the unknowns are the counterfactual quality choices by nursing homes $\tilde{\beta}_j$. Hence, the counterfactual quality choices by firms are given by:

$$\tilde{\beta}_j \approx \beta_j + \gamma_j \tilde{\Delta}\kappa_j,$$

or equivalently,

$$\tilde{\beta}_j \approx \beta_j + \left(\frac{\tilde{\Delta}\kappa_j}{\Delta^*\kappa_j}\right)\Delta^*\beta_j.$$

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

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.[152] I then estimate $\Delta^{5-star}\beta$ 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.[153] 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 my use of a single quality estimate per nursing home (which nets out year effects) in the main analysis. 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 $\left(c_j^{\beta} \cdot \beta_j + c_j^{N}\right) N_j$ or where cost of quality is not multiplicative in quantity

[152] 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.

[153] Conventional bandwidth selection methods (Imbens and Kalyanaraman 2012; Calonico, Cattaneo, and Titiunik 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.

$\frac{c_j^\beta}{2}\beta_j^2 + c_j^N N_j$, as well as ad hoc demand functions (i.e., demand functions that are not microfounded) which 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 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.[154] 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.[155] 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., $\bar{Y}_{\mu(i),-i} \equiv \frac{1}{N_{\mu(i)}-1}\sum_{i'\neq i} Y_{i'} D_{i',\mu(i)}$, 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 β_j as a function of observables (resident outcomes in this case),

[154] 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.

[155] 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.

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 β_j. We can rewrite this in a more familiar form for panel data:

$$Y_{ji'} = \mu_1 + X_{ji'}'\gamma + \beta_j + \epsilon_{ji'},$$

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

$$\begin{aligned} l_j = -\frac{1}{2}\Bigg\{ & \frac{1}{\sigma_\epsilon^2}\sum_{i'=1}^{N_j}(Y_{ji'} - \mu_1 - X_{ji'}'\gamma)^2 \\ & - \frac{\sigma_\beta^2}{N_j\sigma_\beta^2 + \sigma_\epsilon^2}\left[\sum_{i'=1}^{N_j}(Y_{ji'} - \mu_1 - X_{ji'}'\gamma)\right]^2 \\ & + log\left(N_j\frac{\sigma_\beta^2}{\sigma_\epsilon^2} + 1\right) + N_j log(2\pi\sigma_\epsilon^2)\Bigg\}, \end{aligned}$$

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

$$\hat{\alpha}_j = \frac{\hat{\sigma}_\beta^2}{\hat{\sigma}_\beta^2 + \hat{\sigma}_\epsilon^2/N_j}\left[\frac{1}{N_j}\sum_{i'=1}^{N_j}(Y_{ji'} - \hat{\mu}_1 - X_{ji'}'\hat{\gamma})\right] + \frac{\hat{\sigma}_\epsilon^2/N_j}{\hat{\sigma}_\beta^2 + \hat{\sigma}_\epsilon^2/N_j}\left[\frac{1}{N}\sum_{j=1}^{J}\sum_{i'=1}^{N_j}(Y_{ji'} - \hat{\mu}_1 - X_{ji'}'\hat{\gamma})\right], \quad (17)$$

where N is the total number of observations.[156]

[156] 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 unshrunken estimates $\sum_{i'=1}^{N_j}(Y_{ji'} - \hat{\mu}_1 - X_{ji'}'\hat{\gamma})/N_j$, instead of the shrunken 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".

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 λ. Recall that the endogenous variable is the quality estimate of resident i's chosen nursing home $\hat{\alpha}_i$, and the instrument $\hat{Z}_i$ is the quality estimate of the nursing home closest to the resident, where the "hats" emphasize that these are finite-sample estimates.[157] Given that we use a finite-sample estimate $\hat{\alpha}_i = \alpha_i + e_\alpha$ in place of the estimate $\alpha_i \equiv \lim_{N_{j(i)}\to\infty} \hat{\alpha}_i$ that we will obtain in large samples, it suffers from measurement error, and similarly for $\hat{Z}_i = Z_i + e_{Z,j(i)}$ where $Z_i \equiv \lim_{N_{j(i)}\to\infty} \hat{Z}_i$ since it is defined using the $\hat{\alpha}_j$'s. I assume the measurement error is classical, so that $e_{\alpha,j(i)} \sim^{i.i.d.} N\left(0, Var(e_\alpha)\right)$ and $e_{Z,j(i)} \sim^{i.i.d.} N\left(0, Var(e_Z)\right)$.

I will now show that the IV coefficient for my setting can be written as $\dfrac{Var(\alpha)}{Var(\alpha)+Var(e_\alpha)} \cdot \lambda$, and thus it suffers from the same attenuation bias as the OLS estimate does. First, we write the IV coefficient as the ratio of the reduced-form and first-stage coefficients: $\dfrac{Cov(Y_i, \hat{Z}_i)}{Cov(\hat{\alpha}_i, \hat{Z}_i)}$. Using the formula for classical attenuation bias, as well as the basic properties of the covariance operator, we can write the IV coefficient as $\left(1 + \dfrac{Cov(e_{\alpha,j(i)} e_{Z,j(i)})}{Cov(\alpha, Z)}\right)^{-1} \cdot \lambda$. 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 $D_{nearest}$ be the indicator variable for this event. Then, we can write:

$$\begin{aligned} Cov(e_{\alpha_{j(i)}} e_{Z_i}) &= p \cdot \mathbb{E}[e_{\alpha,j(i)} e_{Z,j(i)} | D_{nearest} = 1] + (1-p) \cdot \mathbb{E}[e_{\alpha,j(i)} e_{Z,j(i)} | D_{nearest} = 0] \\ &= p \cdot Var(e_\alpha), \end{aligned}$$

using the fact that $\hat{\alpha}_i = \hat{Z}_i$ if the resident chooses her nearest nursing home, and that e_{α_j} and $e_{\alpha_{j'}}$ are independent if $j \neq j'$. Using similar computations, we find that $Cov(\hat{\alpha}_i, \hat{Z}_i) = p \cdot Var(\alpha)$. 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 3.1 by running the following regression at

[157] For simplicity, I omit the notation indicating that this is a leave-year-out estimate.

the nursing home-day level:

$$admit_{jdmt} = \gamma_0^a + \gamma_1^a \check{occ}_{jdmt} + \delta_{jmt}^a + \xi_{jdmt}^a,$$

where $admit_{jdmt}$ is a measure of nursing home j's admissions on day d of month m in year t, $\check{occ}_{jdmt}$ is the average log occupancy over the seven days preceding this date,[158] and δ_{jmt}^a are nursing home-month fixed effects.[159] A negative estimate for γ_1^a would be in line with prediction 1.

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

$$x_i = \gamma_0^x + \gamma_1^x occ_{\mu(i),d(i),m(i),t(i)} + \tilde{x}_i' \gamma_{\tilde{x}}^x + \delta_{\mu(i)}^x + \xi_i^x,$$

for different resident characteristics x_i, controlling for other resident characteristics $\tilde{x}_i$ and nursing home fixed effects $\delta_{\mu(i)}^x$, where $\mu(i)$ is the nursing home that resident i is admitted to, and similarly $d(i)$, $m(i)$, and $t(i)$ indicate the day, month, and year that resident i was admitted respectively. Evidence that $\gamma_1^x \neq 0$ would be in line with prediction 2, and we would expect γ_1^x 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/occupancy percentile). Regardless of the specification, the estimate of γ_1^a 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. 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,[160] the results show that nursing homes admit different types of residents during periods of high and low occupancy ($\delta_a^x \neq 0$), which is consistent with prediction 2.

[158] 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.

[159] 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.

[160] For example, while certain health conditions make residents more costly to care for, these conditions may also be associated with higher reimbursement rates.

N Simple Model of Imperfect Information About Quality

Suppose that residents do not directly observe nursing home quality β_j, but only a noisy signal of it, $\check{\beta}_j = \beta_j + e_j^{\beta}$, as well as publicly available information about the nursing home $w_{\sim\beta,j}$, which includes reported staffing levels, ownership status, and number of cited deficiencies.[161] Assume also that $e_j^{\beta} \sim^{i.i.d.} N(0, \sigma^2_{e^{\beta}})$, 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, $w'_{\sim\beta,j}\beta_{\beta}^{w_{\sim\beta}}$. In this case, residents' conditional expectation of nursing home j's quality given the signal is:

$$\tilde{\beta}_j \equiv \mathbb{E}[\beta_j|\check{\beta}_j] = \frac{\sigma^2_{e^{\beta}}}{\sigma^2_{e^{\beta}} + \sigma^2_{\beta}} w'_{\sim\beta,j}\beta_{\beta}^{w_{\sim\beta}} + \frac{\sigma^2_{\beta}}{\sigma^2_{e^{\beta}} + \sigma^2_{\beta}} \check{\beta}. \tag{18}$$

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

$$v_{ij} = w'_{\sim\beta,j} \overbrace{\left[\left(\frac{\kappa^{\tilde{\beta}}\sigma^2_{e^{\beta}}}{\sigma^2_{e^{\beta}} + \sigma^2_{\beta}}\right) \cdot \beta_{\beta}^{w_{\sim\beta}}\right]}^{\equiv \kappa^{\tilde{\beta}}_{w_{\sim\beta}}} + \overbrace{\left[\frac{\sigma^2_{\beta}}{\sigma^2_{e^{\beta}} + \sigma^2_{\beta}} \cdot \kappa^{\tilde{\beta}}\right]}^{\equiv \kappa^{\tilde{\beta}}_{\beta}} \beta_j + \kappa^{dist} dist_{ij} + \tilde{\epsilon}_{ij}, \tag{19}$$

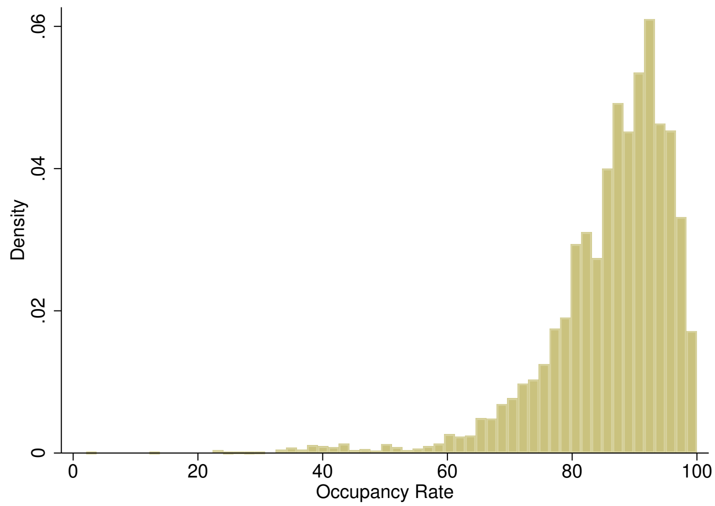
where $\tilde{\epsilon}_{ij} \equiv \kappa^{\tilde{\beta}}_{\beta} e_j^{\beta} + \epsilon_{ij}$ is a composite error term that is conditionally independent of the instruments given the covariates. Then, it is clear from the expression for $\kappa^{\tilde{\beta}}_{w_{\sim\beta}}$ that if residents have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

The results are qualitatively similar if we allow for noise in the estimated quality measure β_j. It is easy to show that residents' preferences for components of $w_{\sim\beta,j}$ 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 β_j.

[161] 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^{β}. I will also explain at the end of this section that relaxing this assumption does not change the model's predictions.

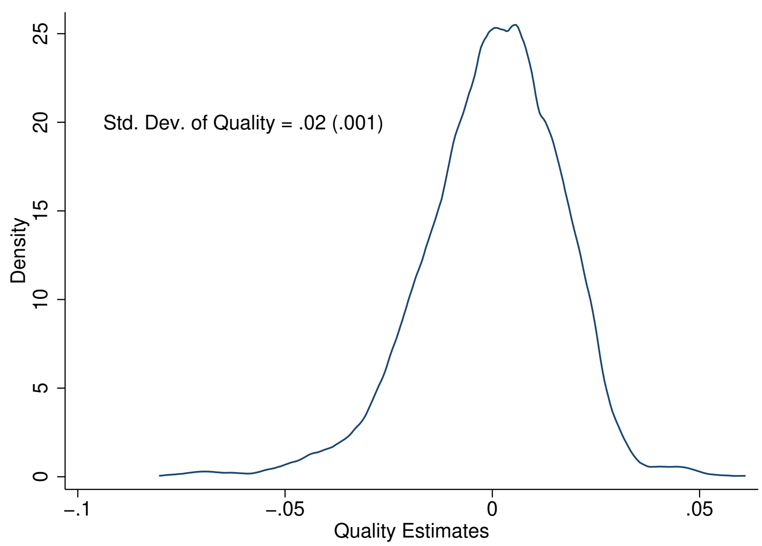
A Appendix Figures and Tables

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.

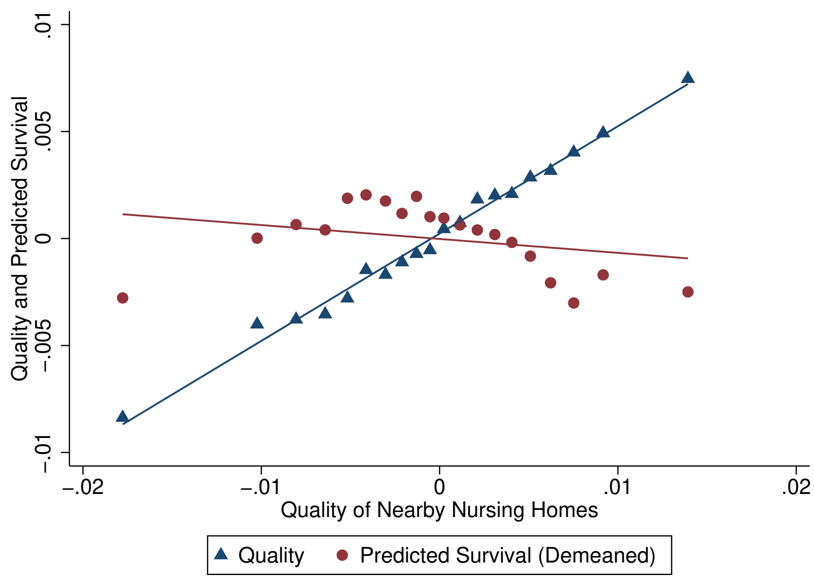
Figure A.2: Kernel Density Plot of Nursing Home Quality Estimates



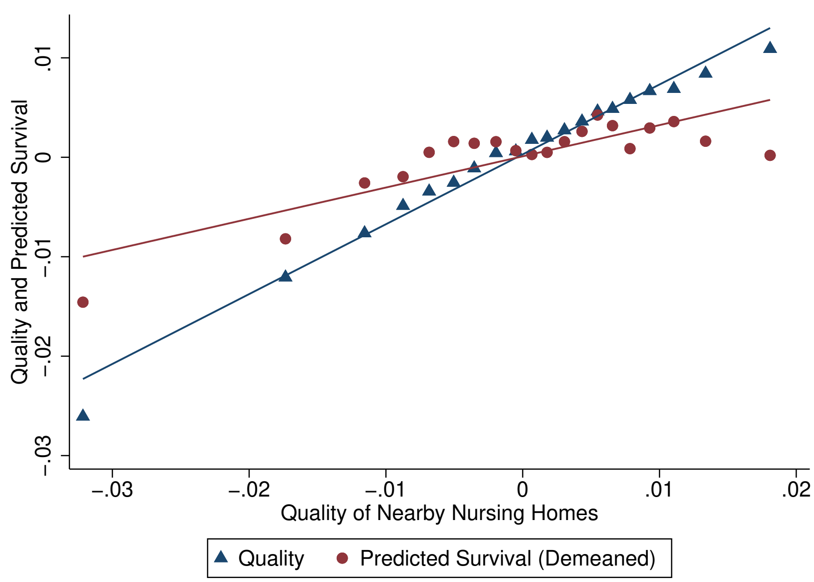
Notes: This figure contains a kernel density plot of the main quality estimates, using an Epanechnikov kernel. The standard error for the standard deviation of nursing home quality displayed in the figure is calculated based on the square root of the variance from the inverse of the Fisher information matrix from the maximum likelihood estimation of the empirical Bayes model in equation (1).

Figure A.3: Robustness of First Stage Assumption and Exclusion Restriction

(a) County Fixed Effects



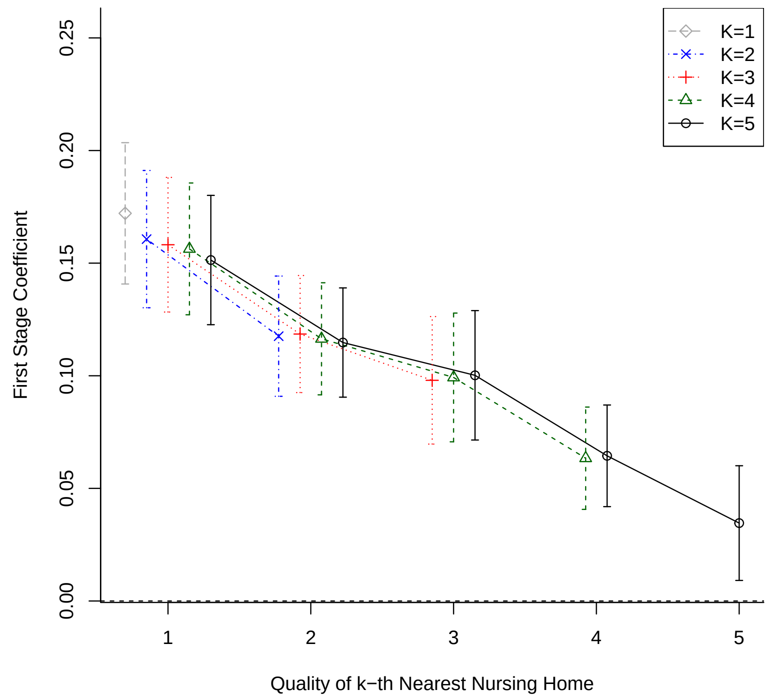
(b) No County Fixed Effects



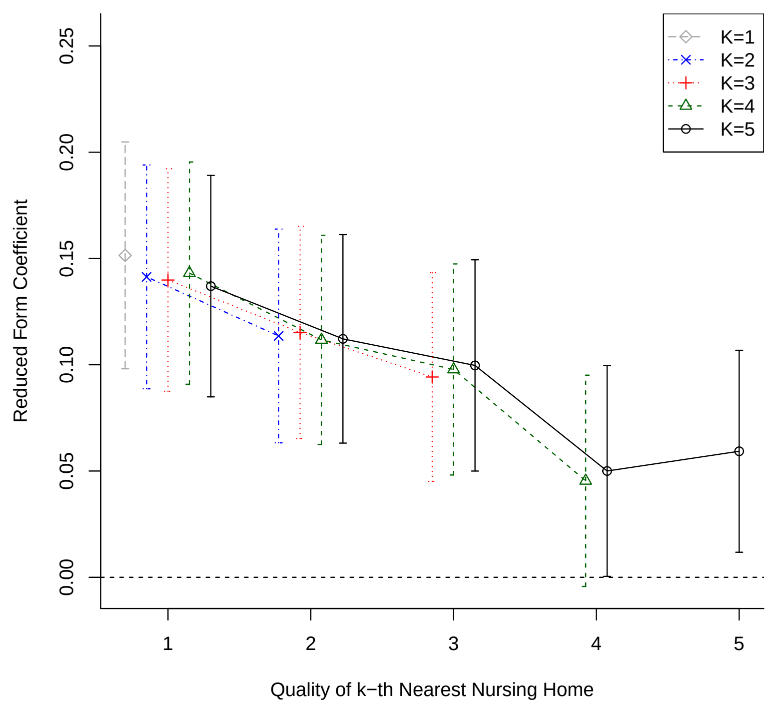
Notes: The x-axis in these figures correspond to the average quality of the five nearest nursing homes to each resident. Variables in panel A are residualized of county fixed effects, whereas variables in panel B are not residualized of county fixed effects.

Figure A.4: First Stage and Reduced Form for Different IV Specifications

(a) First Stage



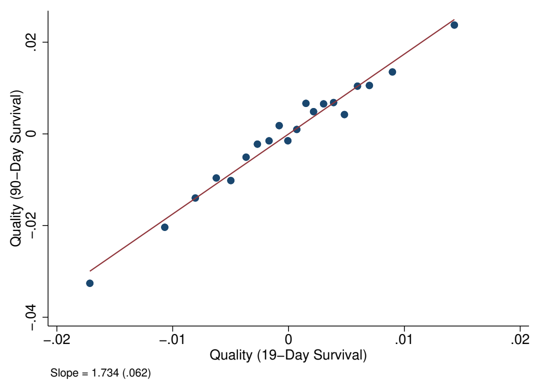
(b) Reduced Form



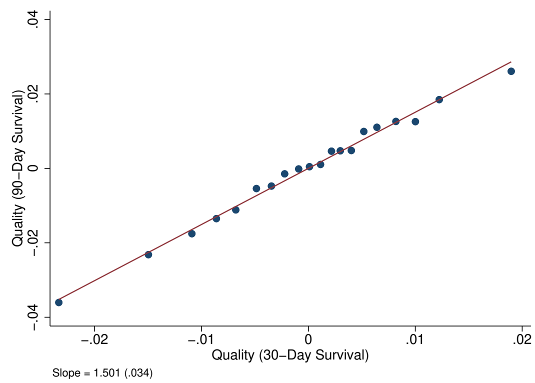
Notes: These figures plot the first-stage and reduced-form coefficients and the associated 95 percent confidence intervals for IV specifications that use the quality of the K nearest nursing homes to each resident as the instrument(s), for K ranging from one to five.

Figure A.5: Relationship Between Quality Estimates Based on Mortality Over Different Time Horizons

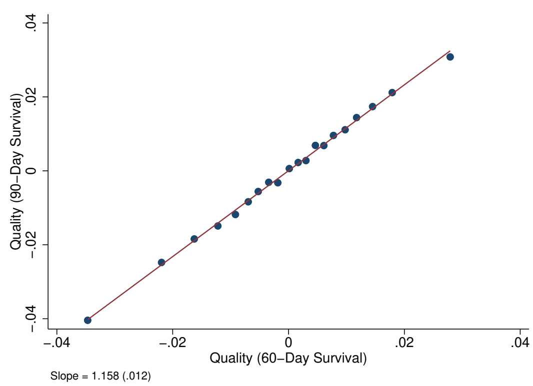
(a) 19-Day Versus 90-Day Mortality



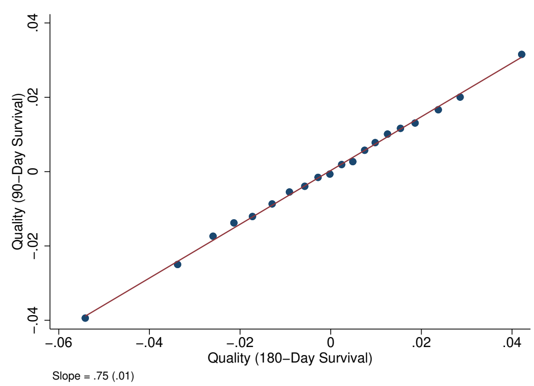
(b) 30-Day Versus 90-Day Mortality



(c) 60-Day Versus 90-Day Mortality

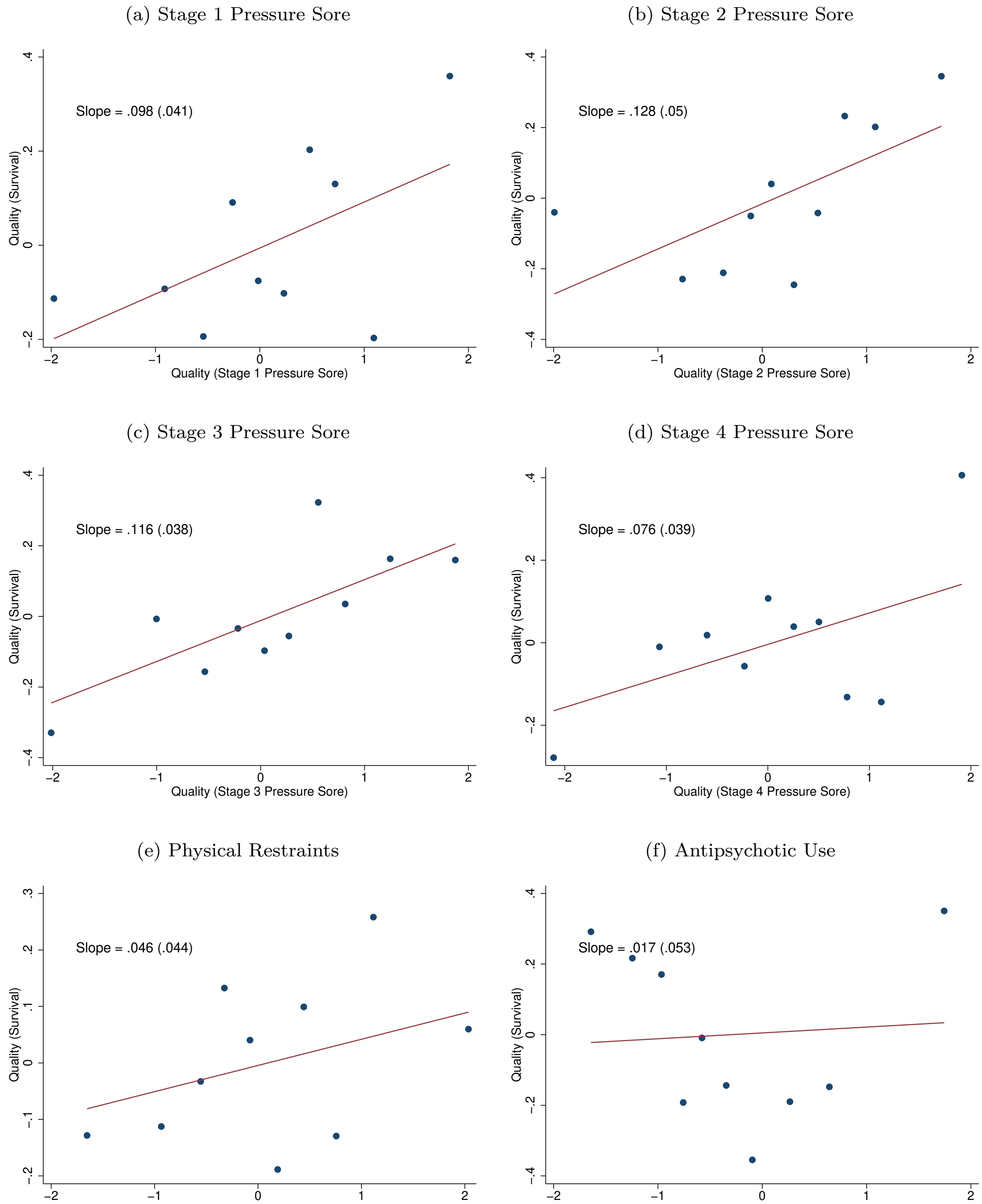


(d) 180-Day Versus 90-Day Mortality



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 bin scatters are weighted by the number of observations for each nursing home, and the standard errors are clustered by nursing home.

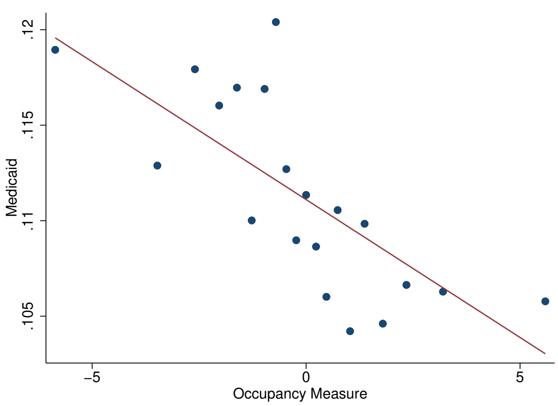
Figure A.6: Relationship Between Main Quality Estimates and Quality Estimates Based on Other Outcomes



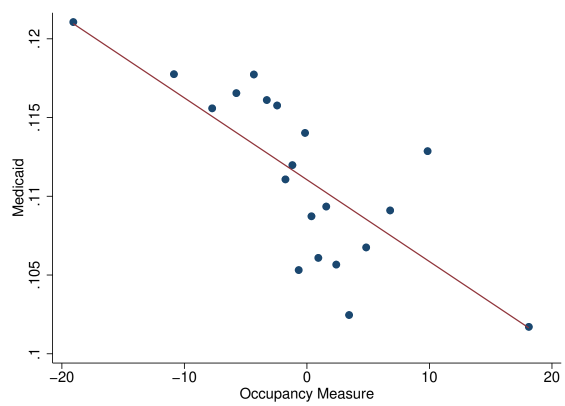
Notes: The x-axis of these figures correspond to quality estimates using other resident outcomes (instead of 90-day survival rate) as the dependent variable. The estimation procedure is the same as for my main quality estimates, except that I use a different outcome variable, and I do not use double machine-learning to select controls for computational reasons. The outcome variables I use are dummies for not developing a stage S or higher pressure sore (for $S \in \{1, 2, 3, 4\}$), no use of physical restraints, and no use of antipsychotics during the first 90 days after admission. The bin scatters 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

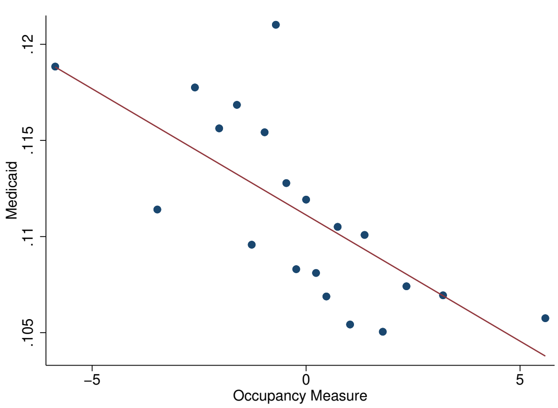
(a) Occupancy (No Controls)



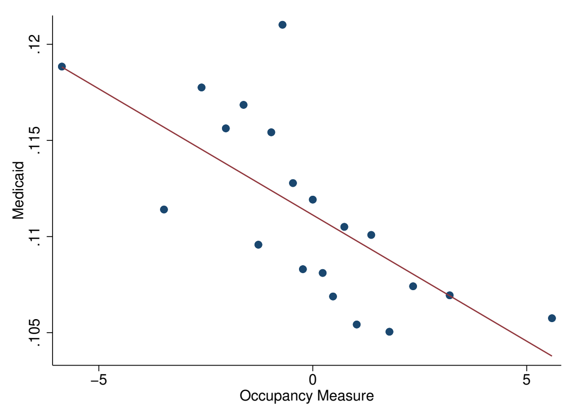
(b) Occupancy Percentile (No Controls)



(c) Occupancy (Controlling for Other Characteristics)

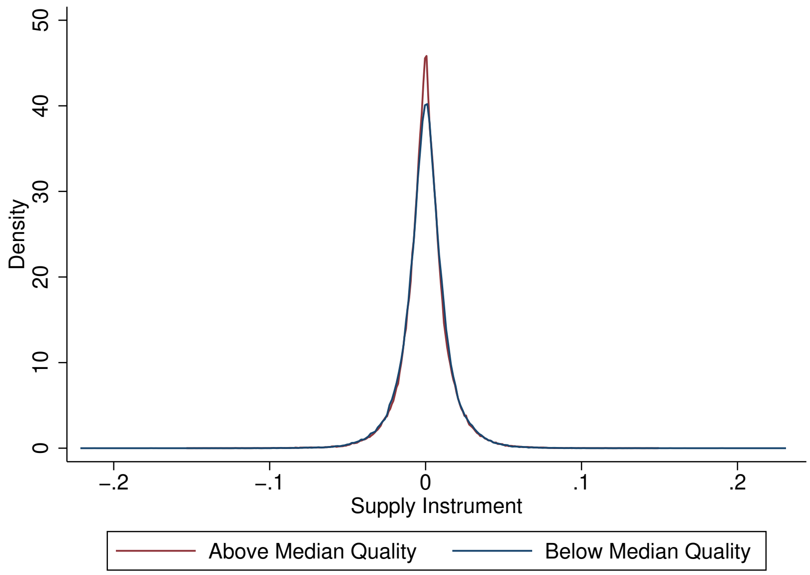


(d) Occupancy Percentile (Controlling for Other Characteristics)



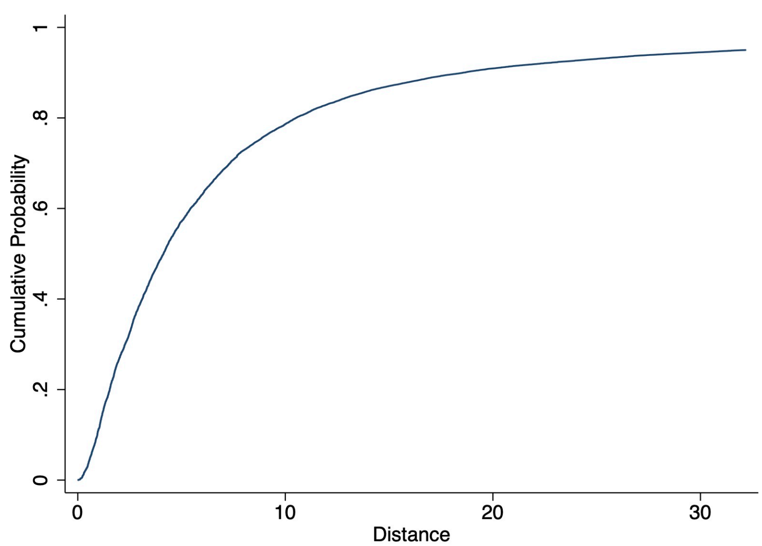
Notes: The occupancy measure is the lagged 7-day average of either occupancy or occupancy percentile (as indicated in the subfigure title) as of the date of admission for the resident, residualized of nursing home-month fixed effects. Occupancy percentiles are computed based on the occupancy distribution within each nursing home. Nursing home fixed effects are included in the bin scatters and regressions, and the unit of observation is a resident.

Figure A.8: Kernel Density Plot of occ_{ij} for Nursing Homes with Above-Median and Below-Median Quality



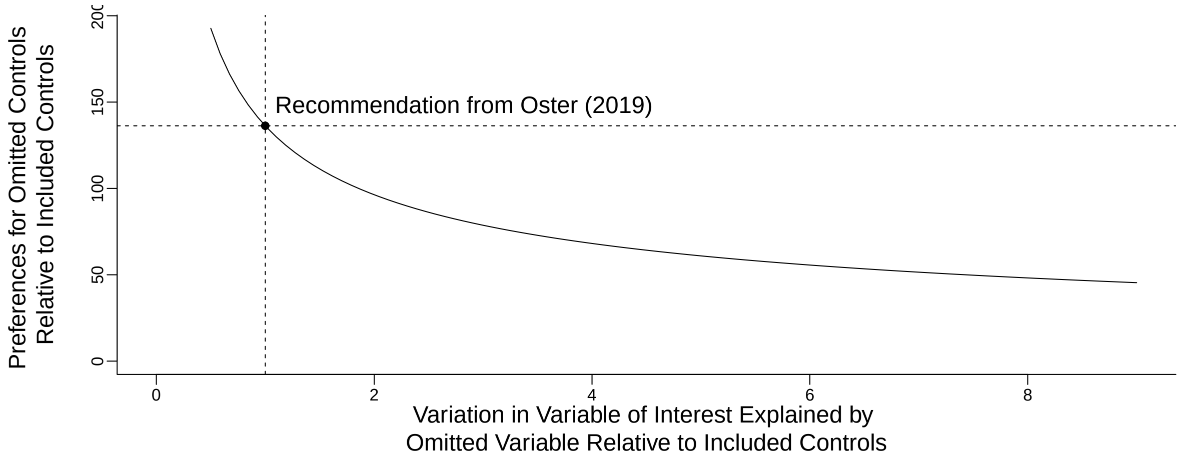
Notes: This figure contains the kernel density plot for the distributions of the supply side instrument occ_{ij} at above-median and below-median quality nursing homes, based on the Epanechnikov kernel. The unit of observation is a resident-nursing home pair (for nursing homes within 15 miles of each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

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



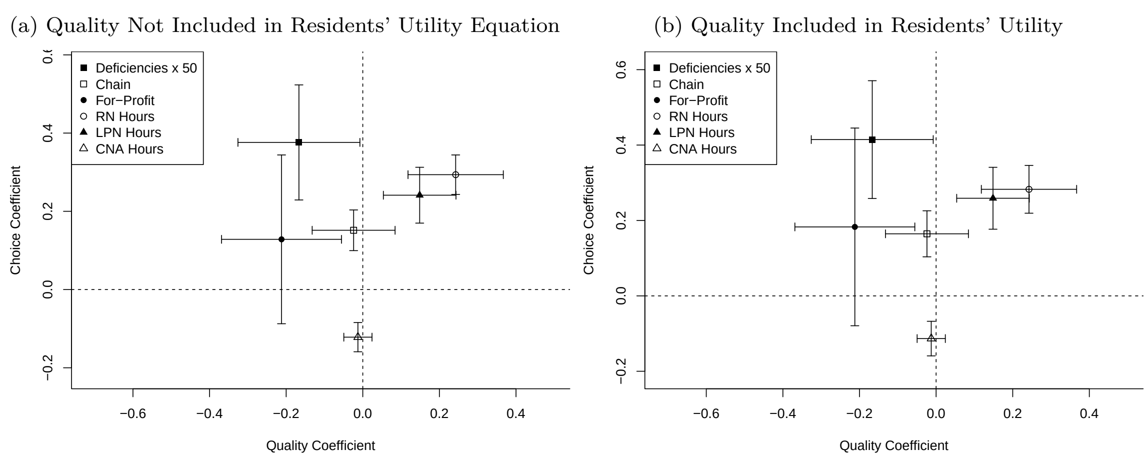
Notes: This figure shows the empirical CDF of distances between residents and their chosen nursing homes. The unit of observation is a resident.

Figure A.10: Conditions Under Which Selection on Unobservables Can Explain the Low Demand Estimate



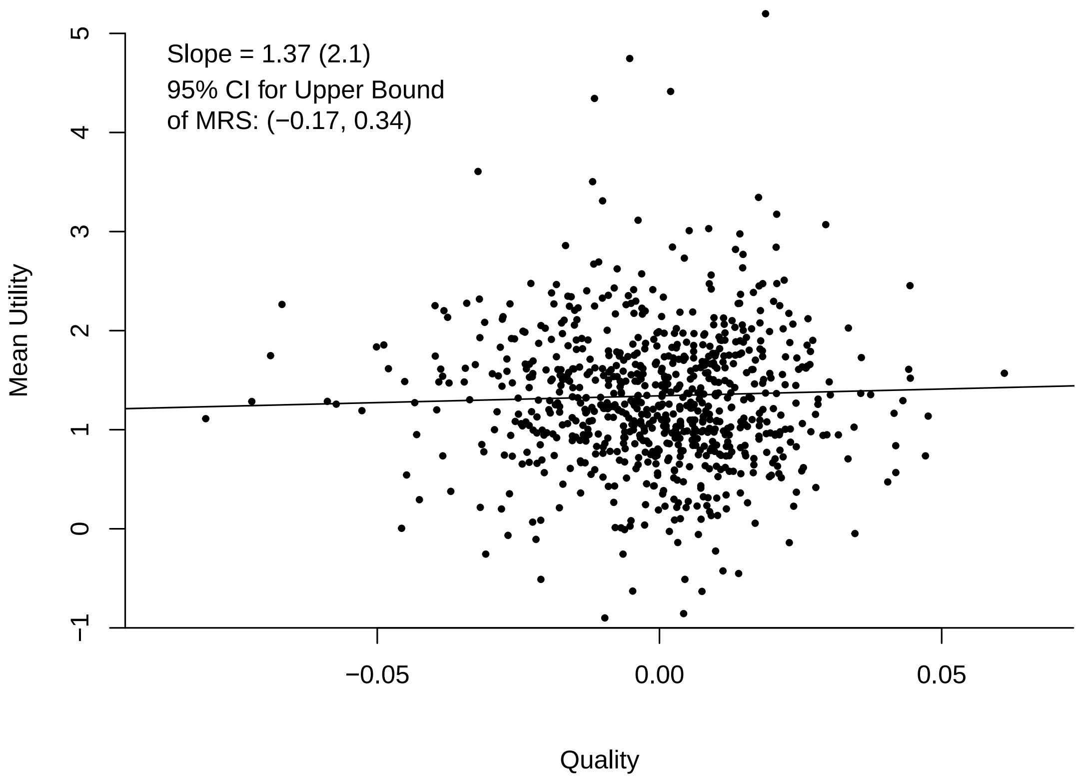
Notes: This figure shows conditions under which omitted variables bias can completely account for the low demand estimate in this paper, using methods from Cheng (2023). The x-axis represents hypothetical values of the proportional selection relationship (Oster 2019), i.e., the omitted variable's importance for explaining the explanatory variable of interest (quality β_j in this case) relative to the included controls (observable nursing home characteristics $w_{\sim\beta,j}$ in this setting) as measured by an R-squared from a hypothetical regression of β_j on $w_{\sim\beta,j}$ and the omitted variable, whereas the y-axis corresponds to hypothetical values for how much more the resident values the omitted variable relative to the included controls. The area to the northeast of the curve corresponds to combinations of these values under which omitted variables bias alone can explain the difference between my MRS estimate and the one from Chandra, Finkelstein, Sacarny, and Syverson (2016). Oster suggests that in many empirical applications, it is reasonable to assume that the omitted variable explains at most as much of the variation in the explanatory variable of interest as the omitted variable does, and the dashed lines in this figure show that under this assumption, the resident must value the omitted variable more than 100 times as much as she values nursing home characteristics $w_{\sim\beta,j}$, in order for omitted variables bias to completely explain my low demand estimate.

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



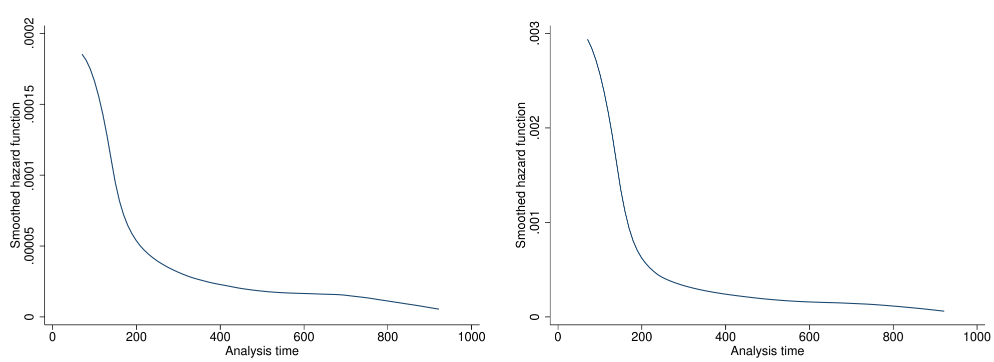
Notes: The y-axis corresponds to estimates of the choice coefficients from the structural demand model where utility is a function of nursing home characteristics (as well as nursing home quality β_j in panel B, but not in panel A). The x-axis corresponds to coefficient estimates from a regression of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

Figure A.12: Scatterplot of Mean Utility Against Quality



Notes: This figure plots mean utilities (net of distance preferences) from the structural demand estimation (i.e. the estimated nursing home fixed effects in residents' preferences) against the quality estimates on the x-axis. The unit of observation is a nursing home, and 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, $h_{c,0}(t)$

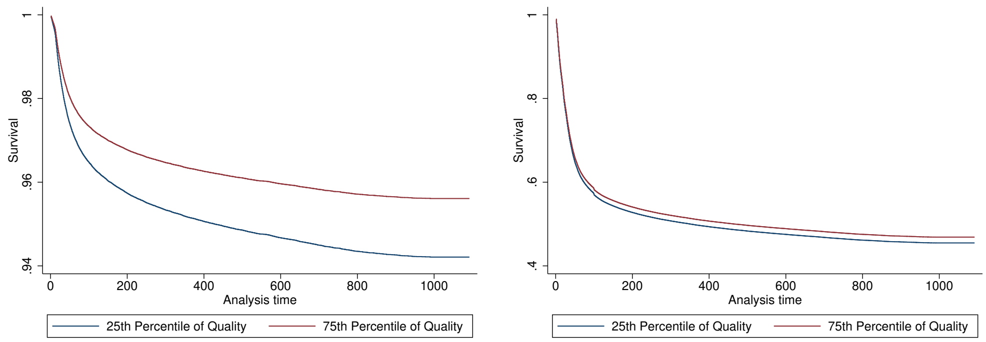


(a) Death

(b) Discharge

Notes: These figures plot the estimated cause-specific baseline hazard functions for death in panel A, and discharge in panel B.

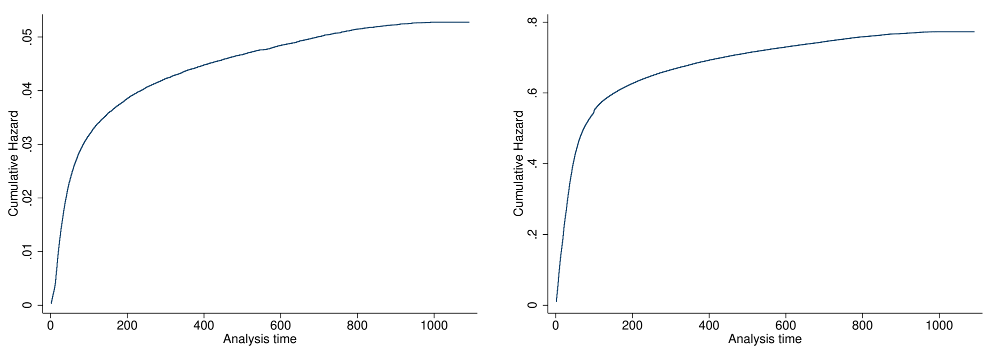
Figure A.14: Survival Curves for Cause-Specific Hazard Model (Split by Nursing Home Quality)



(a) Death (b) Discharge

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).

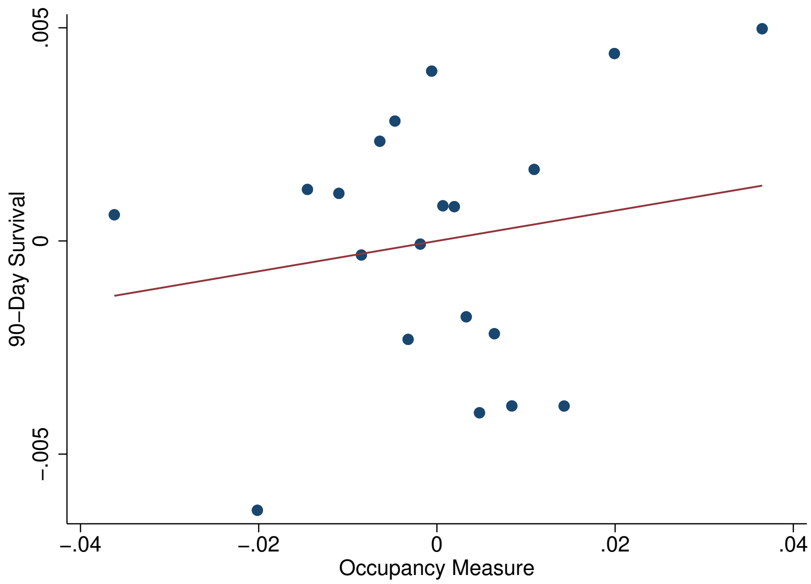
Figure A.15: Cause-Specific Cumulative Baseline Hazard Functions, $H_{c,0}(t)$



(a) Death (b) Discharge

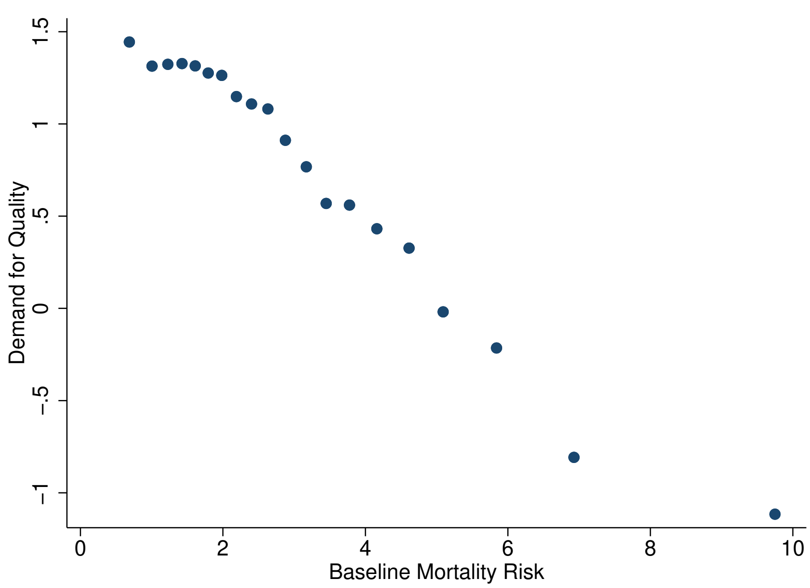
Notes: These figures plot the estimated cause-specific cumulative baseline hazard functions for death in panel A, and discharge in panel B.

Figure A.16: Relationship Between Occupancy Fluctuations and Resident Outcomes



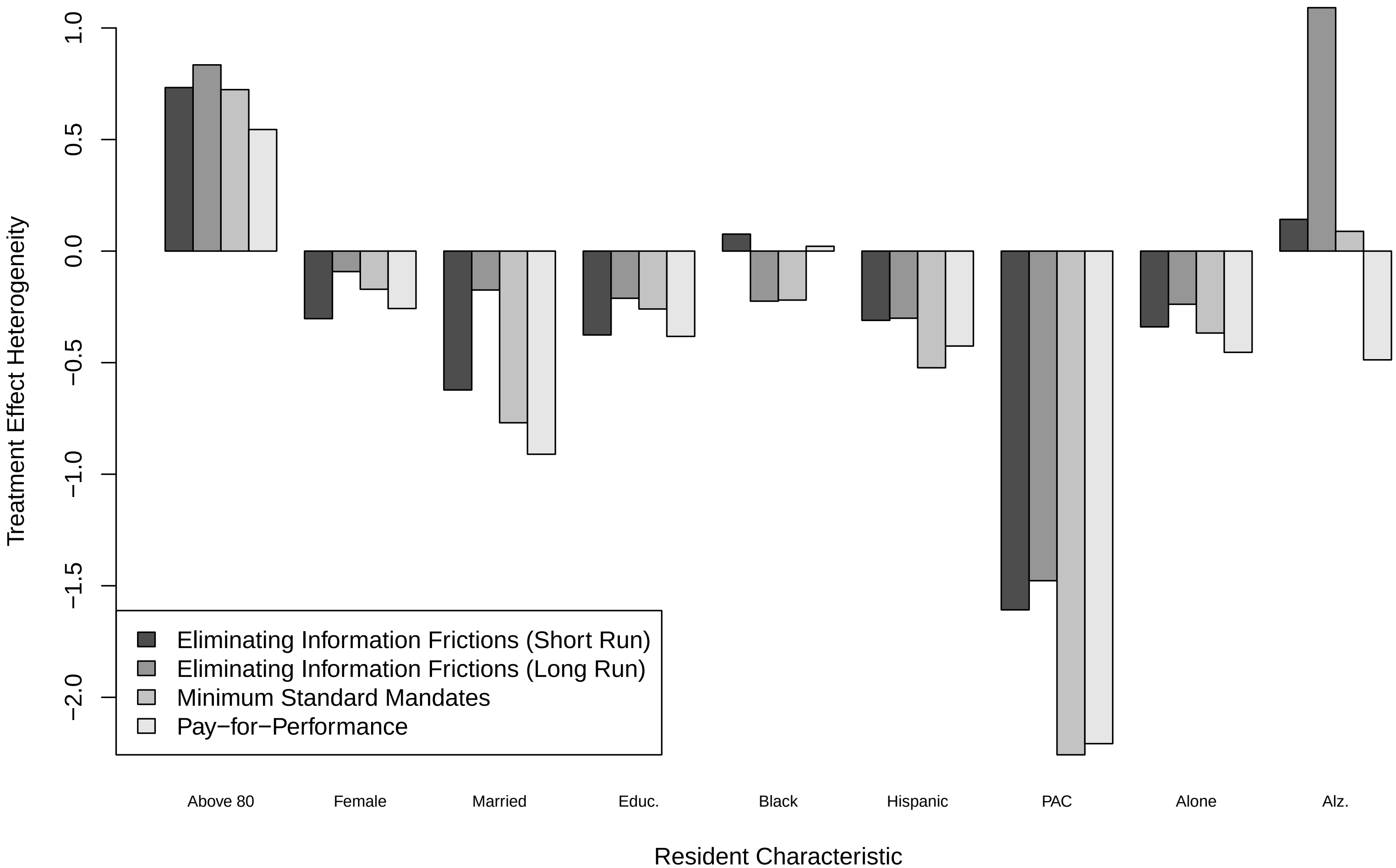
Notes: This figure shows a bin scatter of 90-day resident survival against my occupancy measure, controlling for resident characteristics and nursing home fixed effects. My occupancy measure is defined as the average log occupancy over the 7 days preceding admission residualized of nursing home-month 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. $exp(\mathbf{X}'\beta_{death,haz})$ in equation (12)). The unit of observation is a resident.

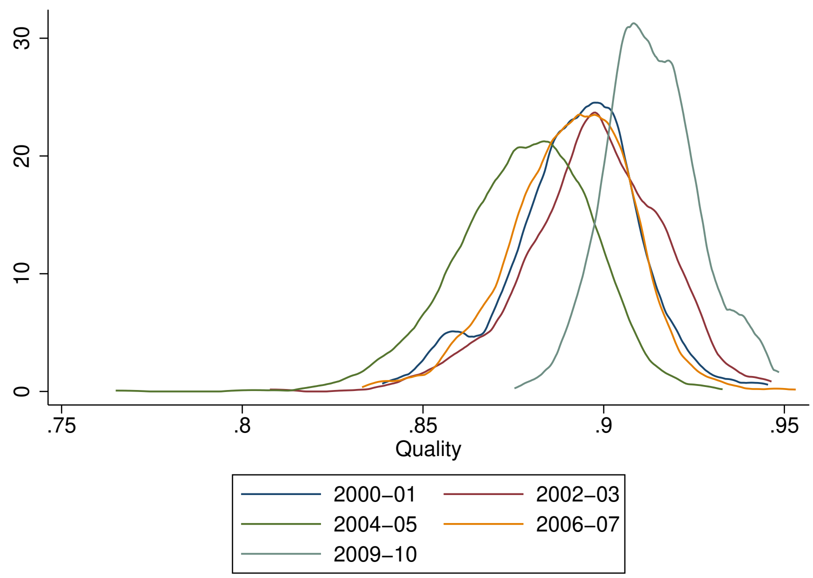
Figure A.18: Treatment Effect Heterogeneity by Resident Characteristics



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.

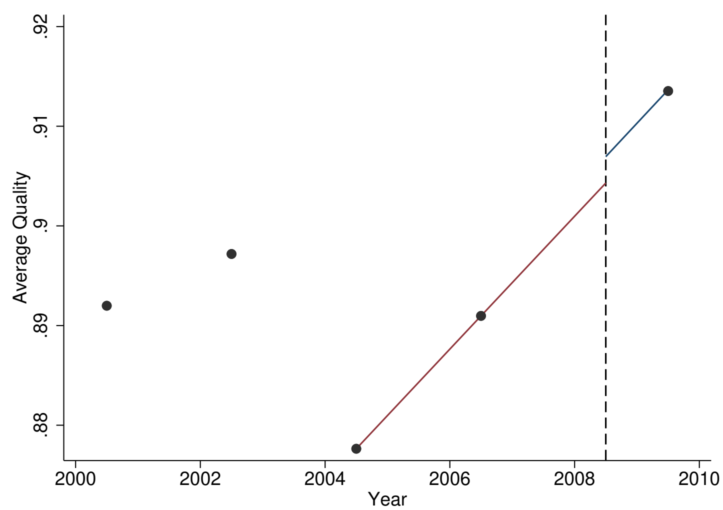
Figure A.19: Effect of Introduction of Five-Star Ratings on Quality

(a) Distribution of Quality Before/After Star Ratings



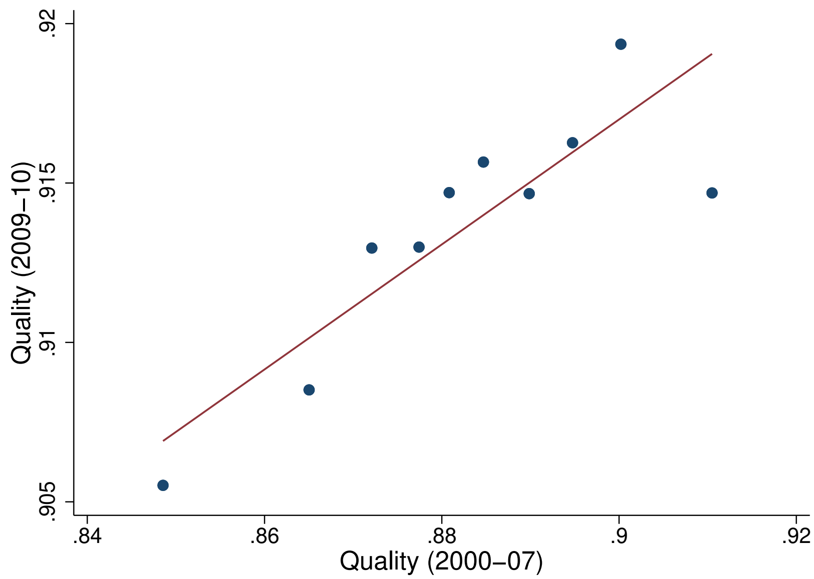
Notes: This figure contains kernel density plots of the quality estimates based on observations from different periods of the sample, using Epanechnikov kernels. The estimation procedure is the same as for the main quality estimates α_j. The unit of observation is a nursing home, and nursing homes are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

(b) Discontinuity in Quality Trend



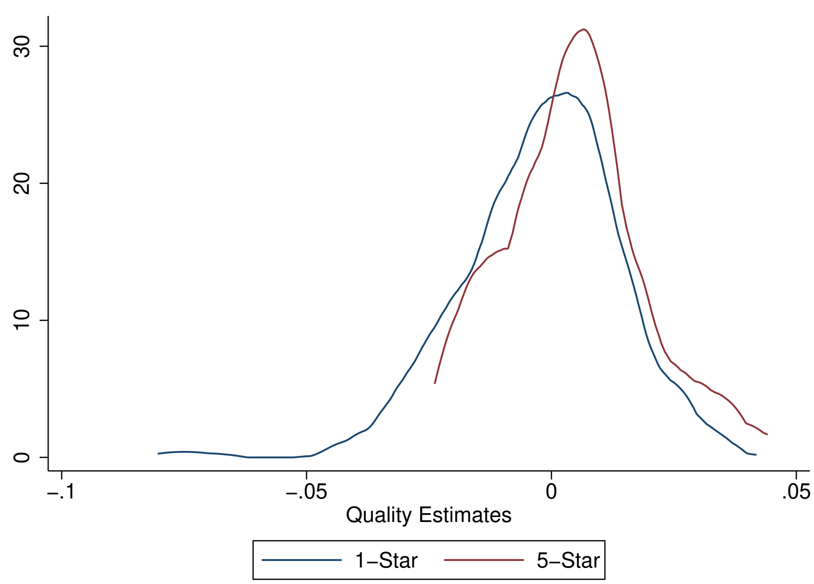
Notes: This figure plots a time series of the average quality estimates for nursing homes during different periods of the sample. A linear trend is added, with the bandwidth chosen so that the estimated discontinuity in the quality before and after the star ratings was introduced at the end of 2008 is 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 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

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

Table A.2: Controls for Resident Characteristics at Admission (continued)

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

Table A.3: Additional Summary Statistics

(a) Additional Summary Statistics for Residents

Residents (N=653,946)			
Age	77.541	Medicare	0.618
	(12.991)		(0.486)
Female	0.611	Medicaid	0.128
	(0.488)		(0.334)
Married	0.329	Admitted from Acute Care Hospital	0.889
	(0.470)		(0.314)
White	0.735	Admitted from Home	0.081
	(0.441)		(0.273)
Black	0.068	Death Within 90 Days	0.075
	(0.252)		(0.264)
Hispanic	0.113	Short Term Memory Issues	0.476
	(0.317)		(0.499)
High School/Some College	0.634	Long Term Memory Issues	0.278
	(0.482)		(0.448)
At Least Bachelor's Degree	0.131	Dementia	0.233
	(0.338)		(0.423)

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)

Nursing Homes (J=840)	
Number of Beds	105.065
	(48.194)
Occupancy Rate	86.958
	(7.493)
Chain	0.600
	(0.411)
For-Profit	0.877
	(0.309)
Deficiencies	6.259
	(3.541)
RN hours per resident day	0.331
	(0.281)

Nursing homes are weighted by number of years that the nursing home was in the sample for.

Table A.5: IV Specification Using Both Variation in Distance and Temporary Occupancy Fluctuations as Instruments

	Dependent Variable: 90-Day Survival				
	(1)	(2)	(3)	(4)	(5)
TSLS					
Forecast Coefficient, λ	0.880	0.911	0.925	0.926	0.958
	(0.115)	(0.097)	(0.088)	(0.085)	(0.083)
Controls for Resident Characteristics	X	X	X	X	X
County Fixed Effects	X	X	X	X	X
Number of Nearest Nursing Homes	1	2	3	4	5
Kleibergen-Paap rk Wald F statistic	116	71	51	39	32
Number of Observations	632,178	625,374	619,845	614,323	607,518

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

	Dependent Variable: 90-Day Survival		
	(1)	(2)	(3)
TSLS			
Forecast Coefficient, λ	0.909	0.909	0.857
	(0.119)	(0.104)	(0.0986)
Demographic Controls	X	X	X
Health Controls	X	X	X
County Fixed Effects	X	X	X
Number of Nearest Nursing Homes	1	2	3
Kleibergen-Paap rk Wald F statistic	39	24	16
Number of Observations	586,483	536,936	490,154

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 clusted at the nursing home level are shown in parentheses.

Table A.7: Relationship Between Unshrunken Quality Estimates and Nursing Home Characteristics

	Quality Estimates (s.d.)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2009 Star Ratings (s.d.)	0.0816							0.0494
	(0.0347)							(0.0356)
RN hours per resident day (s.d.)		0.097						0.211
		(0.0353)						(0.0502)
LPN hours per resident day (s.d.)			0.107					0.117
			(0.0226)					(0.0344)
CNA hours per resident day (s.d.)				0.0962				-0.0386
				(0.0181)				(0.0284)
Deficiencies (s.d.)					-0.0527			-0.0252
					(0.0175)			(0.0148)
For-Profit (s.d.)						-0.0825		-0.061
						(0.0314)		(0.0286)
Chain (s.d.)							-0.0555	-0.0377
							(0.0308)	(0.0296)
N	10,103	10,118	10,119	10,112	10,121	10,121	10,121	10,089
R-squared	0.007	0.010	0.012	0.009	0.003	0.007	0.003	0.032

Notes: This table shows correlations between the nursing home quality estimates (based on the nursing home fixed effects from an OLS regression) 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 residents admitted to the nursing home for their first stay over the sample period. Standard errors are clustered by nursing home.

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

	Indicator for Resident Surviving At Least 90 Days After Admission x 100	
	(1)	(2)
Leave-Year-Out Quality Estimate (s.d.)	1.354	1.339
	(0.0388)	(0.0399)
Deficiencies (s.d.)		0.0239
		(0.0295)
Chain (s.d.)		0.0346
		(0.0504)
For-Profit (s.d.)		0.0714
		(0.0435)
RN hours per resident day (s.d.)		-0.0123
		(0.0507)
LPN hours per resident day (s.d.)		-0.00715
		(0.0405)
CNA hours per resident day (s.d.)		-0.0245
		(0.0304)
2009 Star Ratings (s.d.)		0.0464
		(0.0333)
N	632,223	631,823
R-squared	0.140	0.140

Notes: This table shows results from regressions of 90-day survival on standardized leave-year-out nursing home quality estimates and nursing home characteristics (with coefficients multiplied by 100 for better legibility). 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

	Any New Residents		
	(1)	(2)	(3)
Lagged 7-Day Avg. Log Occupancy	-0.35		
	(0.0451)		
Lagged 7-Day Avg. Occupancy		-0.00518	
		(0.000281)	
Lagged 7-Day Avg. Occ. Percentile			-0.0012
			(5.15e-05)
Nursing Home-Month Fixed Effects	X	X	X
N	1,103,528	1,103,528	1,103,528

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

	Flow of Residents		
	(1)	(2)	(3)
Lagged 7-Day Avg. Log Occupancy	-4.614		
	(0.550)		
Lagged 7-Day Avg. Occupancy		-0.0757	
		(0.00150)	
Lagged 7-Day Avg. Occ. Percentile			-0.0153
			(0.000339)
Nursing Home-Month Fixed Effects	X	X	X
N	1,103,528	1,103,528	1,103,528

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

	Medicaid	Post-Acute Care	Dementia	Age	Female	Married	Black	Hispanic	At Least Bachelor's
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Residualized 7-Day	-0.277	0.0594	-0.0631	4.96	0.0611	0.247	-0.0432	-0.0747	0.0631
Avg. Log Occ.	(0.0459)	(0.0385)	(0.0657)	(1.945)	(0.0799)	(0.0775)	(0.0392)	(0.0502)	(0.0559)
Nursing Home Fixed Effects	X	X	X	X	X	X	X	X	X
N	159,529	159,529	159,529	159,529	159,529	159,529	159,529	159,529	159,529
R-squared	0.218	0.126	0.043	0.145	0.023	0.028	0.180	0.166	0.086

Notes: This table shows results from regressions of various characteristics of admitted residents on fluctuations in nursing home occupancy shortly beforehand. 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

	Medicaid	Post-Acute Care	Dementia	Age	Female	Married	Black	Hispanic	At Least Bachelor's
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Residualized 7-Day	-0.241	0.0269	-0.0889	3.994	0.121	0.244	-0.0346	-0.0522	0.0402
Avg. Log Occ.	(0.0450)	(0.0380)	(0.0631)	(1.849)	(0.0767)	(0.0747)	(0.0386)	(0.0492)	(0.0553)
Nursing Home Fixed Effects	X	X	X	X	X	X	X	X	X
N	159,529	159,529	159,529	159,529	159,529	159,529	159,529	159,529	159,529
R-squared	0.247	0.148	0.118	0.228	0.100	0.097	0.207	0.202	0.107

Notes: This table shows results from regressions of various characteristics of admitted residents on fluctuations in nursing home occupancy shortly beforehand, controlling for other characteristics of the admitted resident. 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

Resident Preferences	
Distance to Facility	-0.159
	(0.015)
Quality	6.879
	(1.264)
Quality x Alzheimer's	-2.293
	(0.419)
Quality x Age	-0.071
	(0.015)
Quality x Female	-0.257
	(0.353)
Quality x Married	-0.192
	(0.371)
Quality x Black	-0.663
	(0.690)
Quality x Hispanic	-2.614
	(0.680)
Quality x At Least Bachelor's Degree	0.997
	(0.509)
Quality x Lived Alone	-0.478
	(0.419)
Quality x Post-Star Ratings	-1.409
	(1.311)
Quality x Post-Star Ratings x At Least Bachelor's	0.522
	(0.635)
Quality x Post-Star Ratings x Black	-1.159
	(0.795)
Quality x Post-Star Ratings x Dementia	-0.160
	(0.495)
Quality x Post-Star Ratings x Age	0.020
	(0.016)
Quality x Post-Star Ratings x Female	0.397
	(0.447)
Quality x Post-Star Ratings x Married	-0.199
	(0.469)
Quality x Post-Star Ratings x Hispanic	0.467
	(0.627)
Quality x Post-Star Ratings x Lived Alone	0.699
	(0.520)
Supply Side	
Temporary Fluctuations in log(occupancy)	-7.49
	(1.08)
Resident Controls in Supply Side Equation	X
Quality x Demographic Variables	X

Notes: This table shows results from the structural estimation using Gibbs sampling. Resident characteristics are included in the supply side estimation, but their coefficient estimates are not shown. 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

	(1)	(2)	(3)	(4)
Temporary Fluctuations in log(occupancy)	-7.34	-7.97	-6.95	-7.38
	(1.07)	(1.24)	(1.37)	(1.10)
Dementia	0.006	-0.069	-0.068	-0.016
	(0.045)	(0.021)	(0.016)	(0.048)
Age	0.027	0.026	0.020	0.027
	(0.009)	(0.008)	(0.008)	(0.009)
Female	0.091	0.074	0.027	0.097
	(0.047)	(0.046)	(0.028)	(0.033)
Married	0.184	0.204	0.164	0.190
	(0.059)	(0.071)	(0.058)	(0.055)
Race (Black)	-0.710	-0.671	-0.594	-0.688
	(0.176)	(0.173)	(0.187)	(0.177)
Race (Hispanic)	-0.145	-0.182	-0.126	-0.180
	(0.033)	(0.045)	(0.024)	(0.026)
Education (Bachelor's)	-0.098	-0.141	-0.118	-0.111
	(0.036)	(0.035)	(0.025)	(0.05)
Admitted from Acute Care Hospital	0.057	0.228	0.179	0.040
	(0.083)	(0.091)	(0.086)	(0.063)
Lived Alone	0.076	0.082	0.052	0.094
	(0.024)	(0.033)	(0.011)	(0.03)
Year (2009)	-0.010	-0.041	-0.001	-0.019
	(0.023)	(0.023)	(0.011)	(0.021)
Year (2010)	0.052	0.063	0.066	0.035
	(0.026)	(0.019)	(0.02)	(0.024)
Constant	-1.31	-1.35	-1.06	-1.32
	(0.56)	(0.48)	(0.46)	(0.54)
Quality in Residents' Utility Equation	X		X	X
Nursing Home Characteristics in Residents' Utility Equation		X	X	
Quality Interacted with Resident Characteristics in Utility Equation				X

Notes: This table shows the supply side estimates from the structural estimation using Gibbs sampling based on the specifications from Table 4 of the main text. A burn-in period corresponding to the first half of the chain was used.

Table A.13: Resident Preferences: Robustness Checks

Panel A: Different Specifications for Instruments	(1)	(2)	(3)
Demand for Quality	0.684	0.752	0.681
	(0.416)	(0.431)	(0.23)
Instruments Interacted with Resident Characteristics	X		
Log Distance		X	
4th Order Polynomial in Distance			X
Panel B: Different Sample Definitions	(1)	(2)	(3)
Demand for Quality	0.647	0.272	-0.411
	(0.436)	(0.284)	(0.215)
Choice Set Includes Nursing Homes Within X Miles	15 miles	10 miles	5 miles
Drop Rapidly Expanding/Contracting Nursing Homes		X	X

Notes: This table shows results from the structural estimation using Gibbs sampling. Demand and supply instruments, the supply side, and the variables indicated in the table are included in the estimation, but their coefficient estimates are not shown. A burn-in period corresponding to the first half of the chain was used.

Table A.14: Demand Heterogeneity: Robustness Checks

	(1)	(2)	(3)	(4)
Resident Preferences				
Utility (Quality x Dementia)	-2.424	-2.221	-2.439	-2.003
	(0.27)	(0.277)	(0.273)	(0.353)
Utility (Quality x Age)	-0.059	-0.067	-0.052	-0.058
	(0.009)	(0.01)	(0.012)	(0.012)
Utility (Quality x Bachelor's)	1.326	1.372	1.376	1.802
	(0.299)	(0.313)	(0.299)	(0.347)
Utility (Quality x Lived Alone)	-0.046	-0.276	-0.067	-0.187
	(0.245)	(0.259)	(0.254)	(0.287)
Utility (Quality x Post-Star Ratings)	0.56	0.518	0.519	0.498
	(0.199)	(0.202)	(0.218)	(0.242)
Demand and Supply Instruments	X	X	X	X
Quality in Utility Equation	X	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. Variables indicated in the bottom section of the table are included in the estimation, but their coefficient estimates are not shown. 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

	Resident Characteristic			
	Dementia	Age	Bachelor's Degree	Post-Star Ratings
Resident Preferences	(1)	(2)	(3)	(4)
Utility (Quality x Resident Characteristic)	-2.523	-0.077	1.021	0.841
	(0.298)	(0.010)	(0.334)	(0.235)
Utility (RN Staffing x Resident Characteristic)	-0.120	-0.004	0.002	0.023
	(0.012)	(0.000)	(0.014)	(0.013)
Utility (LPN Staffing x Resident Characteristic)	-0.126	-0.001	0.066	-0.007
	(0.013)	(0.001)	(0.012)	(0.008)
Utility (CNA Staffing x Resident Characteristic)	0.050	0.003	0.017	0.015
	(0.006)	(0.000)	(0.006)	(0.005)
Utility (Deficiencies x Resident Characteristic)	-0.005	-0.001	-0.010	-0.015
	(0.002)	(0.000)	(0.002)	(0.002)
Utility (For-Profit x Resident Characteristic)	0.024	-0.013	-0.248	0.062
	(0.014)	(0.001)	(0.026)	(0.018)
Utility (Chain x Resident Characteristic)	-0.051	0.001	-0.006	0.010
	(0.009)	(0.001)	(0.016)	(0.009)
Demand and Supply Instruments	X	X	X	X
Quality in Utility Equation	X	X	X	X
Nursing Home Characteristics in Utility Equation	X	X	X	X

Notes: This table shows results from the structural estimation using Gibbs sampling. Coefficient estimates for interactions between various resident characteristics and nursing home variables are shown, with the resident characteristic that the nursing home variables are interacted with being indicated at the top of each column. Demand and supply instruments, supply side variables, as well as main effects for quality and nursing home characteristics in the utility equation are included in the estimation, but their coefficients estimates are not shown. 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

	Dependent Variable: 90-Day Survival	
	OLS	IV
	(1)	(2)
Leave-Year-Out Quality	0.692	0.501
	(0.035)	(0.138)
Leave-Year-Out Quality x Above 80	0.122	0.4
	(0.050)	(0.137)
Leave-Year-Out Quality x Dementia	0.26	0.619
	(0.052)	(0.118)
Leave-Year-Out Quality x At Least Bachelor's Degree	-0.034	0.082
	(0.059)	(0.160)
Leave-Year-Out Quality x Post-Star Ratings	-0.298	-0.562
	(0.056)	(0.133)
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 interactions 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

Assumption on Correlation Between Shocks	Change in Quality	Change in Survival Rate
Perfectly Correlated	0.003	0.005
Independent	0.003	0.021
Number of Observations	57,636	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)

Demand Function	Cost Function	Reduction in Mortality
$\frac{exp(\kappa_j^{quality}\beta_j+\xi_j)}{\sum_l exp(\kappa_l^{quality}\beta_l+\xi_l)}\cdot\bar{N}$	$\left(\frac{c_j^{quality}}{2}\beta_j^2+c_j^N\right)N_j$	44%
$\frac{exp(\kappa_j^{quality}\beta_j+\xi_j)}{\sum_l exp(\kappa_l^{quality}\beta_l+\xi_l)}\cdot\bar{N}$	$\left(c_j^{qual}\cdot\beta_j+c_j^N\right)N_j$	85%
$\frac{exp(\kappa_j^{quality}\beta_j+\xi_j)}{\sum_l exp(\kappa_l^{quality}\beta_l+\xi_l)}\cdot\bar{N}$	$\frac{c_j^{quality}}{2}\beta_j^2+c_j^N N_j$	49%
$A\kappa^{quality}\beta_j+N^p(\bar{p}_{res})$	$\left(\frac{c_j^{quality}}{2}\beta_j^2+c_j^N\right)N_j$	26%
$B\kappa^{quality}log(\beta_j)+N^p(\bar{p}_{res})$	$\left(\frac{c_j^{quality}}{2}\beta_j^2+c_j^N\right)N_j$	26%

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).