The College of Wooster
Open Works

Senior Independent Study Theses

2021

Market Structure and Quality of Service: Investigating Oligopolies and the Quality of Nursing Home Care in California During The COVID-19 Pandemic

Tessa Ireton
The College of Wooster, tireton21@wooster.edu

Follow this and additional works at: https://openworks.wooster.edu/independentstudy

Part of the Economic Theory Commons, Health Economics Commons, Infectious Disease Commons, and the Organizational Behavior and Theory Commons

Recommended Citation

Ireton, Tessa, "Market Structure and Quality of Service: Investigating Oligopolies and the Quality of Nursing Home Care in California During The COVID-19 Pandemic" (2021). *Senior Independent Study Theses.* Paper 9597.

This Senior Independent Study Thesis Exemplar is brought to you by Open Works, a service of The College of Wooster Libraries. It has been accepted for inclusion in Senior Independent Study Theses by an authorized administrator of Open Works. For more information, please contact openworks@wooster.edu.

© Copyright 2021 Tessa Ireton

Market Structure and Quality of Service: Investigating Oligopolies and the Quality of Nursing Home Care in California During The COVID-19 Pandemic

by
Tessa Ireton

Presented in Partial Fulfillment of the
Requirements of Senior Independent Study

Supervised by
Dr. Brooke Krause
Economics

Spring 2021

Abstract

Quality-of-service outcomes in nursing homes are of great social and human importance. However, especially during the COVID-19 pandemic, consistently maintaining markets with high quality care has been a pervading issue in the American nursing home industry. Furthermore, the industry is strongly characterized by oligopolies, a market structure that literature indicates may be less compatible with quality service than competitive markets. With this paper, I aim to investigate the possible intersection of oligopolist market structures and the quality of nursing home care during the COVID-19 pandemic. I start by describing quality of care in nursing homes, particularly during the COVID-19 pandemic, contextualized by literature analyzing nursing home care quality and market structure. Then, I develop a model of profit-maximizing nursing home behavior using theories of oligopolist decision-making, rooted in the basis provided by both Cournot and Bertrand, informed by contemporary models describing the nursing home market. This model demonstrates an opposite correlation between the number of firms in a market and the quality of nursing home care. To test this prediction, I construct a 26-week panel dataset including nursing home attributes and facilities' experiences with COVID-19, using data from California's Agency of Health and Human Services, the California Department of Public Health, the Census Bureau, and the Centers for Medicare and Medicaid Services. Random effects and Hausman-Taylor estimations test the relationship between the number of nursing homes in a market and the outcome of COVID-19 outbreaks in nursing homes, a proxy for quality of care during the pandemic. The results indicate that market structure is not explanatory in understanding differences in quality of nursing home care during the COVID-19 pandemic.

Acknowledgements

I want to thank my parents, Michelle and Bryan, for your love and support throughout my academic career. You instilled in me an appreciation for education that not only drives me in my current pursuits but will continue to motivate me to be a lifelong learner. Thank you for teaching me the value of applying compassion first in all things. I’m so privileged to be your daughter. I also want to thank my sisters, Sophie and Maddy, and my soon-to-be-sister, Sarah. Your friendship has me set for life, which is another great privilege of mine.

Thank you to my advisor, Dr. Brooke Krause, for all your help and guidance, not only in Independent Study, but throughout my economics studies ever since I first sat down in your Principles of Economics class. Ever since then, you have challenged me to sharpen my mind and skills with your wise and patient teaching, which has played an instrumental role in shaping the kind of student I am, and the kind of professional I hope to be.

In addition, I want to thank all of the amazing economics professors at The College of Wooster, particularly Dr. Huiting Tian, Dr. Melanie Long, Dr. Amyaz Moledina, and Dr. Jim Burnell. I know I will continue to be inspired throughout my career by the many lessons I learned from you, and I especially appreciate all of your patience and tenacity in teaching effectively throughout the COVID-19 pandemic. Your leadership and dedication to students during this time is valued and endlessly encouraging.

I am also grateful for the friends I made here at Wooster who have supported me throughout my college career and made my college experience. Thank you for every memory.

Table of Contents

List of Tables

Table of Figures

Introduction

Nursing homes are an increasingly essential healthcare option in the United States. The increased longevity of Americans means that more elderly people survive to the condition of disability that demands long-term care (Glied and Smith, 2011; Grabowski and Norton, 2012). The National Center for Health Statistics estimates that the total number of nursing home users in the United States was about 1.16 million in 2016 (NCHS, 2018). Estimated total national expenditures for nursing home and continuing care services have risen from $800 million in 1960 to about $163 billion to 2016 (CMS NHSG, 2018). Friedberg et al. estimated in 2014 that 27% to 44% of men over 65 years old and 44% to 58% of women over 65 years old will use nursing home services. At the same time, the American nursing home industry is struggling to consistently foster a high-quality market. In 1999, the General Accounting Office released a report of survey data on nursing homes nationally from 1995 to 1996 and indicated that 66% of facilities surveyed were cited for at least one quality deficiency that had the potential to cause harm to a patient or did/could cause actual or serious harm (p. 10).

The implications of lacking care quality are especially dire during the pandemic, as a chief metric of quality of care is infection control. In 2020, nursing homes have been subject to the most severe and tragic institutional COVID-19 outbreaks since the virus' spread in the US began. A CDC case study of a nursing home outbreak of COVID-19 in the Seattle area, which began in late February 2020, reports a staggering fatality rate of about one-third of patients who were infected with COVID-19, with 35 deaths before only twenty days after the first resident tested positive for the virus (McMichael et al., 2020).

Assessing underlying mechanisms as to what contributes to shortcomings in nursing home care quality are all the more relevant during the COVID-19 pandemic. One hypothesis that is under-addressed in the literature is the impact of monopolist or oligopolist market structures on facilities' incentives to provide quality care. The nursing home market may have a strong tendency to be oligopolistic. In Lin's (2015) sample, 87.84% of American nursing home markets had three or fewer facilities. What does this prevailing market structure do to promote better a quality market? Theoretically, fewer firms means fewer choices for consumers, which in turn would disincentivize nursing homes from providing quality services to differentiate themselves and attract market share (Lu, 2014; Grabowski and Town, 2011). This paper investigates the possible intersection between market structure and quality of nursing home care, especially in the context of the COVID-19 pandemic. The main guiding research question is whether oligopolist markets are associated with lower quality of care, and, by extension, whether nursing homes in more competitive markets safer from COVID-19 outbreaks.

This paper will unfold as follows: First, background literature on quality and oligopolies of nursing home care are summarized, which attention to linking these dimensions to the COVID-19 pandemic. Second, this paper will outline a theoretical framework illustrating a for-profit, oligopolist nursing home's quality decision. Third, an empirical analysis will investigate the relationship between market structure and quality of care using data on California nursing homes from both 2019 as a whole and from a 26-week period in 2020. Finally, estimation results are analyzed in light of policy implications for the nursing home industry.

1. Background Literature

This section will provide the literary context that seeks to explain variations in the quality of nursing home care. First, quality is defined, and its connection to COVID-19 outbreaks in nursing homes is established. Then hypotheses that attempt to explain widespread quality shortfalls in the nursing home market are explained, including the market structure hypothesis. Lastly, the causes and indicators of nursing home oligopolies are discussed.

1.1 Quality of care in Nursing Home market

The nursing home industry has a well-observed quality of care problem, which has spanned decades. The U.S. Government Accountability Office has been tracking the problem of nursing home quality for decades based on data from CMS surveys and consumer complaints. They indicated in 2015 that, although average number of serious quality deficiencies per nursing home fell from 2005 to 2014, average consumer complaints per facility rose from 3.2 to 3.9. (Government Accountability Office, 2015). This variance in quality level is also particularly relevant to disease control and public health. A 2020 report from the Government Accountability Office indicated that from 2013 to 2017, 82% of surveyed nursing homes were cited for “infection prevention and control” insufficiencies. Of these, 48% were cited in one or more consecutive years (John, 2020). For positive research, there is much to be investigated as to why this quality problem persists. For normative research, understanding why quality issues persist can lead us to ways the issue can be addressed. The first step in both is understanding what “quality” is and how it could interact with the idiosyncrasies of the nursing home market.

Identifying and defining quality care is an economic challenge. In medical literature, "quality of care" is analyzed from the perspective of individual patients as to whether they "get the care they need and [whether] the care is effective when they need it" (Campbell et al., 2000). In economics, quality is a relative concept that is intrinsically tied to consumer satisfaction with a good or service (Payson, 1994, p. 3). Consumer expectations are critically important to the idea of quality, and different individual consumers have different ideas as to what qualifies as quality. Banker et al. (1998) analyze quality on the basis of a material good's "conformance quality characteristics that are of interest to the consumer when evaluating the product" (p. 1180). The quality of nursing home care could be thought of as the service's conformity to its purpose of purchasing, from the consumers' perspective. This is certainly the case for consumers of nursing home care, as research indicates that nursing home residents have widely variant opinions as to what quality care is. These opinions impact their satisfaction with the care they receive (Bowers et al., 2001).

During a pandemic, the quality of nursing home care is intrinsically associated with controlling disease outbreaks. In response to the ambiguity of patient preferences in defining quality, metrics are often selected to proxy for these preferences. There are, generally speaking, two types of variables that may stand in for quality when quantifiably analyzing quality of nursing home care: "output" or "input" based metrics (Lu, 2014). Notable output-based metrics used include frequency of diagnosed pressure sores and the use of physical restraints, indwelling catheters, and feeding tubes (Grabowski and Castle, 2004). Harrington et al. (2020) find that worse 5-star scores from CMS' surveys, which are mostly based on output quality metrics, predict more COVID-19 cases in nursing homes. The most important output-based metric for this paper's analysis, and used particularly frequently on the part of government scorecards,[1] is infection control. Indeed, it can be reasonably presumed that patients' expectation of quality would include a facility's ability to keep them safe from disease, such as COVID-19.

Further entangling quality of care and COVID-19 outcomes are input-based metrics are frequently used to predict quality of care. The most dominant input-based metric used to proxy quality is some variation of skilled nursing hours per patient day or beds-per-staff ratios (Lin, 2015; Lu et al., 2019; Lu, 2014; Grabowski and Castle, 2004). Facilities' staffing choices likewise are related to severity of COVID-19 outbreaks in nursing homes. Workers are needed in abundance not only to ensure that patients are well-tended to if they are ill, but also to ensure there no shortages of caregivers if infected staff need to stay home (Ouslander and Grabowski, 2020; Bui, 2020). Having ample nursing staff is also important to make sure that infection control protocols are implemented properly (Harrington et al., 2020). Harrington et al. (2020) indicate in their study that greater licensed nursing hours per resident day is associated with fewer COVID-19 patients in nursing home facilities.

In sum, controlling COVID-19 outbreaks is a criterion of quality nursing home care. For the sake of this paper's discussion, this means factors that could have a negative impact on quality of care likewise could have a negative impact on controlling COVID-19 in nursing homes. A few main hypotheses are put forth to explain the inconsistent quality in the nursing home industry, including the role of Medicaid, information transparency, organization type, and, most important for this article's discussion, oligopolist market structures.

[1] CMS' surveys of nursing homes' quality inform facilities' five-star scores on its website Care Compare, formerly Nursing Home Compare, which is allows users to search facilities and review their health, safety, and comfort ratings based on government inspections. It includes infection control as a quality criterion (CMS Care Compare, 2021).

The first of these above hypotheses, the role of Medicaid, is essentially a story of income disparity. Lin (2015) lays out the argument that the nearly monopsonist-like power[2] Medicaid has in buying nursing home care could cause sub-optimal quality outcomes for consumers. Via the purpose of the program and the government's sway in negotiating prices, the government is charged much lower rates than what nursing homes may charge private payers, such as those paying for care out-of-pocket or via a private insurer.[3] Medicaid-insured patients would obviously be partial to high-quality nursing home care rather than low-quality care. However, since the rate the government pays on their behalf is set, even if they have the ability to pay for better care, they may not make use of their willingness to pay.[4] Since providing higher-quality care is expensive for firms, nursing homes may have little incentive, or ability due to high costs, to increase quality level in markets where Medicaid payers are the dominant consumers. As a result, in markets where private payers are dominant, nursing homes may be inclined to provide higher quality care based on consumers' ability, rather than willingness, to pay for it.[5] In these markets, since firms may collect higher fees, it might be more feasible for the nursing homes to provide more expensive, but higher-quality, care. Grabowski and Castle (2004) indicate in their results that higher Medicaid reimbursement rates correlate with better output-based quality measures in nursing homes; that is, when nursing homes are compensated more for providing care to

---

[2] In 2012, 41.7% of nursing home care in the United States was paid for via Medicaid programs (Grabowski et al., 2012, p. 307).

[3] Medicaid reimbursement rates for health care services, including nursing home care, are established by the state governments in which the enrollee resides. These rates are what the government pays to the care provider in return for treating the enrolled patient. These payments to care providers are often much lower than what an insured or otherwise non-eligible person would pay (Bhattacharya, 390).

[4] They likely do not have the ability to pay, as evidenced in their Medicaid enrollment in the first place.

[5] Consumers' willingness to pay is often economically representative of their relative valuation of a good or service. This is a flawed assumption in the case that consumers do not have the ability to pay in accordance with their preferences in the first place.

Medicaid patients, they are more likely to have better quality of care. Keeping with the theoretical hypothesis above, this could be because providing quality is more affordable for nursing homes with higher reimbursement rates. Hackmann (2019) reports similar results for input-based quality metrics.

The second hypothesis, information transparency, pertains to consumers' ability to be selective of the nursing home care they receive. Patients are infrequently able to be discerning when shopping for nursing home care. Older consumer populations are more likely to be infirm and unable to compare nursing homes' quality and be discerning in their purchase (Lin, 2015; Chou, 2002; Grabowski and Hirth, 2003). As a result, nursing homes may have less incentive to provide quality care. In the long run, the level of care quality will eventually be detected by families, but in the short run, consumers must rely on facilities' reputation to choose between them (Lu et al., 2017). This dynamic may give some nursing home patients a disadvantage in detecting poor quality compared to others due to infrequency of family visits (Chou, 2002). This may pose a systemic problem for quality monitoring across the entire industry, since elderly patients with no or few representative family members are more likely to purchase nursing home care (Grabowski and Hirth, 2003).

The third hypothesis to explain systemic shortfalls in nursing home quality is centered on the predominant organization type in the nursing home industry: for-profits. 69.3% of nursing home facilities were run by for-profits as of 2016 (*FastStats—Nursing Home Care,* 2020). Research indicates that for-profit status may be associated with lower quality of care. Lin (2015) observes in their study that for-profits concentrate more in the low-quality facility group. In their study sample, 73% of low-quality nursing homes are for-profit, while only 54% of the high-quality nursing homes are for-profits (1267). Grabowski and Hirth (2003) find that markets with higher shares of nonprofit nursing homes than for-profits exhibit higher quality of care from firms of both organization types. Harrington et al. (2020) indicate that for-profit nursing homes are correlated with greater citations of deficiencies in patient health. There are a couple prevailing hypotheses as to why this is the case. First, perhaps for-profit nursing homes are less inclined to provide high quality of care due to the cost associated with superior nursing home service. For-profit facilities' incentive is to generate a return for their owners, so administration may push for additional profits in lieu of increasing care quality beyond where they think necessary (Davis, 1993; Chou, 2002; Chesteen et al. 2005).[6] Nonprofits, on the other hand, are incentivized by their mission to serve their communities.[7] Moreover, the quality of nursing home care is not easily observed (Chesteen et al., 2005), and Lu et al. (2017) writes that nursing home quality may not even be meaningfully discernable until long after purchase. Because of this, for-profit nursing homes may be able to more easily engage in "opportunistic behavior" and cut quality-related spending while maintaining a high price point, unnoticed (Grabowski and Hirth, 2003; Hirth, 1999). In contrast, seeking nursing home care from a nonprofit may represent a safeguard to consumers against this kind of exploitation because a nonprofit's mission will disincentivize them from opportunistically extracting profit (Chou, 2002; Grabowski and Hirth, 2003; Hirth, 1999).[8] Interestingly, another hypothesis links the problem of information transparency to the dominance of for-profits in the industry. Chou's (2002) research indicates

[6] Lu et al. (2017) argue that nursing homes must offer at least some baseline level of care quality, since nursing homes compete with in-home services and family caregivers.
[7] Service of the community is critical for nonprofits in order to justify their tax-exception benefit.
[8] This idea of a quality guarantee from nonprofit firms in absence of information was originally postulated by Kenneth Arrow in 1963. The idea that nonprofit signals this credibility to consumers was also discussed by Henry Hansmann in 1980.

that information “asymmetric”[9] markets are more likely to exhibit greater differences in quality between nonprofit and for-profit nursing homes. Grabowski and Castle (2004) write that the proportion of nonprofits versus for-profits in a market might actually be an indicator of how informed consumers are; in other words, a high proportion of nonprofits in the market might indicate that the market has a high degree of information asymmetry.

1.2 Oligopolies in the nursing home market

Lastly, market structure could be a contributing factor to sub-quality care in the nursing home industry. Like many health care industries, the nursing home industry is characterized as oligopolistic. This is likely due to several unique factors present in the industry: government regulations exercising control over where facilities may be constructed, extremely high capital and labor costs required to establish and maintain operations, the need for extensive relationships with local doctor and hospital networks to guarantee patients, and “certificate of need” (CON) laws requiring that entrants prove their service provision is necessary for the local community (Gulley and Santerre, 2004). Lack of competition poses possible challenges to promoting a high-quality market. For people contemplating buying nursing home care for themselves or a loved one, it is preferable to have many choices when looking for a suitable care option. When facilities have few competitors, they may be less inclined to compete with each other via service differentiation, i.e., providing higher quality care than their competitors (Lu, 2014; Grabowski and Town, 2011). Even worse, this problem of a captured consumer base may render information-disseminating resources, such as CMS’ Nursing Home Compare, less relevant in producing higher-quality markets

[9] “Information asymmetry” refers to unbalanced access to information regarding the quality of nursing home care between firms and consumers.

(Grabowski and Town, 2011). The lack of competitors means that, even if an oligopolist's care is obviously poor to potential consumers, patients will opt to purchase the facility's care anyway due to a lack of options.

There is little literature that tests the relationship between oligopolist market structures and nursing home care levels directly, and there is no clear consensus. Zinn (1994) finds that more oligopolistic nursing home markets are more likely to exhibit better output-based quality metrics, but Lu's (2014) findings in this area are mixed. In contrast, Lu (2014) also finds that less-oligopolistic markets exhibit better input-based quality-of-care outcomes. Lin (2015) analyzes counterfactuals based to assess the impact of an augmentation of factors that prompt additional firms to enter the market, and Hackmann (2019) analyzes counterfactuals based on their theoretical and empirical models to assess the impact of an artificial addition of nursing homes to markets. Hackmann writes that adding nursing home firms has a small positive impact on input-based quality measures, but Lin's (2015) model indicates that the new entrants would likely be providing low-quality care. The aim of this article is to provide evidence regarding the impact of oligopolist market structure on the quality of nursing home care.

2. Theoretical Framework

This section analyzes the impact of strategic competition between oligopolist nursing homes on the quality of care they offer. As the following demonstrates, an increase in the number of nursing homes in the market may incentivize firms to increase the quality of their care to profit-maximize. This model draws mainly from Perloff's (2017) and Mas-Colell et al.'s (1995) constructions of Cournot's oligopoly model. It also incorporates Perloff's (2017) model of a duopoly with differentiated products, as well as Banker et al.'s (1998) strategy for including quality in the cost and demand functions for oligopolist firms. The price function is informed by Lu et al.'s (2017) linear price function for an oligopolist nursing home firm.

This model assumes linear functional relationships, so that the direction of the relationships between variables can be clearly and explicitly expressed. This section analyzes two scenarios in the demand structure of the nursing home market, namely, the different impacts of publicly- and privately- insured demand segments on nursing homes' decision-making. This approach addresses an idiosyncrasy of the nursing home market, the presence of Medicaid-insured consumers, and their demand segment's impact on quality of care in the context of strategic interaction. First, the market conditions for nursing home firms are established, with special attention to nursing homes' choice parameters. Then two scenarios hypothesizing nursing home decision making are described, one for the "regulated" segment of nursing home care, and the other for the "private-payer" segment. Lastly, the implications of the cost of providing quality nursing home care are discussed in light of this chapter's model.

2.1 The conditions for an oligopolistic Nursing Home Market

A market for nursing home service exists, with boundaries determined by the community. One of the nursing home firms, NH_1, is the firm of interest in this analysis, whose decision-making is this chapter's reference point for analyzing the impact of strategic market conditions on the quality of care it offers. There are others, such as $NH_2, NH_3, \ldots\ NH_n$, as there are n nursing home firms in the market. All nursing homes in the market have identical capacity to provide services; they have equally free reign to choose the levels of their choice variables in order to maximize profit. All nursing homes have the same set of options available to them in crafting their production of service, including equal opportunity to enhance quality.[10]

2.1.1 Nursing Homes' choice variables

NH_1's strategic decision-making is summarized by choosing the levels of four variables: its capacity, q_1, the quality of its services, x_1, and the level of inputs employed in providing care, K_1 and L_1.

The output of a nursing home firm, q_1, is understood in this example as the capacity of their care services. In the context of nursing homes, the term "capacity" could conflate the varying numbers of patients they can serve at a time or the various types of care they can provide. In this model, the term "capacity" has a specific interpretation. This chapter's example assumes that the type of nursing home and the kind of services they offer are

[10] The nursing homes have access to the same kinds of quality-enhancing labor, since they operate in a shared labor market. Realistically, if one firm in any market were to make a decision to enhance quality by hiring as many specialized nurses as possible, that firm would then have the chance to try and absorb all these workers in the local labor market, which would remove this opportunity from the other firm's set of decisions. This model assumes that firms are only competing with each other via the setting of choice variables, and that there would be ample skilled nurses for hire such that all nursing home firms are not concerned about competing over a niche labor pool.

identical between NH_1 and all other facilities. For these firms, capacity refers to the number of patients they may provide their services to at a time.

In producing these services, a nursing home employs two basic inputs: capital, K_1, and labor, L_1. Capital refers to the medical or residential facilities, medical machinery, and equipment employed in providing nursing home care, standardized by the hours of usage and averaged across capital type. Labor refers to the physician hours or nursing hours employed by the firm in caring for patients in the medical or residential setting.

Quality, x_1, is the most relevant variable in this study's research question and is the most abstract of the choice variables. As Payson (1994) describes, in one sense, quality refers to the characteristics or attributes of one firm's good or service that make it more or less desirable than another's. Therefore, levels of service quality must be relative among firms. The literature that analyzes the quality of nursing home care consistently attributes quality of care differences among nursing homes to the variance between skilled nursing employed by facilities. This study will focus on the impact of skilled labor[11] that enhances the quality of care and treatment patients receive.

Notice the initial absence of the price NH_1 charges for care among this initial set of choice variables. As Scanlon (1980) points out, the total demand for nursing home care is "segmented" among different payer types, since nursing home care is rarely directly paid for by those receiving care. The government's role in deciding the rates Medicaid will pay for care sets the price of nursing home care for a large portion of consumers in the "regulated"

[11] In the real world, quality enhancements certainly come from the quality of capital employed in care as well, but this study focuses on the role of skilled labor-hours because of its significance in the literature, simplicity in analysis, and because of the data available. Moreover, the author argues that in many cases skilled labor personnel are required to employ specialized medical instruments that one might identify as high quality or quality-enhancing.

demand segment. Since nursing home firms have limited control over the prices they may charge for Medicaid-insured consumers, price competition is not feasible to attract patients that are Medicaid-insured. This study postulates that competition in service quality and output level, rather than price, is the mode of strategic interaction for firms engaging with this segment of demand. The price NH_1 may charge, whether it is the rate decided by the government on behalf of Medicaid-insured consumers, or by uninsured consumers' willingness to pay, is exogenous to nursing homes. This is the first scenario analyzed by this chapter. However, for the segment of consumers that are not Medicaid-insured, firms are not subject to the same restrictions in terms of pricing. Price competition for personally-paying or privately-insured consumers is possible, so for this "private-payer" segment, nursing home firms are sensitive to consumers' willingness to pay as they freely determine price levels. This is the second scenario analyzed by this chapter. In both cases, other nursing homes' decided levels of output and quality, q_2 and x_2 are exogenous to NH_1.

2.1.2 Determination of the consumers' willingness to pay for $\boldsymbol{NH_1}$

Due to the possibility of differences among NH_1 and the other nursing homes' quality levels, these nursing homes services are imperfect substitutes of each other. The output and quality each nursing home produces impacts consumers' marginal willingness to pay for the other nursing homes' services, but each firm can be thought of as facing different demand schedules for every level of service capacity they offer. In other words, each firm exists in a different, but related, market, each with a different demand function. Consumers' marginal willingness to pay for NH_1's care, p_1, is given by the following price function, where n represents the number of nursing homes in the market:

$$p_1 = \gamma_1 - \frac{1}{x_1} q_1 - \frac{\sigma}{x_1}\left(\frac{q_2}{x_2} + \cdots + \frac{q_n}{x_n}\right) \tag{1}$$

where $x_1, x_2, \dots x_n > 0$, and $\sigma < 0$

This function draws from Banker's (1998), Perloff's (2017)[12] and Lu et al.'s (2017) linear price models. Borrowing from Banker et al.'s (1998) price-demand function, the baseline demand for its nursing home services is given by γ_1. It represents what non-Medicaid-insured consumers would be willing, if able, to pay for nursing home services if all firms produced nothing. Price falls as firms produce more because of the Law of Demand; as a good becomes more available in the market, consumers' relative valuation of it compared to other goods falls, reducing their willingness to pay.

Similar to Lu et al.'s price equation, the reciprocal of NH_1's quality level is used as its capacity's coefficient because the higher a nursing home's service quality is, the more desirable those services are to consumers.[13] As a result, a unit increase in output would result in a smaller negative impact on consumers' willingness to pay for high-quality services than for services of a lower quality level. The variable σ, like Lu et al.'s "degree of horizontal differentiation,"[14] is a factor to represent the relative substitutability between NH_1's services and the services of all other nursing homes.[15] If $\sigma = 0$, there is no substitutability between NH_1's care and the care of other nursing homes, since the quantity other nursing homes provide to the market has no impact on consumers marginal willingness to pay for NH_1's nursing home care, p_1. If $\sigma > 0$, NH_1 and the other nursing homes offer services that are imperfect substitutes to consumers, because a change in other nursing homes' capacities has

[12] Specially, Perloff's linear model of a Cournot oligopoly.
[13] Although, the causality of this relationship would be such that the desirability of the services would determine the level of quality.
[14] See the section in the appendix A.6 for analysis of Lu et al.'s 2017 article, "Does Competition Improve Quality?"
[15] This model assumes that consumers would have identical relative evaluation of NH_1's care to all other nursing home firms; that is, σ is the same number for all nursing home firms.

a nonzero, negative relationship with price consumers are willing to pay for NH_1's care.[16] The variable x_1 denominates σ's term because the higher the quality level of NH_1's care, the less relevant the other nursing homes' capacity levels are in determining consumers' valuation of NH_1's care. In other words, holding the quality levels of other nursing homes' care equal, the quality of NH_1's care mitigates the substitutability of its care and other nursing homes' care. The quality levels, x_1 and x_2, cannot be zero because it is inconceivable for a nursing home to offer no level of quality in this framework. To do so would mean that the nursing homes' services have zero conformity with consumers' expectations of nursing home care. In that case, consumers would not have any willingness to pay for those services, and their demand would be relegated elsewhere by their preferences and budgets.[17]

2.1.3 Nursing homes' production function

This framework assumes that nursing homes produce services based on the levels of two fundamental inputs employed by the firm, labor, L, and capital, K, as they are applied in the Cobb-Douglas production function, per Perloff (2017, Chapter 6) and per Mas-Colell et al. (1995, Chapter 5):

$$q_1 = f(L_1, K_1) = EL_1^a K_1^b \tag{2}$$

in the case of NH_1. Per Perloff (2017, Chapter 6), a and b represent either input's scaled impact on production. They represent the overall significance L_1 and K_1 might have in the production processes among NH_1 and other nursing homes is taken as given. Also given is E, a factor that accounts for innovative efficiencies. This factor allows for two identical

[16] In this model, even when the quality levels of all firms are the same, the capacity of other nursing homes' care services has a smaller impact on consumers' valuation of NH_1's services than NH_1's capacity of service. In other words, even when all nursing homes offer equivalent quality, their services are still imperfect substitutes.
[17] This is important in the real-world context that nursing homes compete with outside industries as well, such as in-home care.

nursing homes with the same input levels to possibly have different service capacities due to other productive processes besides the magnitude of inputs. Per Mas-Colell et al. (Chapter 5, 1995), the Cobb-Douglas production function presents other advantages relevant in the nursing home market,[18] including clear representation of a nursing home's returns to scale from its use of inputs.[19]

2.2 Maximization of NH_1's Profits

Nursing homes in this model are for-profits, so they will engage in profit-maximizing behavior. NH_1's profit, π_1, is given by:

$$\pi_1 = R_1 - C_1 \tag{3}$$

The first term represents the revenue, R_1, reaped by NH_1 and the second term is the total cost, C_1, of the firm to produce.

2.2.1 Revenue

Revenue, R_1, is given by:

$$R_1 = p_1 q_1 \tag{4}$$

Where NH_1's price, p_1, may be substituted by Equation (1).

$$R_1 = \left(\gamma_1 - \frac{1}{x_1} q_1 - \frac{\sigma}{x_1}\left(\frac{q_2}{x_2} + \cdots + \frac{q_n}{x_n}\right)\right) q_1 \tag{5}$$

or

[18] It also abides by a number of other properties necessary for "production sets" (Mas-Colell et al., 1995, Chapter. 5) or "isoquant functions," including the premises that using more inputs guarantees more output and that each input's contribution to an overall increase in output exhibits diminishing returns to scale (Perloff, 2017, Chapter 6).

[19] The nursing home's production process exhibits constant returns to scale if $a + b = 1$, decreasing returns to scale if $a + b < 1$, or increasing returns to scale if $a + b > 1$. Returns to scale refers to the percentage total capacity rises as a result of a percentage increase in inputs. For example, increasing returns to scale means that a percent increase in inputs causes an even greater percentage increase in service capacity.

$$R_1 = q_1\gamma_1 - \frac{1}{x_1}q_1^2 - \frac{\sigma q_1}{x_1}\left(\frac{q_2}{x_2} + \cdots + \frac{q_n}{x_n}\right) \tag{6}$$

more simply,

$$R_1 = q_1\gamma_1 - \frac{1}{x_1}q_1^2 - \frac{q_2\sigma q_1}{x_2 x_1} - \cdots - \frac{q_n\sigma q_1}{x_n x_1} \tag{7}$$

2.2.2 Total Costs

NH_1's total costs of bringing the goods to market are given by:

$$C_1 = rK_1 + wL_1 + \mu x_1 L_1 \tag{8}$$

A nursing home firm's costs are dependent on their input- and quality-associated spending. $rK_1 + wL_1$ represent the firm's baseline spending to deliver service provisions. It represents the production costs the nursing home would incur if it were possible for NH_1 to exist with a quality level of zero. The term $\mu x_1 L_1$ represents the firm's quality-dedicated costs, or the costs the firm incurs to maintain its chosen quality level, x_1. μ represents the premium rate for using quality-enhancing labor hours. As mentioned previously, nursing home quality is associated with skilled labor, so quality-associated labor costs are emphasized in this model.

2.2.3 The Regulated Segment: Profit Maximization via Lagrangean Maximization Method

The expanded function for NH_1's profit, created by combining Equations (4), (7), and (8), is

$$\pi_{1,regulated} = q_1\gamma_1 - \frac{1}{x_1}q_1^2 - \frac{q_2\sigma q_1}{x_2 x_1} - \cdots - \frac{q_n\sigma q_1}{x_n x_1} - rK_1 - wL_1 - \mu x_1 L_1 \tag{9}$$

subject to $q_1 = EL_1^a K_1^b$

NH_1 will seek to maximize its profit, but it is constrained by the condition that q_1 must be equivalent to the production function, $EL_1^a K_1^b$. Given this constraint, the Lagrangian method of maximization can be used similar to Albouy's (2005) technique. From this method, a system of equations may be derived in the context of profit maximization, and a clearer picture of the relationship between x_1 and the other choice variables is drawn.

First, the Lagrangean equation, $\mathcal{L}(\cdot)$, takes the profit function and adds the constraints with the factor, λ:

$$\mathcal{L}(q_1, x_1, L_1, K_1, \lambda) = q_1\gamma_1 - \frac{1}{x_1}q_1^2 - \frac{q_2\sigma q_1}{x_2 x_1} - \cdots - \frac{q_n\sigma q_1}{x_n x_1} - rK_1 - wL_1 - \mu x_1 L_1 + \lambda(L_1^a K_1^b - q_1) \tag{10}$$

Four partial derivatives of $\mathcal{L}(\cdot)$, are taken in respect to each choice variable:

$$\frac{\partial\mathcal{L}(\cdot)}{\partial q_1} = \gamma_1 - 2q_1\left(\frac{1}{x_1}\right) - \frac{q_2\sigma}{x_2 x_1} - \cdots - \frac{q_n\sigma}{x_n x_1} - \lambda = 0 \tag{11}$$

$$\frac{\partial\mathcal{L}(\cdot)}{\partial K_1} = -r + \lambda EbL_1^a K^{b-1} = 0 \tag{13}$$

The first-order condition of profit maximization may be fulfilled by setting each of these derivatives equal to zero. Then, λ may be solved for from Equation (11):

$$\lambda = \gamma_1 - 2q_1\left(\frac{1}{x_1}\right) - \frac{q_2\sigma}{x_2 x_1} - \cdots - \frac{q_n\sigma}{x_n x_1} \tag{15}$$

From Equation (12):

$$\lambda = r\left(\frac{1}{EbL_1^a K^{b-1}}\right) \tag{16}$$

Under the circumstance of profit maximization, Equations (15), (16) and (17) may be solved as a system to isolate the optimal choice variables and identify their relationships with other choice variables. Since x_1 is the primary strategic choice variable of interest to this IS, x_1^*, or x_1 when profit is maximized, is solved for. Equation (15) may be set equal to Equation (15), which sets λ equal to itself:

$$\gamma_1 - 2q_1\left(\frac{1}{x_1}\right) - \frac{q_2\sigma}{x_2x_1} - \cdots - \frac{q_n\sigma}{x_nx_1} = r\left(\frac{1}{EbL_1^aK^{b-1}}\right) \tag{18}$$

From here, x_1^* can be identified as:[20]

$$x_1^* = \frac{2q_1^* + \frac{q_2\sigma}{x_2} + \cdots + \frac{q_n\sigma}{x_n}}{\left(\frac{-r}{EbL^*{}_1^aK^{*b-1}}\right) + \gamma_1} \tag{19}$$

Equation (19) identifies x_1^* as a function of the other choice variables, and the direction of the relationships between them is evident. In the case of the Medicaid-insured portion of the market, the addition of more firms correlates with an increase in x_1^*, NH_1's quality level. The addition of nursing home firms to the market drives down quality level because of the additional capacity they bring to the market, even though their care services are imperfect substitutes of NH_1's care. Consider the following equation, which gives the optimal capacity level for NH_1, q_1^*, derived from Equation (18):

$$q_1^* = \frac{1}{2}\left(x_1\gamma_1 - \frac{rx_1}{EbL^*{}_1^aK_1^{*b-1}} - \frac{q_2\sigma}{x_2} - \cdots - \frac{q_n\sigma}{x_n}\right) \tag{20}$$

As more nursing homes enter the market, the optimal level of service capacity NH_1 may offer falls, since the market service provision demanded by consumers of NH_1, γ_1, is now divvied among more firms, which are relatively substitutable. The higher the quality level, x_1, the stronger the negative effect of an adding another nursing home firm on NH_1's capacity. *Ceteris paribus,* the more capacity NH_1 offers, the more quality would be affordable at the optimal level of capacity, as seen with the positive sign on q_1^* in Equation (19). However, if other nursing home firms as a whole produce an increasing amount of care capacity, NH_1's current capacity level must fall, and the quality of that care must rise, to

[20] Proof included in appendix A.7.

maintain profit-maximization. This mathematical effect is harmonious with the hypothesis that as additional firms enter the market, each nursing home will be more incentivized to provide quality in order to differentiate themselves (Lu et al., 2017).

2.2.4 The Private-payer Segment: Profit Maximization via Lagrangean Maximization Method

For the demand segment comprised of privately insured or out-of-pocket buyers of nursing home care, nursing home firms are less subject to price regulation as when selling care to Medicaid-insured buyers. Therefore, price competition is on the table as a mode of strategic interaction. In this case, NH_1 is subject to a slightly different set of choice variables. Price replaces quantity as a parameter, since NH_1 will provide whatever output the market demands at the optimal price, p_1^*.[21] Before the Lagrangean can be derived, a new profit function that accounts for this change in choice variables must be found. Recall the price function, which determines consumers' willingness to pay for NH_1's care given relevant parameters:

$$p_1 = \gamma_1 - \frac{1}{x_1}q_1 - \frac{\sigma}{x_1}\left(\frac{q_2}{x_2} + \cdots + \frac{q_n}{x_n}\right) \tag{1}$$

where $x_1, x_2, \ldots x_n > 0$, and $\sigma < 0$

Since price, p_1, is now a choice variable, rather than NH_1's capacity, q_1, NH_1's capacity must be isolated from Equation (1):[22]

[21] This assertion is inspired Varian's (1993, pg. 461-2) description of a Bertrand oligopoly, where competing firms "bid" for the business of consumers via price levels, and the firms accept the market's "decision" as to what quantity the firms will provide. However, Bertrand's oligopolist equilibrium presumes that all the products will be identical, which is a central assumption when analyzing how Bertrand equilibrium is the same as the competitive equilibrium (462). This model cannot make this assumption due to the necessity that nursing home firms offer different quality levels, and therefore nonidentical services. Therefore, it is ambiguous whether the equilibrium brought forth by this chapter's case could likewise be close to competitive equilibrium.
[22] See appendix A.7.

$$q_1 = -x_1 p_1 + x_1 \gamma_1 - \frac{\sigma q_2}{x_2} - \cdots - \frac{\sigma q_n}{x_n} \tag{21}$$

This function gives the capacity demanded of NH_1 by the market given the other parameters, including the price NH_1 selects. From here, a profit new function to maximize via the Lagrangian method is derived. Given the baseline revenue function:

$$R_1 = p_1 q_1 \tag{4}$$

The new revenue function that substitutes Equation (21) for q_1 in Equation (4) is

$$R_1 = -x_1 p_1^2 + p_1 x_1 \gamma_1 - \frac{\sigma q_2 p_1}{x_2} - \cdots - \frac{\sigma q_n p_1}{x_n} \tag{22}$$

Therefore, a new profit function for this private-payer scenario would be:

$$\pi_{1,private-payer} = -x_1 p_1^2 + p_1 x_1 \gamma_1 - \frac{\sigma q_2 p_1}{x_2} - \cdots - \frac{\sigma q_n p_1}{x_n} - rK_1 - wL_1 - \mu x_1 L_1 \tag{23}$$

from combining Equation (22) and Equation (8). The new Lagrangian in this case is:

$$\mathcal{L}(p_1, x_1, L_1, K_1, \lambda)$$

$$= -x_1 p_1^2 + p_1 x_1 \gamma_1 - \frac{\sigma q_2 p_1}{x_2} - \cdots - \frac{\sigma q_n p_1}{x_n} - rK_1 - wL_1 - \mu x_1 L_1 + \lambda(L_1^a K_1^b$$

$$+ x_1 p_1 - x_1 \gamma_1 + \frac{\sigma q_2}{x_2} + \cdots + \frac{\sigma q_n}{x_n}) \tag{24}$$

where Equation (21) is substituted for q_1 as the constraint multiplied by λ. The most relevant partial derivative in this case would be the following, set equal to zero to fulfill the first-order condition of profit maximization:

$$\frac{\partial \mathcal{L}(\cdot)}{\partial p_1} = -2x_1 p_1 + x_1 \gamma_1 - \frac{\sigma q_2}{x_2} - \cdots - \frac{\sigma q_n}{x_n} + \lambda x_1 = 0 \tag{25}$$

Solving for λ from Equation (25) yields:

$$2p_1 - \gamma_1 + \frac{\sigma q_2}{x_1 x_2} + \cdots + \frac{\sigma q_n}{x_1 x_n} = \lambda \tag{26}$$

To solve for the optimal quality level, Equation (26) may be set equal to Equation (16), since the maximized $\frac{\partial \mathcal{L}(\cdot)}{\partial K_1}$ would be the same in this scenario as it would be in the regulated case.

$$2p_1 - \gamma_1 + \frac{\sigma q_2}{x_1 x_2} + \cdots + \frac{\sigma q_n}{x_1 x_n} = r\left(\frac{1}{EbL_1^a K^{b-1}}\right) \tag{27}$$

From here, the optimal quality level as a function of the other choice variables can be identified:[23]

$$x_1^* = \frac{\frac{\sigma q_2}{x_2} + \cdots + \frac{\sigma q_n}{x_n}}{\frac{r}{EbL_1^a K^{b-1}} - 2p_1 + \gamma_1} \tag{28}$$

As the number of firms in the market increases, the optimal quality level of NH_1 rises. Again, the mechanism behind this result is the increased levels of output. As more firms enter, the price that NH_1 may charge must fall as other nursing homes bring additional service capacity to the market. This is explicit in Equation (29):

$$p_1^* = \frac{1}{2}\left(\gamma_1 - \frac{\sigma q_2}{x_1^* x_2} - \cdots - \frac{\sigma q_n}{x_1^* x_n} + \frac{r}{EbL_1^{*\,a} K_1^{*\,b-1}}\right) \tag{29}$$

However, as Equation (29) demonstrates, increases in NH_1's quality level mitigates the negative impact of an increase the number of nursing home firms in the market on NH_1's price. When choosing the optimal price and quality level, NH_1 is incentivized to raise its quality level as more rivals enter the market. This finding is again consistent with the hypothesis of product differentiation as a mode of competition among oligopolists. Reducing the substitutability between similar goods or services, such as nursing home care, may act as a prop for price levels by fortifying consumers' willingness to pay for a particular nursing home's care.

[23] See appendix A.3.

2.3 Quality costs and NH_1's decision-making

The ability of Medicaid reimbursement rates to cover the cost of providing quality of care is critically important to understanding the impact of an increase in firms on the overall quality of nursing home care in a market, particularly for the regulated demand segment. Lin (2015) writes that nursing homes may only be incentivized to offer quality care in order to attract privately insured patients, which may be more profitable than Medicaid-insured patients. The implication is that Medicaid reimbursement is not sufficiently incentivizing for profit-seeking nursing homes. Lin writes that, in the case that both privately and publicly insured consumers are competing in the market, the entrance of more nursing homes might cause quality levels to fall because more firms are competing for the same number of private payers. As more firms enter the market, each nursing home will receive less private-payer revenue, which reimburses at higher rates than Medicaid, thus rendering them less able to provide equal levels of quality care. In this scenario, as Lu et al. (2017) write, nursing home firms rely on revenue from private-payer patients to cover the quality-associated costs of providing care to both privately insured patients and publicly insured patients.

What does this hypothesis offer predictions of care quality? Per Lin (2015), this premise would hold that, as more firms enter the market, quality of care will fall if Medicaid reimbursement rates, or service revenue per Medicaid patient, are less than the cost of providing quality level, x_1, per Medicaid patient. This chapter's theoretical analysis is limited in that that it assumes that Medicaid-established prices, and therefore nursing home firms' reimbursement for providing care to the regulated demand segment, are enough to cover the cost of care. Without formally deriving fulfillment of profit-maximization, it is indeed the case that NH_1would be disincentivized from increasing quality level beyond a minimum necessary to maintain consumers' baseline willingness to pay for NH_1's care *in what case*. In the regulated price scenario, the profit function for NH_1 could be described the same as in Equation (23), where profit, π_1, is a function of price, rather than quantity:

$$\pi_1 = -x_1 p_1^2 + p_1 x_1 \gamma_1 - p_1\left(\frac{\sigma q_2}{x_2} - \cdots - \frac{\sigma q_n}{x_n}\right) - rK_1 - wL_1 - \mu x_1 L_1 \qquad (30)$$

If we assume the baseline level of quality is $x_1 = 1$, quality-associated costs would be equal to μL_1. If $p_1 < \mu L_1$, that is, if the Medicaid reimbursement rate is less than quality associated costs, Equation (30) generally denotes that NH_1 could possibly have little no incentive to increase quality level, lest they suffer a decrease in profit. An increase in x_1 by 1 would increase profit by $p_1\gamma_1$, but decrease it by $p_1^2 + \mu L_1$. Normally, NH_1 would increase quality level while $x_1 p_1 \gamma_1$ magnifies the set price sufficiently to cover the costs of increasing quality and net profit. But in this case, this would never happen, since for every increase in x_1, or quality level, the cost terms $p_1^2 + \mu L_1$ would always rise in a manner that outstrips the growth of the term $p_1\gamma_1$.

An important underlying assumption, which is critical to this outcome in the model, is that the costs of providing quality will stay the same as more firms enter. This chapter assumes that quality-associated costs are unique to labor costs, expressed in the term $\mu x_1 L_1$, rather than capital costs. These assumptions are obviously a simplification of real-world input markets. In the real world, there are likely capital-related quality associated costs as well. Also, Banker et al.'s (1995) paper discusses a scenario in which capital costs fall as firms enter the market. Theoretically, this is a logical alteration, since increased competition for factor inputs should put downward pressure on the prices firms must pay for inputs. While Banker et al.'s discussion of this hypothesis is in the context of all inputs, perhaps the same phenomenon could exist in certain circumstances for quality-enhancing costs, both labor and capital. If this is the case, as more firms enter the market and select nonzero quality levels, their increased competition for quality-yielding inputs will lower the individual costs to each firm in providing quality. At a certain threshold, it might be the case that the cost of providing quality could fall below the price in the situation where quality costs initially exceeded the price of quality care.

2.4 Discussion of the model

The above model regarding the case of Medicaid- and non-Medicaid-insured consumers illustrates a possible theoretical avenue though which oligopolist markets, defined by a small number of firms, negatively impact the quality-of-care nursing homes provide. The additional care capacity new firms provide to the market may incentivize individual nursing homes to increase quality either to shield themselves from downward pressure on prices or to differentiate their care capacity from rivals' services. In a real-world sense, firms may tend to differentiate their product in more competitive markets in order to help their good or service stand out among numerous competitors and attract buyers. While this tactic is certainly used in the context of "spurious product differentiation" (Perloff 2017, p.503), in which the difference between goods is only perceived, there's no reason the same push toward differentiation would not lend to a real quality increase. This may be the case in a market environment subject to high exogenous expectations of information transparency,[24]

[24] Nursing homes must report a slew of information to the government, such as number of licensed beds, proportion of payer types (Medicaid, Medicare, private insurance, etc.), ownership type, and "process of care" benchmarks. This information is incorporated into CMS' publicly available on the Online Survey Certification and Reporting (OSCAR) database (pre-2012) and on the Certification and Provider Enhanced Reporting (CASPER) database (2012 and post) (Cowles Research Group, 2020). Also, nursing homes must keep track of summary information on their own patients, including diagnosis-related and socio-economic details, called "minimum dataset (MDS) requirements" (Grabowski and Norton, 2012, p. 310). In 2002, CMS launched its Nursing Home Quality Initiative (NHQI) with the goal of increasing information transparency among nursing home firms, patients, and their families. Central to this Initiative is the Care Compare tool on CMS' website, which allows consumers to search for specific facilities and see how CMS has rate their quality them based on a

such as the nursing home market. Oligopolist markets may be dampening the quality potential of nursing home care. Of course, differences in quality have a heavier implication in the case of nursing home care than they do in consumer goods. The health of vulnerable people is at stake when a nursing home makes its quality decision. In the case of a global pandemic, the outcome of that decision could be a matter of survival. Considering the great expense of providing quality nursing home care, the possible negative relationship between the number of firms on the costs of providing quality care is of great interest to this study's research question and could be considered in a future study's theoretical framework.

five-star scorecard (Nursing Home Quality Initiative, 2020; Find and compare Nursing Homes: Nursing Home Compare, 2020).

3. Data

Six data sources comprise the dataset used for this study. The first is the Long-Term Care Financial Data from the California Office of Statewide Health Planning and Development (OSHPD) (CHHS, 4 December 2020). This dataset contains detailed information regarding supply-side characteristics of long-term care facilities across California in 2019.[25] For analysis and interpretation, this study relied on CHHS' provided disclosure report and data file specification, as documentation (CHHS, 3 May 2018; CHHS, 22 January 2019). Prior to analysis, certain observations were dropped to make the cross sections more comparable. Some facilities face different reporting standards and are not required to submit the full amount of documentation to OSHPD. These are marked as "non-comparable", as opposed to "comparable" datapoints that draw from the fully submitted paperwork. Observations not marked "comparable" were dropped. Also, long-term care facilities with licenses that do not match this study's broad definition of "nursing home" were likewise dropped.[26] Facilities whose reporting periods did not cover the whole year were also dropped. From the remaining observations and values, total services revenue, total Medicaid (Medi-Cal, in California) services revenue, and the proportion of 2019 services revenue attributable to Medicaid were calculated. The number of competitor firms was calculated by subtracting 1 from the total number of nursing homes in the same county as the facility of interest. Each nursing home's total patient census days for 2019 was also collected from this dataset.

[25] This data is in an "as-submitted" state; it contains data self-reported by Californian nursing homes that has not yet been audited by the state.
[26] See appendix A.2 for details on which facility licenses were considered "nursing homes" for this study.

The second source is CMS' COVID-19 Nursing Home Dataset, complied by the Centers for Medicare and Medicaid Services (CMS), which contains information regarding individual nursing homes' responses to COVID-19 and outcomes across the United States.[27] The data used are the archived dataset covering weeks ending in 5/31/20 though 11/22/20, with each of these 26 weeks constituting a period of time for this analysis' time-series (CMS, November 2020). This analysis relied on the data dictionary documentation provided by CMS (CMS, January 2021). All non-Californian nursing homes were dropped. This set provided variables regarding the week-to-week status of each facility's COVID-19 response, including variables indicating different types of staff shortages, the ability of nursing homes to COVID-19 test all residents or get the resources to COVID-19 test all residents within the week, and dummies indicating shortages of resources that prevent disease spread, such as masks and hand sanitizer. When merging the nursing home observations in this dataset and in the Long-Term Care Financial Dataset, some discretion was used on the part of the author.[28] Also, all observations that did not pass CMS' data quality assurance check were dropped (CMS, *Frequently Asked Questions*, 2020).

---

[27] See the following document for details on nursing home's reporting requirements and relevant survey questions: Director, Quality, Safety & Oversight Group. (2020). *Interim Final Rule Updating Requirements for Notification of Confirmed and Suspected COVID-19 Cases Among Residents and Staff in Nursing Homes*. Centers for Medicare & Medicaid Services. https://www.cms.gov/files/document/qso-20-29-nh.pdf

[28] A two-step merging process was used. First, the nursing home observations were merged on the basis of identical facility names, identical names of the California counties in which they are located, and date marking the end of the week. After this step, a large number of nursing homes, which did have a match in the opposing set, could not be matched due to idiosyncrasies in the naming conventions and the formats of other identifying data. For this analysis, nursing homes were only merged if there is reasonable certainty that the two facilities are the same. This included the facilities having the same or nearly the same name, the same county, and the same address. Address and facility name discrepancies were manually checked after coding for common deviations in convention, such as "CENTER" versus "CENTRE," "REHAB" versus "REHABILITATION," and the inclusion or omission of firm type identifiers such as "INC." or "LLC." More nuanced differences between nursing home identifiers included discrepancies in the type of facility evident in the name. Observations showing these kinds of discrepancies were not merged unless it was reasonably clear that the facility was the same between the two, even after the county and addresses were matched. Using a fictional example, if a nursing home from one set was called "Squirrel Hill Care Center," it would not be merged with "Squirrel Hill Post-Acute Rehabilitation" because it is unknown whether the "Care Center" serves the same patients with the same management and staff as the "Post-Acute Rehabilitation" facility.

The next dataset used is the State of California Data Portal Case Data for California COVID-19 Response (California COVID-19 Response, 2021). The surrounding severity of COVID-19 spread is an important control in understanding nursing home outbreaks, according to Ouslander and Grabowski (2020) and White et al. (2020), so this dataset provides a control for the newly confirmed COVID-19 cases in the county. The data are reported on a daily basis, so they are compressed to the weekly level for the time period 5/31/20 though 11/22/20, dropping new cases that are not assigned to a county or are marked "out of the country."

The last three data sources are from the US Census Bureau and aim to cover critical demographic control variables. These variables are necessary for contextualizing the demand-side factors individual nursing homes face. The first is the Small Area Income and Poverty Estimates (SAIPE), State and County Estimates for 2018, the California file (US Census Bureau, December 2019). This analysis uses this dataset's information about each of California's counties' median household income during 2018 to control for income disparity and its plausible impact on the quality of nursing home care available. Second is the Census Bureau's County Population by Characteristics data, Vintage 2019, the California file, for Annual County and Resident Population Estimates by Selected Age Groups and Sex. (US Census Bureau, 2019). Referencing their methodology document (US Census Bureau, March 2020) and data dictionary (US Census Bureau, June 2020), this analysis uses this dataset's information on California counties' elderly populations to extract each county's proportion of its population over the age of 65, for the time period 3/1/2019 to 7/1/2019. The last Census Bureau set used is from the United States Population Estimates, Vintage 2019; the California file for Annual County Resident Population Estimates by Age, Sex, Race, and Hispanic

Origin, likewise for 3/1/2019 to 7/1/2019 (US Census Bureau, 2019). Referencing the same documentation as the prior file, this analysis uses this dataset's information on California counties' total population estimates and estimates of the percentage of the population that is African American.[29]

The following figure models the final, merged dataset used by this paper:

Table 1: Model of Dataset[30]

Nursing Homes id	**Date** (week interval)	**For-Profit dummy**	…	**County 2019 elderly percent**	**...**	**Weekly new COVID-19 cases in the nursing home per 1,000 residents**	**...**
1	Week 1	0	…	20%	...	0	...
1	Week 2	0	…	20%	...	32	...
…	…	…	…	...	...	...	...
926	Week 25	1	…	31%	...	401	...
926	Week 26	1	…	31%	...	445	...

Variables that are the same for all time periods but are wholly unique to each nursing home for all time periods are highlighted in brown.
Variables that have the same value for all observations in the same county are in green.
Variables that are gathered on the weekly basis, and are unique to each nursing home, are highlighted in purple.

[29] Percentage of the estimated population that are "Black or African American alone or in combination" with another ethnicity (US Census Bureau, June 2020).
[30] The full list of variables and their explanations is in appendix A.1

4. Empirical Models

This section covers the empirical models used in this analysis. First, variable choice is justified, with special attention to control variables recommended by the literature. Then, the estimation techniques used are justified. For this dataset, the primary challenge in finding an appropriate empirical model is the presence of both time-variant and time-invariant variables in the dataset. This presents a challenge when controlling for bias. Given the variables in the set, omitted variables could include the foresight of administrators in planning for the pandemic's arrival in the US, relative magnitude of the nursing staff's compensation, the administrator's outside priorities or incentives, etc. These "nursing home effects" would be ideally be captured by a fixed effects model, but many important variables in the model are time invariant. As a result, a fixed effects model would presume perfect multicollinearity among most of this dataset's variables, ruining their predictive relevance. In response to this problem, two alternative specifications are proposed: Random effects and Hausman-Taylor estimation.

4.1 Main independent variable: modeling oligopoly

Different methods may be used to differentiate between competitive versus oligopolist markets, or to identify competitive intensity in a nursing home market. The first method a is a Hirschman-Herfindahl index (HHI), which measures the concentration of market share among the firms in the market. Zinn (1994) calculates this by using the number of beds in the market as the basis of the total market output for nursing home care. Market share is defined as the proportion of total beds in the market each nursing home firm has. The index is a value from 0 to 1 derived by adding the squared proportion of market shares for each firm in the market, where 1 would represent a monopolist's market. Zinn (1994) and Lu (2014) use this method as a dependent variable to represent the structure of a nursing home market. A second method is to simply use the number of firms in the market, which is favored by Lin (2015), and Lu et al. (2017). The bounds of the market can be identified by a radius outward from the location of a nursing home (Lu et al., 2017), or the boundaries of the county in which it resides (Lu, 2014; Hackmann, 2019; Zinn, 1994; Lin, 2015). This paper favors the use of counties to represent market boundaries, since county-level data is more readily available.

This analysis uses both market concentration and the number of rival firms in its modeling, with the primary specification using the number of rival facilities in a nursing home's county as a measure of competition. Also, secondary specifications use market concentration, measured via Hirschman-Herfindahl indices, as the independent variable of interest. In this study, two different HHIs are calculated in separate ways. First, HHI for 2019 as a whole is calculated as the sum of the squared proportion of each nursing home's portion of the 2019 total patient census days in the county. Using this index, market power as it relates to market share is more closely tracked than the number of rivals, but only for 2019 as a whole. The second HHI is calculated on a weekly basis for each California county during the period 5/31/20 though 11/22/20. In this case, market share is defined as the nursing home's proportion of the total bed occupancy in the county during each week. Market concentration is then measured by the sum of the squared proportion of each nursing home's portion of the total current nursing home bed occupancy in the county during the week.

4.2 Main dependent variable: modeling quality of care in respect to COVID-19

To test the relationship between oligopolist market structure and the quality of nursing home care, the dependent variable of interest is the nursing home's weekly new confirmed COVID-19 cases among residents, per 1,000 residents. This will represent an outcome-based proxy for quality of care, since the quality of care during a pandemic includes the expectation that consumers can expect relative safety from COVID-19. Therefore, this paper's analysis aims to address whether a greater number of rival firms in the county, or a lower HHI, may correlate with a higher quality level among nursing homes' care and possibly lower rates of COVID-19 cases.

4.3 Control variables

The next set of variables represent controls for individual nursing homes' attributes or their markets' attributes. Since prior literature indicates that a facility's labor decision-making is correlated to quality outcomes, specifically COVID-19-related quality outcomes, the first of these controls address the facility's employment of labor and the labor supply it accesses. The main specification includes variables indicating whether the nursing home firm is experiencing weekly shortages in RNs, clinical staff, nursing aides, and other staff aim to control each nursing home's access to precious labor while caring for patients during the pandemic. A secondary specification uses the nursing home's 2019 productive labor hours of different specialty types, divided by their total patient census days in 2019. Moreover, a dummy variable that indicates whether the nursing home is a for-profit firm, or "investor-owned," is included. This data is only available in the set for the year 2019, but since the organization type likely does not change frequently, it can reasonably be interpreted as current ownership status in subsequent time periods included in the dataset. The current occupancy of the nursing home is an important control according to He et al. (2020) and Abrams et al. (2020), so the facility's 2019 total patient census days is included. Lin (2015) finds that nursing homes with more beds correlate with greater quality of care after a certain threshold, so the hypothesized sign for patient census days is negative. The main specification will also use a variable for the nursing home's 2019 proportion of total services revenue from California's Medicaid program, Medi-Cal. This is aimed at controlling for the hypothesis that the greater the dependence of the nursing home firm on Medicaid reimbursement, the worse their quality of care will tend to be due to cost prohibitions. To control for nursing homes' varying access to COVID-19 testing, there is also a dummy variable indicating whether a nursing home is able to test all of its residents for COVID-19 within seven days, or if it has the ability to obtain resources to do so. Similarly, to control for the differential in resources that similar-quality nursing homes may have as they manage COVID-19 in their facilities, there is a dummy variable for whether the nursing home has a week's supply of N95 masks.

Next, the model includes variables that aim to control for attributes of a nursing home's market, proxied by the county in which it resides. These variables describe the characteristics of the market the nursing home operates in. Among these are the proportion of the population that are African American in the county. This is an important control due to the well-documented fact that African Americans are disproportionately affected by the COVID-19 pandemic in the United States (Shah et al., 2020), and Lu et al. (2017) use this as a control in their study concerning the quality of nursing home care. The proportion of the county population that are elderly, which is the proxy for the overall demand for nursing home care used by Lu et al. (2017) and Lin (2015). Median household income is important to control since wealthier patient populations are more likely to have greater affordability for better nursing home care. The new COVID-19 cases in the county are also used as a control, which Ouslander and Grabowski (2020) and White et al. (2020) indicate is an important when assessing the causes of COVID-19 severity in long-term care facilities.

Below is a simplified regression equation to illustrate the goals of the empirical models in this paper.

$$New\ confirmed\ COVID\ 19\ Cases_i = \beta_0 + \lambda X_{it} + \gamma Y_{it} + \alpha Z_{it} + \epsilon_{it}$$

On the left-hand side, new confirmed COVID-19 cases in the nursing home i is the dependent variable. On the right-hand side is the intercept, β_0, and the coefficient and independent variable of interest, market structure, shown as λX_{it}. The remaining terms are the coefficients and vectors for the nursing home-specific and market-specific variables, γY_{it} and αZ_{it} respectively, and the error term, ϵ_{it}.

4.4 Estimation technique 1: Random Effects

A random effects (RE) model is proposed because RE allows the coefficients on time-invariant variables to be estimated and these variables can be controlled for. This is because RE presumes that the constant, or intercept, of each nursing home is on a random distribution, the mean of which is the same for all time periods. The downside of this model is that it does not control for omitted variables associated with the cross-sections, i.e., unique effects attributable to individual nursing homes. RE assumes that the impact of individual nursing homes' uniqueness on the dependent variable is random and uncorrelated with any of the other independent variables. This is problematic because individual nursing home effects that may exist could correlate with other independent variables, such as the organization type of the firm, the scale of their operation, and the percent of revenue from different sources, such as private-payer revenue and Medicaid revenue. Nevertheless, RE could be a viable option for testing the data because it allows for meaningful inclusion of the main independent variables.

4.5 Estimation technique 2: Hausman-Taylor

A Hausman-Taylor (HT) model allows both time-variant and time-invariant variables to be used, and it addresses possible correlation between the nursing home error terms and the time-invariant variables (Muck, 2018). HT is able to do this by instrumenting the coefficients of the time-invariant variables based on the residuals of a fixed effects estimation of only the time-variant variables, via a two-stage least squares regression (Muck, 2018; Aadland, 2009; *xthtaylor,* 2021).[31] In order for this method to presume accuracy, there must be the same amount, or more, of variables that are exogenous and time-variant than there are variables that are endogenous to nursing home effects and are time-invariant (*xthtaylor,* 2021). The nursing homes' for-profit status and the total patient census days in 2019 are both probably closely endogenous to nursing home effects. This analysis assumes that new COVID-19 cases in the county and the weekly COVID-19 testing ability of the facility are exogenous to the nursing home's decision-making.[32] Therefore, the number of exogenous, time-variant variables could be the same as endogenous, time-invariant variables, and fulfill this condition.[33] Endogenous time-variant variables include the staff shortage dummies, the dummy indicating whether the nursing home is experiencing a shortage in N95 masks, and

[31] This paper will not delve into the mathematics of the HT estimation; see Muck (2018), Aadland (2009) and the STATA manual (*xthtaylor*, 2021). This analysis referred to these sources in its execution of the HT estimation and its robustness tests in section 5.
[32] Nursing homes' ability to test for COVID-19 arguably has less to do with the decision-making of the firm, especially early in the pandemic when supplies were scarcer; testing capacity for asymptomatic people is triaged to nursing homes (CHHS California Department of Public Health, 22 September 2020).
[33] The proportion of revenue that comes from Medicaid reimbursement could be argued as endogenous or exogenous. Nursing homes payer-mix could have an impact on their quality decision,

the weekly HHI, which is the independent variable of interest in this model. The next assumption necessary for employing HT is that the exogenous variables must be correlated with the endogenous, time-invariant variables. This assumption is tested in section 5.2 as a robustness test of the HT used.

5. Estimation Results

This section presents the results and insights presented by the estimation of the dataset. First, descriptive statistics are reviewed, then the estimation results from random effects and Hausman-Taylor models are discussed.

5.1 Summary statistics

The dataset contains 939 nursing homes, with data spanning the 26-week period between 5/31/20 and 11/22/20, inclusive. The set is unbalanced, with 23,205 observations, each noting the state of a nursing home at one of 26 weeks. These nursing homes reside in 49 of California's counties.[34] 87.9% of these nursing homes are for-profits.

Reviewing the dataset gives some insights concerning the main variables of interest to this study. There are only nine counties in which there is a nursing home monopoly, and only 117, or 12.5%, of the nursing homes in the dataset reside in a county with less than 10 nursing homes. More importantly, oligopolies and monopolies appear to pertain to a relatively small number of nursing home residents in California. For the time period between 5/31/20 and 11/22/20, the nursing homes in the set reported 1,778,951 weekly occupied beds; that is, the sum of each nursing home's total occupied beds at each of the 26 weeks.[35] Only 26,855, or 1.5%, are weekly occupied beds in nursing home markets with weekly HHIs greater than 0.5.

---

[34] The counties not included in this dataset are Alpine, Glenn, Mariposa, Modoc, Mono, Plumas, San Benito, Sierra, and Trinity counties. These counties are not in the set because there are no nursing homes in the dataset that are located in these nine counties.

[35] This does not mean 1.8M total patients were served by the set's nursing homes during the 26-week time period. If the same patients occupied a NH's bed two weeks in a row, they would be counted twice in this measure. This number is simply meant to capture the total output, or capacity of the nursing homes on a weekly basis, during the 26-week time period.

The following table gives the descriptive statistics for the time-invariant variables in the set.[36]

Table 2: Summary statistics: Time-Invariant Variables[37]

	Minimum	Maximum	Median
Total patient census days 2019	718	133,275	31,868
Medicaid (Medi-Cal) services rev. / total services rev.	0	.779	.005
Productive Hrs / Tot. 2019 patient census days, GNPs	0	.035	0
Productive Hrs / Tot. 2019 patient census days, RNs	0	6.642	.341
Productive Hrs / Tot. 2019 patient census days, LVNs	0	5.239	.962
Number of rival firms	0	335	63
HHI in 2019	.004	1	.019
% population in county elderly	10.48	28.376	14.082
Total population in county	18,039	10,039,107	2,470,546
% population African American	1.253	17.657	9.977
Median household income 2019	45,086	125,933	67,986

The following table gives the descriptive statistics for the non-dummy time-invariant variables. The time variance is on a weekly basis between 5/31/20 and 11/22/20.

Table 3: Summary statistics: Time-Variant Variables, not dummies[38]

	Min.	Max.	Mean	Std. Dev.	Median
HHI weekly	.004	1	.063	.125	.022
Total occupied beds	0	555	76.333	39.77	71
New cases per 1000 residents	0	902.4	8.942	41.237	0

The following table gives the means of the time-variant dummy variables. Again, variance is on a weekly basis between 5/31/20 and 11/22/20.

[36] Mean and standard deviation are not shown, since they constitute the means and medians for all time periods. Since the values covered by the time-invariant variables are the same for all of the NH's time periods, and the dataset is unbalanced, mean and standard deviation are not reliable for describing the nursing homes in the dataset.
[37] GNPs are geriatric nurse practitioners, RNs are registered nurses, and LVNs are licensed vocational nurses.
[38] Weekly new COVID-19 cases in the NH's county are not included in this table, since the data would be heavily skewed toward later weeks, and it is arguably fairly exogenous to both the NH's decision-making and the demand for the NH's care.

Table 4: Summary statistics: Time-Variant Variables, dummies

	Mean
Testing ability of NH	.989
RN shortage	.012
Clinical staff shortage	.003
Nursing aide shortage	.019
Other staff shortage	.005
Has week supply N95	.917
Has week supply surgical mask	.947
Week supply hand sanitizer	.933

The measures of market structure captured above again suggest that the California nursing home market is not very oligopolistic. The median for both HHI in 2019 as a whole and for the weekly HHI during the time period, 0.019 and 0.022 respectively, indicate that the county-based nursing home markets tend to be fairly competitive. These measures of HHI are based on market share defined as both number of occupied beds (weekly HHI) and the patient census days (2019 as a whole). The number of rival firms tells a similar story, with the median at 63 rivals. Californian nursing homes also appear to have ample ability to test their residents for COVID-19, with 98.9% of nursing homes able test all their residents or able to obtain the supplies to test all their residents within a given week. Reported staff shortages are uncommon, and facilities appear to often have ample access to surgical masks, hand sanitizer, and N95 masks. Also, nursing homes in this set appear to have little reliance on Medicaid revenue as a whole, with the 2019 median percent services revenue from Medi-Cal at 0.5%. This suggests that, should quality of care shortfalls pervade among nursing homes in the dataset, as a whole they might not be largely attributable to the low Medicaid reimbursement rates.

5.2 Estimation results

The table below gives the results of four random-effects regressions. The first two columns use the number of rival firms as the independent variable to account for market competition. The first column uses the weekly dummies indicating a staffing shortage to account for each firm's labor decisions, and the second column substitutes 2019 productive labor hours per patient census day as the skilled labor control. While the staff shortage dummies provide insight into the availability of skilled labor as nursing homes combat COVID-19, they are perhaps less reflective of the firms' typical quality-related decision-making than the 2019 labor hours, since facilities' behavior could be unusual given the pandemic. The third and fourth rows address competition in terms of market concentration; the third row uses 2019 HHI to account for market structure, and the fourth row uses weekly HHI. Some control variables used are omitted from the table below for brevity. These variables' coefficients and standard errors are included in the appendix A.3.

Table 5: Random effects Estimation of COVID-19 outbreaks in nursing homes[39]

	Main IV: # of rivals	Controls with 2019 Labor Hours	**Main IV:** 2019 HHI	**Main IV:** Weekly HHI
# of rival firms	-0.0173	-0.0260		
	(0.0392)	(0.0379)		
NH is for-profit	4.461***	3.903***	4.491***	4.484***
	(1.503)	(1.325)	(1.500)	(1.498)
Total Pop. county 2019 Apr.-Jul.	-1.64e-07	1.59e-07	-8.16e-07***	-8.32e-07***
	(1.37e-06)	(1.33e-06)	(1.79e-07)	(1.80e-07)
Median household income in county, 2018:	-7.96e-05***	-7.90e-05***	-8.98e-05***	-9.26e-05***
	(2.93e-05)	(3.03e-05)	(3.18e-05)	(3.22e-05)
New confirmed COVID-19 cases in county	0.000282***	0.000278***	0.000281***	0.000282***
	(9.56e-05)	(9.53e-05)	(9.58e-05)	(9.58e-05)
Shortage of RNs	16.91		16.87	16.83
	(12.37)		(12.37)	(12.37)
Shortage of clinical staff	-37.66*		-37.77*	-37.77*
	(21.42)		(21.39)	(21.38)
Shortage of Aides	8.267		8.255	8.259
	(6.021)		(6.015)	(6.012)
HHI in 2019			-5.238	
			(5.780)	
Productive Hrs. / Tot. 2019 patient census days, GNPs		565.5***		
		(78.40)		
P. Hrs. / Tot. 2019 PCD, RNs		-0.144		
		(1.960)		
P. Hrs. / Tot. 2019 PCD, LVNs		-0.303		
		(1.337)		
Weekly 2020 HHI, 5/31-11/22				-6.126
				(5.243)
Constant	8.406	8.873	9.258	9.491
	(7.262)	(7.202)	(7.082)	(7.102)
Observations	11,088	11,093	11,088	11,088
Number of nursing homes	807	807	807	807

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

For-profit nursing homes appear to be positively correlated with increased COVID-19 cases in nursing homes, which was predicted by the hypothesis that for-profit nursing homes may be less incentivized to provide quality care. However, the magnitude of this effect is striking. The coefficients in the four columns above indicate that, when there are

[39] A number of control variables were included in these specifications but not listed: see appendix A.3 for the coefficients and standard errors for these variables.

new COVID-19 cases in a nursing home, there are about three to five more new COVID-19 cases per 1,000 residents in for-profit nursing homes. Median household income also predicably shows negative association with COVID-19 cases; counties with more affluent households are less likely to see more cases. This could be due to residents' increased ability to pay for quality care, which is consistent with the Medicaid-reimbursement hypothesis. Interestingly, however, Medicaid reimbursement as a proportion of total services revenue is not significantly associated with new COVID-19 cases in either direction.[40] Evidently, a nursing home's relative reliance on Medicaid revenue may be less associated with the facility's quality decision. Per the aforementioned hypothesis, this lack of effect may indicate that Medicaid reimbursement in California is possibly sufficient to cover the costs of providing quality care.

The regression estimates indicate some unexpected results. The variables accounting for facilities' skilled labor decisions are included as a control for input-based metrics of quality, but it is unexpected that the results above indicate there is little correlation between shortages of staff and new COVID-19 cases. In fact, employment of GNPs in 2019 has a positive correlation with new cases, and the coefficient is ludicrously high. This is could be due to a bias introduced by virtue of the fact that only 47 observations had nonzero values for this variable. These 47 observations are likely a poor sample of nursing homes (and time periods) where GNPs are employed by facilities and reflect outcomes that are not representative of the observations as a whole. Moreover, the direction and magnitude on the variable indicating a shortage in clinical staff is large and counterintuitive. Likewise, only 75 observations report a shortage of clinical staff, biasing the coefficients. Lastly, metrics of

[40] See appendix A.3 for the SEs and the coefficients for this control.

market power and competition have no significant correlation with quality outcomes proxied by COVID-19 outbreaks.

The table below reports the results of the HT estimation of new COVID-19 cases in nursing homes, per 1,000 residents, with some included variables omitted from the table, likewise omitted moved to the appendix A.3. No specification was constructed using the number of rivals nor the 2019 HHI as the independent variable of interest, since the inclusion of more endogenous, time-invariant variables would leave the HT underspecified. Below, the independent variable of interest is the weekly HHI from 5/31/20 to 11/22/20.

Table 6: HT estimation of New COVID-19 Cases in NH, per 1,000 residents

	HHI is Independent Variable of Interest, Hausman-Taylor Estimation
Time variant, exogenous variables	
New COVID-19 cases in the county	0.000290***
	(9.54e-05)
Testing ability of the NH	9.448
	(13.36)
Time variant, endogenous variables	
Weekly HHI	-55.63
	(69.38)
RN shortage	27.64*
	(15.83)
Clinical staff shortage	-87.37***
	(33.89)
Nursing aide shortage	25.04**
	(10.83)
Other staff shortage	42.09
	(28.84)
Has week supply of N95 masks	-0.692
	(2.423)
Time invariant, exogenous variables	
Medicaid services rev. / total services rev.	-40.83
	(237.5)
% county population that is elderly, 2019	3.400
	(9.126)
Total population of the county, 2019	1.33e-06
	(1.02e-05)
% county population African American	-1.514
	(5.004)
County median household income, 2018	-6.33e-05
	(0.000561)
Time invariant, endogenous	
NH is for-profit	-308.2
	(1,171)
Total patient census days, 2019	-0.00143
	(0.00606)
Constant	289.4
	(1,064)
Observations	11,088
Number of nursing homes	807

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Per the hypotheses discussed in the Background section, some results in the table above are expected. New COVID-19 cases in the county, RN shortages and nursing aide shortages are predicably correlated with worse COVID-19 outbreaks. Similar to the RE regressions, the proportion of Medicaid services revenue has an insignificant correlation with the severity of outbreaks, but median household income is not significant in the estimation above. This could perhaps indicate that patients' ability to afford quality care has little impact on whether or not they receive quality care, proxied by COVID-19 cases in their nursing home. The hypothesis that Medicaid reimbursement might be sufficient to cover the costs of quality care is not rejected by the results above. Clinical staff shortages are significantly negatively correlated with outbreaks, which likely is caused by the same effect in the RE regression: the dataset contains only a small sample of nursing homes that ever experience a shortage in clinical staff. Also, market structure proxied by market concentration appears to have little association with quality-related outcomes.

5.3 Robustness tests

The presumption of a randomly distributed intercept in the random effects regression possibly biases the estimation results. Given implications of the discussion above, it is worthwhile to document the extent of the bias with a comparison between a fixed effects and random effects estimation. Although a fixed effects model cannot be constructed with the same variables used in the RE regression above, two side-by-side FE and RE regressions can be assessed using the pertinent time-variant variables in the main models used. The table below shows the results of RE and FE estimations using only the time-invariant variables, still predicting the new confirmed COVID-19 cases in the nursing home per 1,000 residents. While it does control for market structure using the weekly HHI, it does not control for important factors such as the organization type of the nursing home and the demographics of the county.

Table 7: FE vs. RE: Estimation of New NH COVID-Cases, Time-Variant Variables

	Fixed Effects	**Random Effects**
Weekly 2020 HHI, 5/31/20-11/22/20	-39.90	4.733
	(46.57)	(4.075)
New confirmed COVID-19 cases in county	0.000363***	7.24e-05
	(9.75e-05)	(7.60e-05)
Testing ability of the NH[41]	9.919	1.914
	(13.22)	(4.300)
Shortage of RNs.	28.95*	21.58*
	(14.81)	(11.95)
Week supply of N95	0.0287	0.0478
	(2.991)	(2.131)
Total occupied beds in NH	-0.338***	-0.0163
	(0.125)	(0.0101)
Shortage of clinical staff	-84.79***	-39.60**
	(29.56)	(20.16)
Shortage of nursing aides	27.76***	7.606
	(10.63)	(6.069)
Shortage of other staff	34.71	19.71
	(25.31)	(21.31)
Week supply of surgical masks	3.563	2.313
	(2.422)	(2.051)
Week supply of hand sanitizer	-5.576	-3.344
	(3.679)	(2.327)
Constant	24.51	5.716
	(17.81)	(5.034)
Observations	12,554	12,554
Number of Nursing Homes	913	913
Adjusted R-squared	0.014	

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Review of the table above reveals that there are dramatic differences between many of the coefficients, suggesting a strong bias introduced due to the random effects estimation. This is especially the case for the weekly HHI, the independent variable of interest, and the dummies indicating a shortage in staff, which are critically important control variables. A Hausman test can be used to assess the difference between the coefficients of a random effects estimation and a fixed effects estimation, with the null hypothesis being that the difference between the two is zero.

[41] A dummy indicating whether the nursing home has the ability to test all of its residents for COVID-19 within the week or has the ability to acquire the resources to do so.

The results shown in the table above use robust standard errors. Running the same regression without heteroskedasticity-adjusted standard errors[42] allows a Hausman test to be conducted to evaluate the level of bias introduced by RE rather than FE. The result of the test is a p-value of nearly zero. This indicates that the error term is endogenous to the control variables, so nursing home effects are not randomly distributed (Torres-Reyna, 2007). The level of bias introduced by the use of random effects is significant when only time-variant variables are included in the model. Although the regressions above omit important controls, the result of their comparison still suggests that, in the main RE regression using both time-variant and time-invariant variables, a good deal of bias may be present due to unaccounted nursing home effects. This may be particularly important when assessing the impact of competition on COVID-19 outcomes in nursing homes, since the sign on weekly HHI is opposite between the random effects and fixed effects regressions above due to the bias introduced by random effects.

As previously mentioned, the HT instruments and their appropriateness are dependent upon the exogenous variables' correlation with the endogenous, time-invariant variables. The table below shows the correlation matrix for the time-variant, exogenous variables and the time-invariant endogenous variables:

[42] The results of this regression are in the appendix A.3

Table 8: Correlation Matrix for HT

	For-profit dummy	Patient census days 2019
For-profit dummy	1	
Patient census days 2019	0.1403	1
New confirmed COVID-19 cases in the county	0.0395	0.0913
Testing ability of the NH	0.0339	-0.0355
Proportion of services revenue from Medicaid	-0.0184	-0.0306
% of population elderly in county, 2019	0.0591	-0.0561
Total population of county, 2019	0.0515	0.1227
% of population African American, 2019	-0.0214	-0.0073
2018 Median household income in the county	0.0108	-0.0020

The proportion of African Americans in the county in 2019 (Apr.-Jul.) and the 2018 median household income in the county appear to be weak instruments for the patient census days in 2019, but both appear to correlate with the for-profit dummy enough that they could be a basis for an HT instrument. The rest of the exogenous variables appear to sufficiently correlate with both of the time-invariant, endogenous variables to justify their use in the HT.

Heteroskedasticity, or nonconstant variance in the error term, is possibly a concern for the models used here due to the likely omitted variable bias due to omitted nursing home effects. To test for heteroskedasticity, a Wald statistic was computed based on a generalized least squares estimation. This estimation used used the same variable specification used in the fourth column of the random effects estimation table, with weekly HHI as the independent variable to capture market structure.[43] The chi squared output of this test was 3.7e+10, giving a p-value of nearly zero. This model assumes homoskedasticity, so a p-value of 0 rejects null hypothesis of homoskedasticity, implying evidence for heteroskedasticity. It is likely the case that the models used have biased standard errors, harming the hypothesis testing implicit in this paper's discussion (Studenmund, 2017).

[43] See Appendix A.4 for the coefficients and standard errors of this GLS estimation.

Multicollinearity, or strong linear correlation between one or more variables, could cause upwardly biased standard errors and likewise adversely affect hypothesis testing (Studenmund, 2017). Evaluation of pairwise correlations of the main specification variables[44] indicates no extreme correlations between variables, with the highest correlation being 0.658 between the dummy variables indicating a shortage of RNs and a shortage of nursing aides. This is logical considering that, if a nursing home is experiencing staff shortages of people who care directly for patients, both of those staff types would be encompassed in that category. As an additional for multicollinearity, variance inflation factors were computed for each variable in the weekly HHI random effect specification based on an OLS estimation. Each VIF represents the impact of multicollinearity on increasing the variance of the respective coefficient (Studenmund, 2017).[45]

Table 9: VIFs for OLS estimation of weekly new COVID-19 cases in nursing homes

Variable	**VIF**
Population of the county	3.30
New confirmed COVID-19 cases in the county	2.83
Shortage of RNs	1.97
Shortage of nursing aides	1.79
Weekly HHI	1.74
Shortage of clinical staff	1.54
Proportion of the county population that are elderly	1.52
Proportion of the county population that are African American	1.46
Shortage of other staff	1.36
Median household income in the county, 2018	1.29
Nursing home is for-profit	1.09
Total patient census days 2019	1.05
Week’s supply of N95 masks	1.02
Proportion of services revenue from Medi-Cal (Medicaid)	1.01
Nursing home has materials COVID-19 test or get materials to test all residents within week	1.01

All VIFs are relatively small, or less than 5, indicating that multicollinearity is likely not responsible for bias in the model’s standard errors (Studenmund, 2017).

[44] See Appendix A.4 for the complete table showing the pairwise correlations.
[45] See Appendix A.4 for the coefficients and standard errors of this OLS estimation.

Conclusion

The empirical results above suggest that market structure has little to do with COVID-19 outcomes in California nursing homes, although they are given by imperfect estimations.[46] Still, there are insights available from this analysis. Note that the estimations only agree that new COVID-19 cases in the county are significantly correlated with outbreaks in nursing homes. Community spread of the virus could be sufficiently able to explain most of the variance among nursing homes' outbreak outcomes. It could be the case that the COVID-19 pandemic is largely exogenous to nursing homes' quality outcomes, despite the intrinsic relationship between infection control and quality of care. According to the estimates above, not even nursing homes' labor-related decision making has a strong correlation with new COVID-19 outbreaks.

COVID-19 in nursing homes, as a disease, human catastrophe, and administrative nightmare, could be too novel a phenomenon to meaningfully proxy for quality of care, and nursing home behavior could be abnormal as a result. Furthermore, some literature may contextualize this finding. Despite the connection between input-based quality measures and quality-of-care outcomes in nursing homes, output-based measures of quality do not universally align with favorable outbreak outcomes in nursing homes. For outcome-based quality measures, CMS' five-star ratings on Nursing Home Compare are important indicators. However, in terms of predicting severity of COVID-19 outbreaks, these five-star ratings seem to be less reliable. Bui (2020) finds that nursing homes with one-star ratings

[46] If there truly is a relationship between market structure and care quality, California might not be the best place to investigate; per the summary statistics, the vast majority of California nursing homes are contained by competitive markets. Also, the assumption that counties constitute separate markets may be flawed in a state where individual counties are quite large compared to other states in terms of population and square milage.

were more likely to face a COVID-19 outbreak. He et al. (2020) find that nursing homes with four- and five-star ratings were less likely to have high rates of COVID-19 related deaths, but one-star nursing homes were less likely than two- and three-star facilities. Harrington et al.'s (2020) study indicates that nursing homes with higher star-ratings, specifically under CMS' RN staffing-related categories, were more likely to have lower proportions of COVID-19 cases in their facilities. Abrams et al. (2020) and White et al. (2020) find no significant correlation between five-star rating and the likelihood of COVID-19 cases being reported in a facility. Therefore, COVID-19 outbreaks in nursing homes could be an ineffective predictor for care quality deficiencies. Despite infection control's intrinsic association with quality of care, it could have little association with the typical quality level of a nursing home's care. In this case, variation in market structure type could have little relevance in predicting COVID-19 outbreak outcomes.

It may also be the case that market structure itself has little to do with quality of care. Holding other quality-associated firm decisions equal, monopolies, oligopolies, or perfectly competitive markets may be equally capable of providing quality. Variations in quality available could be attributable to demand-side factors, such as income or, in the case of COVID-19, community spread, only.

The other notable result from this study's estimations is the significance of the for-profit dummy in the set of random effects regressions, which suggest a strong positive correlation between for-profit facilities and COVID-19 outbreaks. This is a concerning finding in light of the prevalence of for-profits in the industry. Attention should be given to decoupling this relationship. If similar effects explain both the relationship between for-profits and COVID-19 outbreaks and the literature's explanation of quality disparity between for-profits and nonprofits, then policy responses that have potential benefits to information transparency could potentially help patients during the pandemic too. The information asymmetry hypothesis regarding the prevalence of for-profits would contend that efforts to promote information transparency should promote better quality of nursing home care. Lin (2015) notes that their predictions of nonprofits' and for-profits' decision-making demonstrate fairly similar behavior in terms of entry, exit, and quality of care. They write that this similarity could be due to outside factors aligning the incentives of nursing homes of different organization types, such as the push for information transparency in the industry. Resources for consumers that disseminate information about COVID-19 in nursing homes, such as the reporting requirements required by CMS, could be beneficial in this respect.

If a nursing home's COVID-19 response is not correlated with the quality of care they offer under "normal" circumstances, then it could be the case that COVID-19 outcomes in nursing homes depend on individual effects, after controlling for income, ownership type, and community spread. There are attributes that clearly interact with quality-determining factors enough to be endogenous but that cannot be observed. These attributes include the foresight of administrators in preparing for the novel virus, the compliance of patients and their families with quarantine protocols, the local political climate's response to the virus, and the community's understanding of how infectious disease spreads. Expanding the data collection on the individual decision-making of nursing home administrators could provide an avenue to test these aforementioned factors.

What are the policy implications of these findings for nursing home patients right now? The majority of factors to explain variation in nursing home care mentioned in this paper hinge on consumer choice. The hypothesis concerning oligopolies simply refers to the number of options available to consumers. The information asymmetry and organization structure hypotheses center on what information consumers have available and are able to interpret. However, in terms of policy implications, neither of these theories are useful for current patients in nursing homes, whose nursing home choice has already been decided by their budget, physician, or family, if not themselves. Therefore, policy implications for the current pandemic comprise of damage control based on known factors that ward facilities against severe outbreaks, for example, the emergency movement of funds, personnel, and an infusion of resources toward facilities in low-income communities.

Given this reality, it is essential that as many nursing home residents and nursing home staff get vaccinated for COVID-19 as possible. There are clearly stumbling blocks on that front. A recent CDC report indicated that, when given the chance, about 78% of nursing home residents and only about 38% of nursing home staff members received at least one COVID-19 vaccine dose (Gharpure et al., 2020). The latter percentage is particularly devastating, considering the inadvertent role nursing home staff likely play in the spread of the virus throughout a facility. The virus usually presents asymptomatically at first, meaning infected staff can spread the disease unaware, and lack of ubiquitous paid sick leave available to industry workers could prompt caregivers to show up to treat residents despite being ill with the virus, increasing transmission (Barnett and Grabowski, 2020; McMichael et al., 2020; Ouslander and Grabowski, 2020). The CDC report implies that the mechanism explain low vaccination rates among nursing home staff could be a lack of trust in the vaccine's safety. 37% of nurses surveyed "were not confident that the COVID-19 vaccine would be safe and effective," and only "34% agreed that they would voluntarily receive a COVID-19 vaccine" (Gharpure et al., 2020). Clearly, a needed policy response to address COVID-19 in nursing homes is to confront the information and trust gap between the medical institutions that created the vaccine and the nurses and staff who actually deliver care to the country's most vulnerable patients.

Appendix

This section includes explanations and estimation results referenced, but not included, in the main body of the paper.

A.1 Detailed variable explanations

The following table is a color-coded list of the variables included in the data file used by this paper in the regressions detailed in this section. The color code is to identify the different types of time-variant and time-invariant variables:

Variables that have the same value for all observations in the same county are in green.
Variables that are the same for all time periods but are wholly unique to each nursing home for all time periods are in gold.
Variables that are gathered on the weekly basis, and are unique to each nursing home, are in purple.

<u>Variable name</u>	<u>The Meaning of the Data Included</u>
From CHHS Long-Term Care Financial Data, or calculated with this data in post	
NHnum	A number assigned by the author to identify each nursing home.
COUNTY	The county (in California) in which the nursing home resides
cnty_num	the number assigned by the author to each county that has nursing homes in the set
ForProfit_dum	a dummy to indicate whether the nursing home is a for-profit firm
DAY_TOTL	total patient census days in the year 2019
DAY_MCAL	Gross Routine Services Revenue from Medi-Cal in 2019
GR_RT_TOTL	total Gross Routine Services Revenue in 2019
GR_AN_TOTL	total Gross Ancillary Services Revenue in 2019
GR_RT_MCAL	Gross Routine Services Revenue from Medi-Cal in 2019
GR_AN_MCAL_IP	Gross Ancillary Services Revenue from Medi-Cal for Inpatient services in 2019
GR_AN_MCAL_OP	Gross Ancillary Services Revenue from Medi-Cal for Outpatient services in 2019
medi_services_revenue	total services revenue from Medi-Cal (Medicaid) in 2019
tot_services_revenue	total services revenue in 2019
Medical_pro_services_revenue	proportion of services revenue in 2019 that was from Medi-Cal
competition	number of other nursing homes (also included in the dataset) in the same county as the nursing home

HHI	2019 Hirschman-Herfindahl index of the nursing homes' market, assuming that the county denotes the market boundary. HH index = adding the squared proportion of each nursing home's portion of the total patient census days in the market.
PRDHR_GNP_ppd	Productive Hours of Geriatric Nurse Practitioners in 2019, divided by total patient census days in the year 2019.
PRDHR_RN_ppd	Productive Hours of Registered Nurses in 2019, divided by total patient census days in the year 2019.
PRDHR_LVN_ppd	Productive Hours Licensed of Vocational Nurses, divided by total patient census days in the year 2019.
From Census Bureau's County Population by Characteristics data, Vintage 2019, 1 April 2019 to 1 July 2019, or calculated with data in post	
cnty_elderlypercent	the percentage of the population in the nursing home's county that are elderly April-July 2019
cnty_totpop	the total population in the nursing home's county that are elderly April-July 2019
cnty_popAAper	the percentage of the population in the nursing home's county that are African American April-July 2019
From Small Area Income and Poverty Estimates (SAIPE) State and County Estimates for 2018, the California file, or calculated with data in post	
cnty_PovertyPercent2018	the percentage of the population in the nursing home's county that are below the poverty line in 2018
cnty_MedianHouseholdIncome2018	the median household income of the county in 2018
From State of California Data Portal Case Data for California COVID-19 Response, or calculated with data in post	
cnty_newcountconfirmed	new COVID-19 cases in the county in that week.
endwk	the date at the end of the week in which the data was collected.
From CMS COVID-19 Nursing Home Dataset, or calculated with this data in post	
totoccup_beds	the nursing home's total number of beds occupied that week
testable_dum	a dummy indicating whether the nursing home had the ability to test all residents or get the resources to test all residents that week
RN_short_dum	a dummy indicating whether the nursing home reported a shortage in nursing staff (RNs, LVNs, or LPNs) that week
clinical_short_dum	a dummy indicating whether the nursing home reported a shortage in clinical staff that week
aides_short_dum	a dummy indicating whether the nursing home report a shortage in aides (certified nursing assistant, nursing aides, medication aide, medication technician) that week
otherstaff_short_dum	a dummy indicating whether the nursing home reported a shortage in other staff that week
n95_wksupp_dum	a dummy indicating whether the nursing home reported having at least a week's supply of N95 masks that week
surgicalmasks_wksupp_dum	a dummy indicating whether the nursing home reported having at least a week's supply of surgical masks that week

sanitizer_wksupp_dum	a dummy indicating whether the nursing home reported having at least a week's supply of hand sanitizer that week
wklynewcases_1000res	new COVID-19 cases in the NH in that week per 1000 residents.
endwk	the date at the end of the week in which the data was collected.

A.2 Heterogeneity in Nursing Home Care

There is much heterogeneity in the services sold in the long-term care market, and much of the literature analyzing the industry treats the market as such, especially literature analyzing quality of care (Nyman, 1989; Bishop, 1998; Chou, 2002; Grabowski & Castle 2004). This study, from a theoretical and empirical standpoint, uses a broad definition for "nursing home" as facilities providing skilled care beyond helping residents/patients with day-to-day activities, including having nurses on hand 24/7 to ensure medical care for physical and mental disability (*Residential Facilities, Assisted Living, and Nursing Homes,* 2020). The State of California's Health and Human Services Agency identifies three subtypes of long-term care that strictly abide by this definition, identified via the facility's type of license. Individual facilities offer different menus of services depending on their own capacity to provide them. For SNFs (skilled nursing facilities) alone, the California Department of Aging lists the following as "general services" provided by those facilities:

> Dietary services, social services, pharmaceutical services, recreational therapy services, access to dental care, emphasis on rehabilitation, such as gait training and bowel and bladder training, specialized units for dementia residents, administration of potent injectable medications and intravenous medications and solutions on a regular basis, … [and] may also provide ancillary services such as, physical, occupational, and speech therapies. (*Skilled Nursing Facilities—Care Options,* 2020)

The following definitions are from the California Health and Human Services Lookup Table for Certified Healthcare Facility Listing (CHHS, 20 February 2020). The author deems that CHHS' following definitions for specific types of long-term care facilities fall under the aforementioned definition of nursing home care and therefore ought to be considered "nursing homes" for empirical purposes. The following definition of long-term care facility licenses are directly quoted from the Lookup table, but re-formatted for ease of reading here, with words and phrases indicating alignment with this study's definition of nursing home bolded by the author:

> **Congregate Living Health Facility (Value: CLHF):** A residential home with a capacity, of no more than 18 beds (pursuant to Health and Safety Code section 1250(i)(4)(A) a city or county operated facility delivering the same congregate living health facility services may have a capacity of 59 beds), that **provides inpatient care**, including the following basic services**: medical supervision, 24-hour skilled nursing and supportive care**, pharmacy, dietary, social, recreational, and at least one type of the following services: **services for persons who are mentally alert, persons with physical disabilities, who may be ventilator dependent**; services for persons who have a diagnosis of terminal illness, a diagnosis of a life-threatening illness, or both**; services for persons who are catastrophically and severely disabled**. The primary need of congregate living health facility residents shall be for availability **of skilled nursing care on a recurring, intermittent, extended, or continuous basis**. This care is generally less intense than that provided in general acute care hospitals

but more intense than that provided in skilled nursing facilities. (Ref: Health and Safety Code section 1250(i)(1)).

…

Skilled Nursing Facility (Value: SNF): A health facility that provides skilled nursing care and supportive care to patients whose primary need is the availability of skilled nursing care on an extended basis (Ref: Health and Safety Code section 1250(c).)

…

SNF/RES (Value: SNF/RES): A SNF/RES facility is licensed for skilled nursing or intermediate care, but it is an integral part of a residential care facility (Mocilla, 2009 pp. 109).

Likewise, the documentation for the CHHS dataset records:

SNF/RES and ICF/RES indicate facilities that are licensed for skilled nursing or intermediate care but are an integral part of a residential care facility.

A.3 Background regression results omitted for brevity

The tables below include regression results that were used in the estimations discussed in this paper but were not included in the body of the text for the sake of brevity.

Table 10: Random effects Estimation of COVID-19 outbreaks in nursing homes, background controls

	Main IV: # of rivals	Controls with 2019 Labor Hours	**Main IV:** 2019 HHI	**Main IV:** Weekly HHI
Total patient census days in 2019	1.91e-05	2.00e-05	2.25e-05	1.97e-05
	(2.63e-05)	(2.64e-05)	(2.58e-05)	(2.64e-05)
Medicaid services rev. / total services rev.	5.857	5.872	5.505	5.768
	(10.98)	(10.90)	(12.48)	(10.90)
County elderly pop. / total pop.	0.0228	0.0858	0.0673	0.101
	(0.255)	(0.251)	(0.253)	(0.251)
Testing ability of the NH	1.974	1.986	1.660	1.988
	(4.753)	(4.767)	(4.652)	(4.772)
County African American pop. / total pop.	-0.0658	-0.105	-0.0454	-0.114
	(0.176)	(0.162)	(0.173)	(0.163)
Shortage of other staff	21.20	21.23		21.24
	(22.96)	(22.95)		(22.95)
NH has week supply of N95	-0.666	-0.617	-0.917	-0.615
	(1.753)	(1.757)	(1.718)	(1.753)

Table 11: FE vs. RE: Estimation of New COVID-Cases, Time-Variant Variables, Non-robust SEs

	Fixed Effects	Random Effects
Weekly 2020 HHI, 5/31-11/22	-39.90	4.733
	(51.00)	(3.560)
New confirmed COVID-19 cases in county	0.000363***	7.24e-05
	(0.000101)	(7.58e-05)
Testing ability of the NH[47]	9.919*	1.914
	(5.699)	(3.713)
Shortage of RNs.	28.95***	21.58***
	(6.436)	(5.150)
Week supply of N95	0.0287	0.0478
	(2.395)	(1.943)
Total occupied beds in NH	-0.338***	-0.0163
	(0.0600)	(0.0109)
Shortage of clinical staff	-84.79***	-39.60***
	(12.04)	(8.827)
Shortage of nursing aides	27.76***	7.606**
	(5.172)	(3.647)
Shortage of other staff	34.71***	19.71***
	(10.01)	(7.532)
Week supply of surgical masks	3.563	2.313
	(3.297)	(2.716)
Week supply of hand sanitizer	-5.576*	-3.344
	(3.219)	(2.457)
Constant	24.51***	5.716
	(8.346)	(4.294)
Observations	12,554	12,554
Number of Nursing Homes	913	913
Adjusted R-squared	-0.064	

A.4 Testing for Heteroskedasticity and Multicollinearity

The tables below show the standard errors, coefficients, and correlations utilized when testing for heteroskedasticity and multicollinearity.

[47] A dummy indicating whether the nursing home has the ability to test all of its residents for COVID-19 within the week or has the ability to acquire the resources to do so.

Table 12: Generalized Least Squares Estimation, Weekly HHI

NH is for-profit	4.274***
	(1.217)
Total Pop. county 2019 Apr.-Jul.	-8.28e-07***
	(1.51e-07)
Median household income in county, 2018:	-9.23e-05***
	(2.08e-05)
New confirmed COVID-19 cases in county	0.000282**
	(0.000111)
Shortage of RNs	10.81**
	(4.757)
Shortage of clinical staff	-21.04***
	(7.848)
Shortage of Aides	4.117
	(3.224)
Weekly 2020 HHI, 5/31-11/22	-5.826*
	(3.540)
Total patient census days in 2019	2.10e-05
	(2.23e-05)
Medicaid services rev. / total services rev.	5.607
	(8.928)
County elderly pop. / total pop.	0.0986
	(0.171)
Testing ability of the NH	-0.415
	(3.289)
County African American pop. / total pop.	-0.114
	(0.118)
Shortage of other staff	15.00**
	(6.953)
NH has week supply of N95	-0.647
	(1.400)
Constant	12.16**
	(4.856)
Observations	11,088
Number of nursing homes	807

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 13: Ordinary Least Squares Estimation, Weekly HHI

NH is for-profit	4.274***
	(1.218)
Total Pop. county 2019 Apr.-Jul.	-8.28e-07***
	(1.51e-07)
Median household income in county, 2018:	-9.23e-05***
	(2.08e-05)
New confirmed COVID-19 cases in county	0.000282**
	(0.000111)
Shortage of RNs	10.81**
	(4.761)
Shortage of clinical staff	-21.04***
	(7.854)
Shortage of Aides	4.117
	(3.226)
Weekly 2020 HHI, 5/31-11/22	-5.826
	(3.543)
Total patient census days in 2019	2.10e-05
	(2.23e-05)
Medicaid services rev. / total services rev.	5.607
	(8.935)
County elderly pop. / total pop.	0.0986
	(0.171)
Testing ability of the NH	-0.415
	(3.291)
County African American pop. / total pop.	-0.114
	(0.118)
Shortage of other staff	15.00**
	(6.958)
NH has week supply of N95	-0.647
	(1.401)
Constant	12.16**
	(4.860)
Number of Observations	11,088
R-Squared	0.008

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 14: Pairwise Correlations

Variables	Weekly new NH COVID-19 cases	Weekly HHI	For-profit	Total patient census days 2019	Prop. rev. Medicaid	% elderly pop. in cnty	Total cnty pop	% AA pop. in cnty	Median HH income in cnty	New COVID cases in the county	NH COVID-19 testing ability	RN short	Clinical staff short	Nursing aides short	Other staff short	Week supply N95
Weekly new NH COVID-19 cases	1.000															
Weekly HHI	-0.001	1.000														
For-profit	0.035	0.001	1.000													
Total patient census days 2019	0.024	-0.091	0.199	1.000												
Prop. services revenue from Medicaid	0.009	-0.057	-0.016	-0.032	1.000											
% elderly pop. in cnty	-0.019	0.482	0.042	-0.067	-0.081	1.000										
Total county pop	0.007	-0.433	0.075	0.138	0.011	-0.296	1.000									
% AA pop. in cnty	0.005	-0.378	0.006	0.017	0.053	-0.443	0.417	1.000								
Median HH income in cnty	-0.047	-0.154	-0.049	-0.054	-0.023	0.163	-0.279	-0.149	1.000							

Variables	Weekly new NH COVID-19 cases	Weekly HHI	For-profit	Total patient census days 2019	Prop. services revenue from Medicaid	% elderly pop. in cnty	Total county pop	% AA pop. in cnty	Median HH income in cnty	New COVID cases in the county	NH COVID-19 testing ability	RN short	Clinical staff short	Nursing aides short	Other staff short	Week's supply N95
New COVID cases in the county	0.034	-0.337	0.065	0.112	0.009	-0.247	0.843	0.362	-0.290	1.000						
NH COVID-19 testing ability	-0.001	-0.004	0.016	-0.028	-0.011	-0.036	0.021	0.021	-0.034	0.018	1.000					
RN shortage	0.101	-0.015	0.019	0.037	-0.006	-0.020	0.000	0.020	-0.021	0.002	-0.029	1.000				
Clinical staff shortage	0.015	-0.006	0.012	-0.002	-0.009	-0.013	-0.005	0.017	-0.022	-0.003	-0.025	0.481	1.000			
Nursing aides shortage	0.085	-0.010	-0.075	0.018	-0.019	0.003	0.014	0.037	-0.020	0.010	-0.008	0.658	0.370	1.000		
Other staff shortage	0.068	-0.006	0.007	0.010	-0.002	-0.007	0.025	0.015	-0.014	0.019	-0.022	0.434	0.491	0.402	1.000	
Week's supply N95	-0.027	0.034	0.020	0.014	0.022	0.022	-0.002	-0.051	0.004	-0.015	0.029	-0.027	-0.010	-0.049	-0.018	1.000

A.5 Banker et al.'s 1998 article, "Quality and Competition"

Banker et al.'s 1998 paper, "Quality and Competition," looks to address the relationship between competitive intensity and the quality level of goods with a theoretical model based on price competition between oligopolist firms. They note that the relationship between these two variables would differ depending on what is meant by "competitive intensity," pointing to different scenarios presented in the airline, auto manufacturing, and computer manufacturing industries. The third of these scenarios looks to explain possible impacts on quality by competitive intensity, where competitive intensity is defined as an increase of the number of firms in the market, which is the independent variable of interest to this study. Therefore, this analysis will focus on their theoretical analysis in this case. Banker et al.'s model is extremely complex, so the following summary is this intended to focus on the parts of their analysis that lend to this study's quest to investigate the relationship between oligopolist market structures and quality of output, specifically, sparing the reader an in-depth mathematical analysis.

Banker et al. start by defining quality as the "conformance and design" of a good that "are of interest to the consumer" (1180). Quality is identified as the attributes of a good that make it more desirable to consumers than a competitor's good, where "conformance" is the good's ability to meet performance expectations, and "design" refers to the features of performance "such as reliability, durability, and serviceability" (1180). It directly relates to market price, or consumers' willingness to pay, for a firm's good. An increase in a firm's quality level causes two functional shifts; consumers have more willingness to pay for the firm's goods, meaning an increase in market demand for that good, and consumers are less willing to pay for competitors' goods, meaning a decrease in market demand for those goods.

Also, rather than invest in real quality increases, a firm may choose to invest in advertising to only improve the "perceived" quality, versus the real quality, of a good. Since this strategy affects consumers' expectations of a good and its desirability, it has the same impact of a real quality increase on the price.

To discern the impact of adding additional firms to the market on firm i's quality level, Banker et al. solve for a first-order condition of profit maximization in respect to a change in quality. In other words, they solve for the firm's quality level by finding the optimal profit function's derivative in respect to quality and setting it equal to zero. The following is an overview of their process.

There are n firms in the market. The demand, or inverse price function, for firm i is a linear function of price p_i, quality level x_i, the sum of competitors' prices, $\sum_{j \neq i} p_j$, and the sum of competitors' quality, $\sum_{j \neq i} x_j$, where firm i's choice variables are quality level and price. It is given by:

$$q_i = \alpha_n - \beta_n p_i + \gamma_n \sum_{j \neq i} p_j + \lambda_n x_i - \mu_n \sum_{j \neq i} x_j \tag{1}$$

Note the subscript n on Equation (1)'s intercept and variable coefficients, α_n, β_n, λ_n, and μ_n. After showing the reader Equation (1) above, Banker et al. sum all individual firms' price functions to derive a function for total market demand. However, just because firms may enter or exit the market does not mean there is more or less market demand for the good as they do so. Banker et al. wish to constrain the total market demand function so total market demand does not change as firms enter or exit. They do this by averaging each firm's "intrinsic demand potential,"[48] or baseline demand for each firm's good, α, and own price and quality coefficients, β_n and λ_n, by the number of firms in the market, n:

$$\alpha_n = \frac{\alpha_i}{n} \tag{2}$$

$$\beta_n = \frac{\beta_i}{n} \tag{3}$$

$$\lambda_n = \frac{\lambda_i}{n} \tag{4}$$

Where α_i [49] is the baseline demand for firm i's good, β_i is the factor by which a $1 increase in firm i's price causes the quantity demanded of firm i's good to fall, and λ_i is the factor by which an increase in the quality level of firm i's good by 1 causes the quantity demanded of the good to increase.[50] Similarly, they average the coefficients for all the sum of the competitors' quality level, μ_j, and the coefficient for the sum of the competitors' price level, γ_j such that:

$$\mu_n = \frac{\mu_j}{n(n-1)} \tag{5}$$

$$\gamma_n = \frac{\gamma_j}{n(n-1)} \tag{6}$$

[48] Banker et al. do not use detailed language in describing the exact interpretation of the demand function's intercept, α. In another of the paper's scenarios using different criteria for competitive intensity, they provide an individual duopolist firm's demand function as $q_i = k_i\alpha - \beta p_i + \gamma p_j + \lambda x_i - \mu x_j$ and state that $k\alpha$ represents the "intrinsic demand potential parameter for firm i" (p. 1182). This review interprets this term as representative of a "baseline" demand for firm i's good. In other words, what quantity consumers would demand of firm i's good if all firms in the market gave away the good with zero quality for free, were it possible. They remove the coefficient k in the scenario of interest to this IS, which analyzes the impact of the increase of the number of firms in the market on the quality firm i provides; therefore, there is no k included in this section's analysis.
[49] Since they do not say specifically, this review presumes that Banket et al. base their average coefficients defined by Equations (2), (3), and (4) on their established definitions for these coefficients in their individual duopolist scenario in which they give the individual firm's demand function as $q_i = k_i\alpha - \beta p_i + \gamma p_j + \lambda x_i - \mu x_j$. This review includes the subscript is and js in Equations (2), (3), (4), (5) and (6) to help establish more clearly for the reader that these are the coefficients and intercepts faced by individual firms in the market.
[50] The positive and negative directions of the relationships follow the signs used in Equation (1) and in Banker et al.'s duopolist firm example where the inverse demand function is $q_i = k_i\alpha - \beta p_i + \gamma p_j + \lambda x_i - \mu x_j$.

Where μ_j is the factor by which an increase in 1 of the sum of all other firms' quality level decreases the price firm i charges for its product; γ_j is the factor by which an increase in the sum of all other firms' prices by $1 causes the price firm i charges to increase.[51] With their demand function for individual firm i's output decided, Banker et al. sum the demand function of all firms:

$$\sum_{i=1}^{n} Q_i^n = n\alpha_n - n(\beta_n - (n-1)\gamma_n)\overline{p} + n(\lambda_n - (n-1)\mu_n)\,\overline{x} \qquad (7)$$

where $\overline{p}$ is the average price charged by firms and $\overline{x}$ is the average quality level. This price function essentially gives an average market price rendered by the levels at which individual firms set their choice variables, price and quality. This price function is what they use when solving for the optimal profit function later.

Firm i's costs, c_i are a function of variable and fixed costs,

$$c_i = fixed\ costs + variable\ costs \qquad (8)$$

where fixed costs are

$$fixed\ costs = \ f + \phi_n x_i^2 \qquad (9)$$

and

$$variable\ costs = (v + \epsilon x_i)q_i \qquad (10)$$

where f represents fixed costs with no quality investment, $\phi_n x_i^2$ represent fixed costs associated with quality investment, where ϕ_n represents the average individual factor of cost increase to firms per an increase in x_i^2.[52] v is the variable cost added per additional unit of

[51] See note above.

[52] Banker et al. provide for the impact of quality level, x_i, to be convex, hence the squared term. This choice does not matter for this analysis, since this study is primarily interested in Banker et al.'s findings regarding the directional relationships of their variables rather than the scale of impact in changing each one.

output q_i, and ϵx_i is the quality-associated increase in variable costs, with ϵ as an exogenous market-wide factor by which an increase in quality level causes an increased in fixed costs per additional unit of q_i. Banker et al. impose several necessary constraints on the bounds of v, ϵ, and ϕ_n, but this analysis will refrain from analyzing that level of detail in their highly complex model. Likewise, from here this analysis describes an overview of their logic to the best of the author's understanding without digging into their intricate extraction of the optimal quality level, x_i^*.

Banker et al. look to solve for the optimal profit function, revenue minus costs, c_i, where revenue is given by q_i times the price solved for via the average market price function, Equation (7). From here, they solve for the optimal π_i^* as a function of quality level, x_i, and optimal price level, p_i^*. They then take the derivative of π_i^* in respect to x_i, or $\frac{\partial \pi_i}{\partial x_i}$, and set it equal to zero to find the quality level, x_i^*, that yields the optimal profit, π_i^*. They then solve for x_i^* as a function of n, the number of firms in the market, where there it is evident that an increase in the number of firms causes the profit-maximizing quality level of the firm to fall as firm i's market share falls with an increase in n.

From here, Banker et al. introduce a new avenue for modeling the theoretical relationship between oligopolist market structures and quality. They write that, in a more realistic example, the entry of additional firms in the market might cause a reduction in a firm's quality-associated costs, holding total costs equal. This would occur because of knowledge spillovers from competitors that may increase efficiencies and because of increased demand in the factor market for fixed inputs. To control this effect and demonstrate its implications, they state that rather than having the average factor of cost increase to firms per an increase in x_i^2, represented by ϕ_n be exogenously fixed, they write it as

$$\phi_n = \frac{\phi_i}{n^{1-\rho}}$$

where $0 < \rho \leq 1$

where ϕ_i is firm i's individual factor of fixed cost increase as a result of a 1 unit increase in x_i^2, n is the number of firms in the market, and ρ is an measure of how strongly an increase in the number of firms in the market might tend to bring down the average factor for quality-associated fixed costs. If $\rho = 1$, there is no tendency for an increase in the number of firms in the market to cause quality-associated fixed costs to fall. If $0 < \rho < 1$, some tendency exists for an increase in the number of firms in the market to bring down quality-associated fixed costs. Banker et al. demonstrate that, depending on what ρ and n's relative magnitudes are, the profit-maximizing quality level, x_i^*, can rise or fall as n grows. This impact may change depending on the magnitudes of n and ρ. In other words, if there exists some tendency for an increase in the number of firms in the market to cause quality-associated fixed costs to fall, that effect must overpower the conflicting effect of each firm's smaller market share bringing down the optimal quality level while maintaining the same quality-associated costs.

A.6 Lin's 2015 article, "Quality Choice and Market Structure"

Lin's 2015 paper, "Quality Choice and Market Structure: A Dynamic Analysis of Nursing Home Oligopolies," addresses three questions regarding market structure and nursing homes: (1) investigating what factors influence how nursing homes select their "product type, or "quality level," (2) whether a forced exit of low-quality nursing homes, perhaps due to fines or forced shutdown by the government, would result in consumer welfare being lost due to surplus demand in the local market or in a price increase that would be prohibitive for consumers; and (3) if there is a relationship between the quality level firms select and the level of competition in the market. While the Background section touches on

Lin's investigation of the first two questions, this section will focus primarily on their findings regarding the third question as they pertain to the topic of this study, if oligopolist market structures in the nursing home industry adversely impact their care quality. First, this review will outline Lin's theoretical and empirical analysis relevant to this study. Then, this review will turn to specifically addressing Lin's experiment that focuses on the third research question above.

To address their questions, Lin constructs a Markov-perfect equilibrium (MPE)[53]-based theoretical model to describe strategic interaction among oligopolist nursing home firms. They select this equilibrium type due to its "intertemporal dependence;" that is, firms' decisions in one period impact their states and possible actions in the next. This is a realistic interpretation of the real-world infinite game nursing homes play, in which actions selected in one year impact the firms' state in the next. Lin's empirical model uses CCP and a two-stage fixed effects method to control for the uniqueness of individual nursing home markets. Lin retrieved their data from the 1996-2005 Online Survey Certification and Reporting Standards System (OSCAR) surveys, which assess nursing homes' compliance to federal regulations concerning outcome-based and input-based quality metrics among other characteristics. Also, Lin uses Census data and Medicaid reimbursement data for identifying and controlling demand-side characteristics of county-level nursing home markets. Lin makes adjustments to the sample to ensure appropriateness of the data, including removal of

[53] A Markov-perfect equilibrium is a type of subgame equilibrium in dynamic games with pure strategies, developed to address dynamic games' "multiplicity of equilibria" (Gitmez, 2017). This analysis does not derive Markov-perfect equilibrium as Lin does, but it does discuss Lin's maximization problem in terms of relevant probabilities. For more information concerning the derivation of Markov-perfect equilibrium, see Lin (2015) and Gitmez (2017).

markets in which more than six nursing home firms operate.[54] After this weeding process, 132,138 observations remain, which encompass individual nursing homes at different time periods from 2,417 counties.

Lin outlines the most critical variables in their model. They use counties as representatives of separate individual markets. Time is measured in years, and the cross sectionals are individual nursing homes. To track when a nursing home firm is a market newcomer or an incumbent, Lin notes the entry/exit of firms based on the date when Medicare/Medicaid participation is allowed. Alternatively, entry occurs at the time when the nursing home is included in the data but not in the years prior. Exit is at the time a nursing home is no longer in the dataset.

Lin uses a discrete system for generating their quality variable. They start by selecting an input-based proxy for quality of care, licensed nursing hours per resident day, which is a continuous metric. For computational convenience later in their model's execution, they discretize the distribution into two categories, "low-quality" nursing homes and "high-quality nursing homes" based on which side of the sample median the observations are on. After this process, Lin identifies 24,413 low-quality nursing homes and 24,733 high-quality nursing homes in the dataset.

While Lin notes that using licensed nursing hours as a proxy for quality is an approach well-tested by the literature, there is a concern associated with input-side quality metrics, which they address. While the literature indicates nursing homes have better quality of care use more skilled labor, and it is logical to presume that is the case, it is also likely true

[54] Since they observe that adding more firms after five and six does not have a strong impact on "competitive impacts" they note later in the paper. Also, thus ensures that the counties used to proxy for individual markets does not actually contain two or more markets instead of one.

that nursing homes of different scale could tend to use more or less skilled labor. Lin partially controls for this by using licensed nursing hours *per patient day*, which standardizes skilled nursing hours over the volume of patients the facility is tending to. However, there could be unaccounted for economies of scale involved. A larger nursing home might have management differences or organizational procedures meaning fewer or more skilled nursing hours are needed. Moreover, nursing homes of different size will likely have different profitability levels or different quality levels. In this case, it is a concern that using licensed nursing hours per patient day might just be proxying for an omitted variable, such as nursing homes' internal efficiencies. Lin addresses this by testing with an "output" based quality metric, the amount of quality citations by government surveyors the nursing home received. This is a robustness test in addition to skilled labor hours used in their main regression. They find that poor performance in this output-based quality measure has a strong correlation with fewer licensed nursing hours per patient.

Lin then constructs their theoretical model, which is an infinite multi-stage dynamic and repeating game. This review provides an overview of this model. Individual firms, i, operate in M number of markets, and are either a potential entrant or an incumbent firm, and operate with either a high- or low-quality level. Each time period, t, is a year. N_M is the number of nursing homes in market M. Each stage of the game has four parts. In the beginning of the time period t, nursing home i notes the outcomes existent in its market M at the end of $t-1$, the previous year, such as the number of nursing homes in the market and those firms' quality levels. Nursing home i also notes whatever private information it has, unknown to the other firms, called ε_{it} later in this analysis. Next, nursing home i picks its two actions during t; that is, whether it will enter or exit the market, and whether it will maintain or change its quality level. Then entry costs and quality-adjustment costs are incurred. Lastly, profits, the period payoff, are collected. Nursing home i is only aware of what happened in the previous time periods, and all decisions are simultaneous for all nursing homes. For this model, Lin fixes the number of firms in each market to be six, or $N_M = 6$ for all M. These six firms include potential entrants.

The market's "observed state" in the model at time t is stored in a vector, x_t, which can be used as an expression of a market's status in terms of firm quality level characteristics. It is representative of the state of the market known to all players. In other words,

$$x_t = (M_t, I_t, s_{1t}, s_{2t} \dots s_{Nt}) \tag{1}$$

where M_t is the market size represented as the size of the elderly population in the area, I_t is income per capita,[55] and $s_{it}, s_{2t} \dots s_{Nt}$ represent the each nursing home's quality status prior to their decision regarding possible quality level adjustment, or exit, in time t. Each firm's action in time t, that is, what level of quality they will transition to in the next period $t+1$, is represented by a_{it}. Note that each time period's action directly leads to a firm's quality state in the next period. Therefore, $a_{it} = s_{i(t+1)}$. Moreover, if a nursing home decides to maintain the same quality level in $t+1$ as in time t, it is true that $s_{it} = a_{it} = s_{i(t+1)}$.

Note that x_t is only the "observed state" of the market, versus the "unobserved state." x_t represents what each nursing home at time t may observe about the market and their

[55] Lin chooses to discretize both income per capita and local elderly population. By divvying the distribution of elderly populations into ten groups, where each group encompasses 10% of the total number of markets, Lin designates ten sections each assigned a categorical ordinal label 1 through 10. Lin uses the same process to divide up the sample markets' incomes per capita into four groups rather than ten.

competitors, but their perception may not accurately represent the true state of the market due to the fact that individual firms have their own "private information" about the market. In a real-world sense, this is the proprietary information that individual nursing home firms may have about themselves or the market in which they operate. Lin holds that each firm's individual "unobserved state," at time t, called ε_{it}, is only known to itself. Each nursing home's unobserved states can be incorporated into a vector to represent a market-wide unobserved state at time t, such as

$$\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots \varepsilon_{Nt}) \tag{2}$$

Given x_t and ε_t, the two dimensions through which the state of the market is decided, Lin creates a vector to represent the overall state of the market, using a capital $S_t = (x_t, \varepsilon_t)$. The complete action profile of an individual nursing home, meaning the history of all of nursing home i's actions up to time t, is A_{it}.

As indicated, profit payoffs in time t are determined by the nursing homes' decision-making, which is a function of the market state parameters:

$$\pi_{it}(x_t, a_t, \varepsilon_t) \tag{3}$$

Profit can be thought of as a function of both known and unknown market states, as well as the firms' actions that result from those states. The decision of primary relevance for payoff/profit maximization in Lin's model is whether or not a nursing home chooses to maintain or adjust the quality level of its care. Lin expresses this decision as a probability function describing the likelihood of S_{t+1}, the state of the market resulting from the transition decisions of nursing homes in the previous period t, contingent on the state of the market at that time and the actions firms take, S_t and a_t. In other words,

$$probability\ of\ quality\ adjustment = F(S_{t+1} | S_t, a_t) \tag{4}$$

If $F(S_{t+1}|S_t, a_t)$ = 1, the state described within the vector S_{t+1} occurs.

Lin notes that, at face value, there is endogeneity between known states x_t and ε_t in this model. This is problematic since the unobserved state parameter, ε_t, serves as an error term in the model, meaning that it seeks to account for variation in the market outcomes S_t that are not attributable to known market states, x_t. Unobserved states must not be correlated to observed states, x_t, lest unobserved states lose explanatory significance in predicting the probability of S_{t+1}. To address this in the profit function, Lin assumes additive separability of the known and unknown states' parameters such that

$$\pi_{it}(x_t, a_t, \varepsilon_t) = \pi_{it}(x_t, a_t) + \pi_{it}(\varepsilon_{it}) \quad (5)$$

Nursing home i's profit at time t can be expressed as a function of x_t, a_t, ε_t which is equivalent to the sum of two profit functions, one contingent on only x_t, a_t, and the other on ε_{it}. In other words, ε_{it} may cause variation in profit that is independent of both the known states and actions' impacts. Lin also assumes conditional independence between known and unknown state variables, insisting that the likelihood of a particular market state, S_{t+1}, occurring is contingent of the probability of both x_t and ε_t, but the probability of x_t is not influenced by the probability of ε_t. Also, Lin holds that ε_t is exogenous and identically distributed for all nursing homes at time t. With these assumptions, Lin expands Equation (4) to the following:

$$probability\ of\ quality\ adjustment = F_x(x_{t+1}\,|x_t, x_a) \cdot F_\varepsilon(\varepsilon_{t+1}) \quad (6)$$

Where $F_x(x_{t+1}\,|x_t, x_a)$, the probability of known state x_{t+1}, the nursing home's observed state in the next period, is given by x_t and a_t, the nursing home's current state and actions. $F_\varepsilon(\varepsilon_{t+1})$ is the nursing home's unknown state factors, which give the nursing home's unknown state in period $t + 1$.

Nursing home's strategy at time t, called σ_{it}, is given by both its unobserved and observed states, x_{it} and ε_{it}, since its status in the market prompts its decision-making in search of period payoffs, their period profit. As a result, its strategy informs its action profile at time t,

$$\sigma_{it}: (x_{it}, \varepsilon_{it}) \rightarrow A_{it} \tag{7}$$

In other words, as each nursing home in each period makes decisions based on its observations of the market, they define their strategies, and those strategies become their individual action histories.

Next, Lin selects the subgame equilibrium payoffs that a nursing home would seek to maximize. Using Equation (6), they formulate firm i's "value" probability function, the probability of the nursing home having a known state x_t given its strategy:

$$V_i(x_t|\sigma) = E\{\pi_i(\sigma, x_t, \varepsilon_{it}) + B\} \tag{8}$$

V_i is the nursing home i's value, which is the expected profit of nursing home i plus another parameter, B.[56] Nursing home i's expected profit is the primary determinant of the value function. The nursing home seeks to maximize its value probability via its decision-making. Lin defines their subgame equilibrium as the value probability that is equal to or greater than all others, given the firm's known state and strategy. In other words, the maximization constraint that yields an optimal subgame outcome for nursing home i is

$$V_i(x_t|\sigma_{it}^*, \sigma_{-it}) \geq V_i(x_t|\sigma_{it}', \sigma_{-it}) \tag{9}$$

[56] Lin derives the full function of $V_i(x_t|\sigma)$ in their model, where $B = E\{\int V_i(x_{t+1}|\sigma)dF(x_{t+1}|x_t, \sigma)\}$. This is the expected value of the integration of the value function of the next period given nursing home i's strategy in respect to the $F(\cdot)$, probability function of the next period's known state, given the known state of this period and the nursing home's strategy. This is difficult to interpret meaningfully for the purpose of this review, and it is not highly relevant, so this study will emphasize the fact that expected profit, $E\{\pi_i\}$, is of most importance in determining nursing home i's value function. This review uses the relationship between "value" and "expected profit" in an almost interchangeable sense, although they are not the same concept in Lin's article.

The optimal subgame value is the probability function of the current known state x_t given the firm's optimal strategy, σ_{it}^*, and the strategies of other nursing homes, σ_{-it}. It is always equal to or in excess of the probability of another current observed state x_t given non-optimal strategies σ_{it}' and the strategies of other nursing homes, σ_{-it}.

For empirical analysis, Lin uses an adaptation of MPE developed by Maskin and Tirole (1988) and Ericson and Pakes (1995). They use a multinomial logit model to estimate conditional choice probabilities (CCP) of the nursing home firms picking the optimal strategies to generate maximum value per Equation (9). In this case, the CCP is the integral of an indicator function for all nursing home firms,

$$\int I(\sigma_{it}(x_t, \varepsilon_{it}) = a_{it}) dF(\varepsilon_t) \tag{10}$$

which equals 1 if each firm i's action at time t is consistent with the action that their strategy would predict, given their strategy, σ_{it}. This is differentiated on the function of the nursing home i's unknown state, $F(\varepsilon_t)$, which is stochastic on the distribution for all firms. Loosely, the CCP is the probability that each nursing home at time t will pursue an action aligned with a strategy given the observed state of the market at the time, x_t, and unobserved states of the firms, ε_{it}. In Lin's optimization model, they seek to specifically estimate the CCP for the likelihood that the firms will take the optimal actions to maximize their value. Their primary independent variable is instances of entry and exit in the markets, as well as nursing homes' quality level adjustments. Their estimation technique is complex and will not be detailed here, but this literature review is interested in Lin's technique to control for possible endogeneity between nursing home markets, the quality they offer, and their profitability over time. The following outlines their process.

Lin observes that each individual nursing home market is likely to have unique supply-side attributes that could help or hinder each firms' quest for profits. Moreover, there could very likely be a correlation between which market a nursing home is in and the likelihood of nursing homes choosing a specific quality type. Local regulations on the nursing home industry may be present and could realize both of these possibilities in the data. Omitting individual markets' uniqueness could likely amount to an omitted variable bias. Omitted variable bias in this case refers to the attribution of variation between the market outcomes of different nursing homes in different markets to other independent variables such as quality level. This variation could also be brought into the error term. If this trend is serial throughout the dataset, this could result in heterogeneity; that is, the error terms could be correlated with firms' quality level.

To control for heterogeneity or a possible correlation between nursing home quality and the individual characteristics of different nursing home markets, Lin employs a two-stage fixed effects method. The following summary is the author's best understanding of how Lin's method is executed. Stage one is including a dummy variable for each nursing home market included in the sample minus one, $M - 1$. Each dummy would equal 1 if the nursing home i exists within this market, proxied by county. These dummies are meant to control for the unobservable traits of different nursing home markets that might lend to greater profitability of operating in a specific market. Lin then runs two regressions, one with only high-quality nursing homes, and the other with only low-quality nursing homes. The coefficients on the dummies for the high-quality firms' regression estimate individual markets' fixed effects for high-quality firms, or "high fixed effects." The coefficients on the dummies for the low-quality firms' regression estimate the individual markets' fixed effects for low-quality firms, or "low fixed effects." If a low-quality nursing home i is in market M, the coefficient on market M's dummy bumps up or down the intercept of the nursing home's predicted profitability among other low-quality firms in different markets. The same is the case for another high-quality nursing home in the same market M. Lin argues that the relative profitability of a nursing home's selection of quality level in a particular market is indicated by putting these two fixed effects for the same market M side-by-side in the case of markets with both high- and low-quality firms.

The second stage of the fixed effects begins by sorting each market's high- and low-quality fixed effect coefficients into two categories, high profitability and low profitability, based on whether the coefficient is above or below the sample median. Then, each market is sorted into four discrete categories. The first of these categories contains Group I markets, where both high- and low-quality nursing homes have high profitability. Group II markets are where low-quality firms are of low profitability, but high-quality firms are still profitable. Group III markets have profitable low-quality nursing homes, but high-quality nursing homes are of low profitability. In Group IV, both high- and low-quality nursing homes have high profitability. Lin puts a dummy variable for each group into their main regression to control for endogeneity that might exist between quality type and uniqueness of different markets, and how these factors may combine to impact profitability. They write that this method accounts for how markets face different local regulations that might impact nursing homes' profitability.

To account for demand-side factors that may differ between different markets, Lin includes income per capita and the size of the elderly population in the area as control variables. When discussing robustness tests conducted, Lin notes the importance of controlling for income was indicated by their results. They find that higher income per capita in the market is associated with a greater proportion of nursing homes in the market providing high-quality care. Likewise, controlling for the size of the elderly population was critical. Lin find that having a larger elderly population was associated with a greater proportion of lower-quality nursing homes in the market.

Lin uses the parameters estimated in their theoretical and empirical models to test the impact of three hypothetical exogenous changes may have on the structure of nursing home markets and the quality levels they offer. The first of these “simulations” is an increase in the growth rate of the elderly population, justified by a similar real-world trend in the United States. The second is a scenario in which low-quality nursing homes are forced to shut down by the government. The third looks to simulate the impact of a decrease in the cost of nursing home entry. Lin is curious as to what the theoretical impact of each of these policies would be on the number of nursing homes in the market and the level of quality they offer, high or low.

To render the first experiment, Lin presumes that the periods in the game, or years, 6 to 16 would see an increase in the growth rate of the elderly population at 3% each year. The result is that an increase in the elderly population would prompt an increase in the number of firms in the market, which in turn leads to an increase in nursing homes in the market overall, but most of the new nursing homes choose to provide low-quality nursing care. Earlier in their paper, Lin outlines an avenue of possible correlation between the elderly population and quality of care. They argue there is a correlation between the size of the elderly population and the proportion of patients on Medicare, and that higher rates of Medicare-insured patients in a market may be associated with lower quality of care. This could be because the fixed

Medicaid rates chosen by the state might not provide sufficient incentive or cost coverage for nursing homes to provide high-quality care.

Lin imposes the second experiment on the model by imposing the restriction that nursing homes may not choose to operate with a low-quality level; they must always choose to maintain high quality each period or exit. Incumbent low-quality facilities are removed. As a result, most of the nursing home markets, 65%, either are monopolies or are void of any nursing home firm at all. Lin does not speculate in-depth as to why this is the case, but it simply could be cost-prohibitive for some incumbent nursing homes to switch to providing high quality. Also, higher quality costs could result in smaller profit payoffs, disincentivizing entry.

For the third scenario, Lin decreases potential newcomers' entry costs. This prompts another mass entry of nursing homes into markets, but again most of these new entrants choose to offer low-quality care. Lin argues that this is because there are not enough patients with the ability to pay for high care quality in the market. They speculate that the primary reason that nursing homes would choose to offer high-quality care, beyond government regulation,[57] is to attract the few privately paying patients who offer better profitability in exchange for providing service. More nursing homes in the market due to lower entry costs might cause issues with widespread low quality because more nursing homes are competing for the same pool of the privately insured our out-of-pocket payers.

A.7 Lu et al.'s 2017 article, "Does Competition Improve Quality?"

The primary question Lu et al. investigate in their article, "Does Competition Improve Quality? The Case of Nursing Homes Where Public and Private Payers Coexist" is whether

[57] This is admittedly a cold characterization of nursing home administrators and firm executives.

competition is positively or negatively correlated with quality in the nursing home industry, specifically in the case where consumers buy care via private *and* public insurers. Lu et al. point to a greater debate about the role of competition in the health care industry in general. They postulate that competition should be beneficial for the quality of care because it puts the pressure on firms to continually revise their services and push for higher-quality output in order to compete with other firms. A higher number of firms in the market should also drive down the cost of inputs due to increased competition in input markets. However, for the health care industry, there are concerns that having too many providers int the market lead to an over-investment in service provisions, as well as unnecessary inducement of demand,[58] which would represent a shortfall in quality of care.

In addition to this, Lu et al. consider the possibility of nursing home care as a “credence good,” and how it could complicate the impact of oligopolist market structures on nursing home care. A credence good is a good whose quality is unable to be fully observed by the consumer neither before nor after being bought. Nursing home care could be a credence good due to the possible infirmity of people receiving care or the lack of expertise of those who evaluate the care, such as family members. Government inspections and consumer advocacy groups will eventually identify the true quality level of a nursing home’s care, but detection is lagged. Therefore, Lu et al. argue, the reputation a nursing home establishes is critical in consumers’ evaluation of their care prior to and after purchasing it. A nursing home’s quality level adjustment does matter in terms of consumers’ willingness to pay but not in the same time period.

[58] Demand inducement occurs when healthcare infrastructure prompts unneeded treatments for patients who could go without, and these patients might not be able to discern that the treatment is unnecessary due to infirmity or lack of expertise (Khorasani et al., 2015).

Lu et al. looks to address the mix of payer type due to a frequently hypothesized relationship between payer type and quality of care provided. Nursing homes have downward-sloping demand curves; consumers are sensitive to the price they are charged in terms of the quantity of care they demand available. Therefore, those who have the ability to pay for nursing home care at the price set by the nursing home firms, have willingness to pay, given a level of service available in the market. Many patients, however, obviously do not have the ability to pay for nursing home care at the market-established price, which is evident in the very high rate of Medicaid enrollment among nursing home patients. Lu et al. argue that many nursing homes would prefer private-payer patients, whom they may charge higher prices, than the publicly insured patients. They write that this is evident in nursing home behavior in the context of falling Medicaid reimbursement rates since the 1990s, which prompted competition for private payers and subsequent facility bankruptcies. It is illegal for nursing homes to refuse to provide services to patients of different payer types, so nursing homes may find themselves using income from privately insured or out-of-pocket patients to finance equivalent care for publicly insured patients who are charged low, fixed rates.

Lu et al.'s theoretical framework comprises of a repeated game in which there are n nursing home firms, i, in a market, each selecting an output level, x_{it}, and quality level, u_{it} in the beginning of period t. Output is the quantity of services each nursing home i decides to provide. Quality level may be high, h, or low, l.[59] If a nursing home chooses to offer low-quality care, its marginal cost per each additional unit of x_{it} is zero. If it chooses to offer high quality, its marginal cost per each additional unit of x_{it} is c, which is a positive number

[59] Lu et al. technically specify that quality is a continuous value $l \leq u_{it} \leq h$, leaving the specific categorization as to what levels of quality count as "low" or "high" abstract, but when discussing their model in general terms it is practical to simply differentiate as "low" versus "high" as being two distinct categories.

greater than zero. Quality and output level are chosen at the beginning of the time period, and the nursing homes' quality choice in time t is not discernable to consumers or other nursing homes until $t + 1$, not even when consumers have decided to purchase care. An alternative way of visualizing this is to say that consumers are unable to immediately detect whether the quality of care they are receiving is high are low, given the idea that nursing home care may be a credence good.

Lu et al. presume the following linear inverse demand function, which gives consumers' marginal willingness to pay for nursing home i's care at time t:

$$p_{it} = 1 - \frac{2x_{it}}{u_{it}^2} - \frac{2\sigma}{u_{it}} \sum_{j \neq i} \frac{x_{jt}}{u_{jt}} \tag{1}$$

Where $0 < \sigma \leq 1$

Consumers' marginal willingness to pay for nursing home i's care is primarily a function of the quality and output levels provided by i and other nursing homes. As nursing home i produces more services, x_{it}, the Law of Demand holds that consumers are willing to pay less for more abundant care. The quality level of firm i's care inhibits this effect since higher-quality care is more desirable, hence u_{it}^2 in the denominator beneath x_{it}. As other firms produce more output, nursing home care likewise becomes more abundant, lowering consumers' willingness to pay for it, expressed in the term -$\sum_{j \neq i} x_{jt}$. Similar to the effect of firm i's care quality, as other firms increase the quality of their care the negative impact of adding additional service output to the market on price falls, seen by u_{jt} in the denominator of $-\sum_{j \neq i} \frac{x_{jt}}{u_{jt}}$. u_{it}^2 is squared in the second term because the quality level of i's services is more important in determining the impact of additional output on consumers' evaluation of firm i's care than other firms' output, $\sum_{j \neq i} x_{jt}$, and quality, u_{jt}, where the denominator is not squared. The factor on the second term, $\frac{2\sigma}{u_{it}}$, is intended to account for the relative substitutability between nursing home i's services and the services of competitors. Lu describes σ as the "degree of horizontal differentiation,"[60] which refers to how consumers weigh the quality and quantity of other nursing home's services against their valuation of nursing home i's quantity and quality. $\sigma > 0$ presumes that other nursing homes' services are substitutes, albeit imperfect, to nursing home i's. Therefore, their qualities and quantities of services impact consumers' valuation of nursing home i's services. $\sigma = 0$ would hold that there is no substitutability at all between nursing home i's services and others,' so the quantities and qualities of other firms' services have no impact on consumers' valuation of nursing home i's services.

Lu et al. specify two separate scenarios in their theoretical model to account for the segmented demand in the nursing home industry between private and public payers. The purpose of these two theoretical approaches is to illustrate what Lu et al. hypothesize as the determinates of the impact of the number of firms on the levels of care quality firms select, the "reputation-building effect," and "the rent extraction effect." Their model describing the impact of the number of firms in the market on quality of care provided in the context of public payers illustrates the reputation-building effect. The later inclusion of private payers who respond to changes in output and price levels demonstrates the rent extraction effect. These effects are born from the existence of firms' evaluation of punishment and payoff in a

[60] “"Horizontal differentiation" has varying definitions depending on context. According to See (2005), it indicates a situation where rival firms compete on the basis of price with perfectly substitutable goods. According to Machado (2012), products are horizontally differentiated if they are perfect substitutes priced the same, and consumers choose one over the other based on preference. Lu et al. *are* differentiating among different firms' nursing home services on the basis of high or low quality. Also, in their example, consumers' willingness to pay for one firm's services is impacted by the abundance of other firms' in the market, implying imperfect substitutability. Therefore, this review interprets σ as a factor to include the degree of substitutability between nursing home i's services versus others'.

repeating game. The following is a summary of Lu at al.'s reasoning to the best of the author's understanding.

The first scenario, "the regulated segment," presumes that all firm's prices are exogenously fixed, such as the case with Medicare-set reimbursement. Lu et al. isolate x_{it} from the inverse price-demand function, Equation (1), such that

$$x_{it} = \frac{(1-p)u_{it}\left[\left(1+\sigma(n-1)\right)u_{it} - \sigma\sum_{j\neq i} u_{it}\right]}{2(1-\sigma)(1+n\sigma)} \tag{2}$$

At the beginning of the game, every nursing home has no expectation of what the other nursing homes will do. Each is contemplating what quality level to select. Lu et al. illustrate what the profit function for each nursing home would be if all chose to produce high-quality service:

$$\pi_h = \frac{(1-p)(p-c)h^2}{2(1+n\sigma)} \tag{3}$$

If every nursing home chose to produce high quality service, each would reap a payoff of π_h. Next, Lu et al. describe what would happen if a nursing home were to deviate from this strategy for a period. A nursing home might be incentivized to produce low quality services instead because consumers would be unable to tell within that period their quality is low, so the amount of services demanded from that nursing home would be unaffected by this decision. As a result, the "deviation" profits would be

$$\pi_d = \frac{(1-p)ph^2}{2(1+n\sigma)} \tag{4}$$

A deviating nursing home would reap a payoff greater than the other, high-quality nursing homes by ch^2.

In subsequent periods, consumers and other firms will respond to this tactic, beginning in what Lu et al. call the "punishment phase." Some number of firms, m, will switch to providing lower quality in subsequent periods, leaving $n + 1 - m$ nursing homes to offer high-quality. Lu et al. solve for the level of output the m deviating nursing homes would produce, x_{pm},

$$x_{pm} = \frac{(1-p)l\big[h\sigma\big(m-(n+1)\big) + l\sigma(n-m) + l\big]}{2(1-\sigma)(1+n\sigma)} \tag{5}$$

From this, Lu et al. indicate that as more firms decide to switch to low quality, m, the service output of each of those firms, x_{pm}, increases. This is logical because switching to providing low quality service causes the nursing homes' marginal costs to fall to zero. In that case, they gain additional revenue opportunity by producing more services. From this, Lu et al. show that in period t the m deviating nursing homes providing low quality reap a profit payoff of

$$\pi_{pm} = px_{pm} \tag{6}$$

which is simply the output they produce times the exogenously fixed Medicaid price. Costs are not a factor here because Lu et al.'s model holds that their marginal costs would be zero as a result of producing low-quality care.

From these assertions, Lu et al. construct an "incentive compatibility constraint," which gives the condition that must be true in the case that any nursing homes produce quality care:

$$\frac{\pi_h}{1-\delta} \geq \pi_d - B + \frac{\delta}{1-\delta}\pi_{pm} \tag{7}$$

B represents sunk cost associated with adjusting between quality levels in period t. δ represents the "discount factor"[61] for the relative valuation firms place on their future payoffs, $0 \leq \delta \leq 1$. The closer to 1 the discount factor is, the closer a nursing home's valuation of future payoff is to its valuation of payoff in the current period. All nursing homes have equivalent discount factors; δ is the same for each nursing home. The left-hand side of the inequality indicates future profits from continuing to provide high-quality care, discounted for future periods. The righthand side indicates the payoff from deviating to provide low-quality care instead, π_d, minus adjustment costs, B, with punished payoffs in future periods, π_{pm}, discounted via the nursing home's relative valuation of future profits, δ. Simply put, the level of the firms' discount factor δ determines the likelihood that a firm will opt to deviate to offering low quality instead of high quality, conditional on Equation (7).

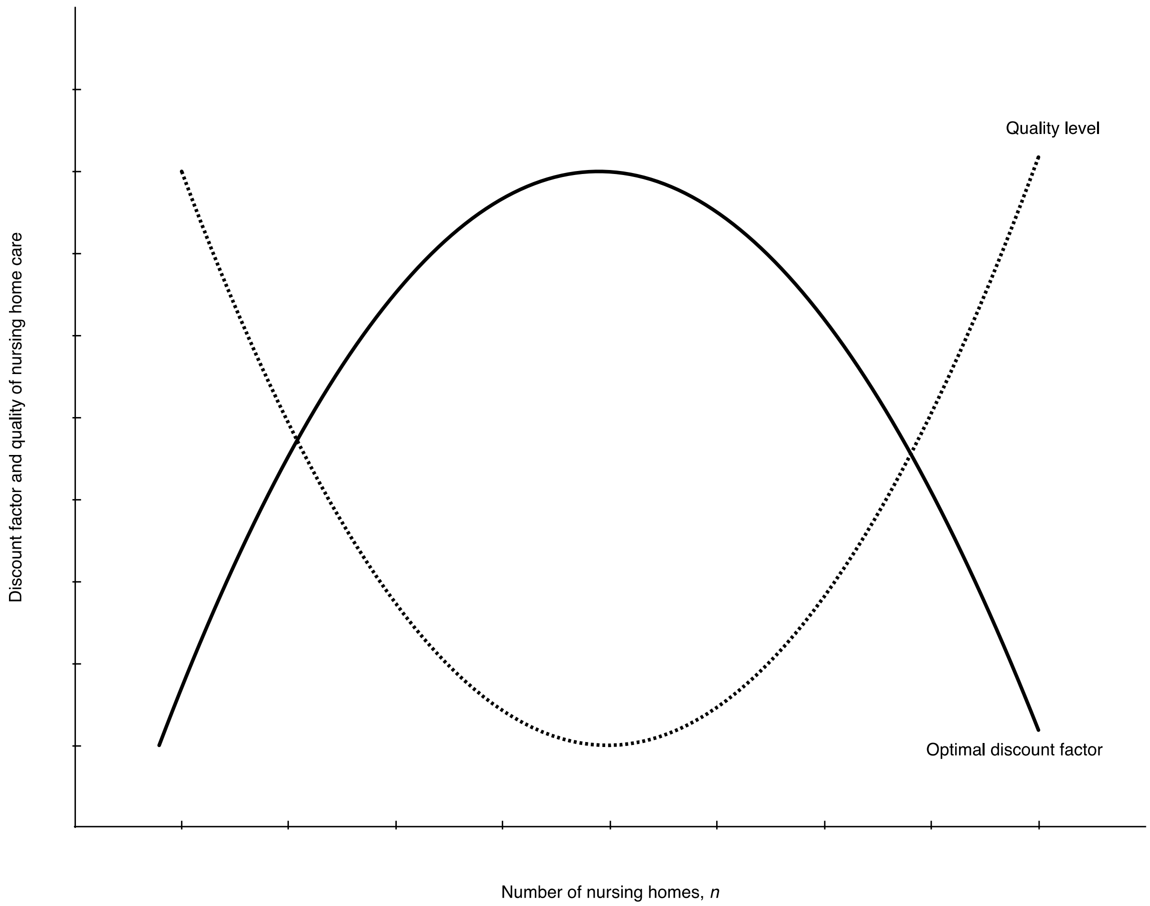
Lu et al. hold that there is a critical discount factor threshold, $\hat{\delta}$, nursing home firms must have to maintain an equilibrium in which all nursing homes consistently choose to provide high quality, and they demonstrate that the more firms there are in the market, n, the optimal discount ratio $\hat{\delta}$ decreases. As more nursing homes enter the market, all nursing homes have smaller relative valuation of their payoffs in the next period compared to their valuation of payoff today. In sum, this is due to the fact that market share x_{it} falls for every firm as more firms enter the market, lowering the incentive to cut costs by deviation since costs decrease overall with a fall in the number of patients. Also, deviators' profits fall faster than non-deviators. They call this the "reputation building effect," since firms will have a tendency to promote consumers' expectations of their goods over time, which leads to an

[61] This Literature Review will interpret "discount factor" similarly to how it is defined by Huang (2015), as a ratio to describe how an actor relatively values a payoff in the current time period versus payoffs in the future.

increase in their market share, and therefore increased profit. They eschew cutting costs and providing low quality. Therefore, incentive to opt for a possibly higher future payoff by alternating to providing low quality falls as more firms enter the market, in the case that the exogenously fixed price cannot be altered by any of these trends.

For the nonregulated case, Lu et al. construct a similar model to determine nursing homes' market shares, high-quality profitability, and deviation profitability are when price may change in tandem with consumers' willingness to pay per Equation (1). Using the same incentive compatibility constraint, Equation (7), Lu et al. illustrate that, as n increases, $\hat{\delta}$ rises and falls in a parabolic, upside-down "U" shape. Markets that have very large or small numbers of firms have nursing homes with a very low optimal discount factor, where markets in the middle (specifically, about 15 nursing homes) have high relative valuation of future payoffs. Lu et al. argue that as firms' discount factor changes, so too does their incentive to provide quality care, similar to the regulated example above. When there are very large or small numbers of nursing homes in the market, nursing homes will have greater incentive to provide a higher quality of care. Lu et al. attribute this relationship both to the reputation building effect and to the "rent extraction effect," in which nursing homes earn less return from being in an oligopolist market as a result of increased competition for patients. On the other side of the spectrum, when the market is very competitive, producing higher quality of care differentiates nursing home i from other firms, attracting market share. The following figure aims to summarize the relationship between the number of firms in the market, their discount factor, and the quality of care they offer. The illustration imitates Lu et al.'s figures describing their theorized relationship between the discount factor and the number of nursing homes but adds a curve to show the resulting trend relating quality to market competition:

Figure 1: Lu et al.'s hypothesized relationship between the number of nursing homes in the market, the optimal discount factor, and quality of care



Lu et al. postulate that the parabolic shape of this effect is due to two factors. First, when there are very few firms in the market, nursing homes' ability to get a return via profits, called rent extraction, causes prices to be so high due to low quantities available in the market. Therefore, they are able to sustain more cost-intensive quality investment per unit of output. When there are very many firms in the market, the need for individual nursing homes to differentiate themselves and attract market share via providing quality is more intense, causing the same effect. With the middle amount of nursing homes in the market, neither effect is strong enough to incentivize nursing homes to offer high quality of care over the cost savings from offering low quality of care. The competition among nursing homes is fierce enough to disincentivize providing quality due to the promotion of lower prices, but not fierce enough to push firms to differentiate themselves via providing distinguished levels of care quality. Lu et al. discuss the theoretical outcome when the regulated and unregulated segments are combined, simulating a market where there is a mix of Medicare and non-Medicare patients. They write that combining the two segments mitigates this U-shaped relationship since the reputation-building effect prevails under more circumstances regarding the number of firms in the market. However, the parabolic relationship remains due to price competition for private paying patients.

Lu et al.'s empirical analysis focuses on two different hypotheses: First, they want to investigate whether markets with a higher proportion of publicly-insured patients, namely Medicaid patients, have a stronger positive correlation with the number of firms and quality. Second, they look to test whether there is a parabolic relationship between the number of firms in the market and the level of quality they produce. They use data from the 2000-2005 Online Survey Certification and Reporting (OSCAR) CMS database, which contains a multitude of datapoints on thousands of facilities in the United States, including financial and quality-relevant characteristics. These include 190 quality of care metrics used by CMS to construct their five-star rating scorecard Nursing Home Compare. The sample has 62,228 observations, which are each nursing homes at different points in time throughout the 6-year period. The cross sections also include characteristics from Medicare's cost reports for skilled nursing facilities and the Census Bureau.

To measure quality of care, Lu et al. use output-based quality metrics from OSCAR. They keep their focus on reputation as a product of care quality, since they interpret quality deficiency citations as representative of both reputation changes and lagged real quality change. Deficiency citations impact consumers' expectations of care,[62] which is intrinsically important to the impact of firm behavior in Lu et al.'s model. They use preliminary probit regressions with a lagged, $t - 1$, variable for deficiency citations to see their impact on the probability that a nursing home will change its name in subsequent time periods. They demonstrate that there is an association of a 7.2% increase in the likelihood of a facility changing its name or exiting with an increase in one standard deviation of deficiency citations in prior periods. Also, similarly to Lin (2015), Lu et al. test their output-based measure of quality deficiency citations for robustness against an important input-based metric, licensed nursing hours per patient day.

Lu et al. measure competitive intensity in their primary regressions as the number of other nursing homes within a five-mile market radius from nursing home i. While they test for robustness with Hirschman-Herfindahl indices,[63] they favor the number of firms in the market because they view it as more exogenous. They also ensure they do not count other nursing homes within the radius that are a part of the same brand chain of facilities, insisting that different facilities should be considered the same if they are operated by the same firm and do not represent competitive rivals. Lu et al. use various control variables, such as the proportion of Medicaid insured patients in the market, the number of beds in a facility,

[62] They are featured in Nursing Home Compare scorecards.
[63] The index is a measure of market power via oligopoly or monopoly. The index is calculated by summing the squared proportions of each firms' market share. An index of 1 represents a monopoly (Zinn 1996).

generic descriptors of patients' relative health, income per capita, and percent of African Americans in the county relative to the whole population.

Lu et al. opt for instrumental variables and fixed effects in their regressions to control for possible endogeneity. Firms might be more willing to enter a market due to its unique consumers and their preference to pay for nursing home care. These preferences may impact the same consumers' relative valuation of different levels of care quality. Also, nursing homes might be less willing to enter if the market already has many firms operating at high quality, suggesting tighter competition among rivals if they were to enter. Lu et al. select four variables to be instruments: The total number of patients in nursing homes within the market radius, lagged by a year, the population of the elderly in the county, lagged by a year, and the squares of both of these metrics. They choose these metrics because they argue these market characteristics are unlikely to be directly predictive of the quality of care nursing homes provide, but still would be associated with the level of competition between firms in the nursing home market. More nursing homes patients would indicate more nursing home facilities, and a greater elderly population may represent of the provision of demand for local nursing home care. The equation for their empirical model is as follows,[64] estimated via a two-stage least squares method:

$$\begin{aligned} Y_{i,m,t} = \beta_0 + \beta_1 \ln\left(options_{i,t}\right) + \beta_2\left[\ln\left(options_{i,t}\right)\right]^2 + \\ \beta_3 X_{i,m,t-1} + \beta_4 Z_{m,t} + \gamma_t + \gamma_i + \varepsilon_{i,m,t} \end{aligned} \tag{8}$$

[64] Equation (8) represents Lu et al.'s simplification of their true primary regression equation, where the variables $X_{i,m,t-1}$, $Z_{m,t}$, γ_t and γ_t are in place of a multitude of variables with separate coefficients. Lu et al. utilize many control variables, as well as fixed effects for time- and observation-dependent impacts on the quality of care each nursing home provides in each time period. In addition, Lu et al. run many iterations of this regression for robustness testing and for implementation of their various instrumental variables. They include an alternative linear specification not detailed in this review.

On the left-hand side of the regression equation sits the dependent variable, $Y_{i,m,t}$, nursing home i's level of care quality in market m at time t. The second and third terms on the right-hand side represent the natural log of the squared and unsquared number of other nursing homes within a five-mile radius of nursing home i at time t, called $options$.

Lu et al. distinguish between supply-side and demand-side control variables with the terms $X_{i,m,t-1}$ and $Z_{m,t}$, where $X_{i,m,t-1}$ is encompasses number amount of beds, general health level of patients, and proportion of Medicaid-insured patients, all in nursing home i during time $t-1$. $Z_{m,t}$ represents demand-relevant control variables including income per capita, Medicaid price levels mandated by the state, and the proportion of African Americans in the population, all for the market m at time t for nursing home i. Lu et al.'s fixed effects for each cross-sectional i and year t in the sample are represented in Equation (8) by γ_t and γ_i. In another specification, Lu et al. include an interaction term where the number of other firms in the market is multiplied by lagged quality of care indicators. The substitutability of different nursing homes' care within the same market plays a role in determining consumers' willingness to pay, and a relationship between substitutability and nursing homes' care quality may exist. The interaction term accounts for different markets possibly demonstrating differing levels of substitutability between firms of different or similar quality among the markets.

From this baseline regression, and other alternative specifications not detailed here, Lu et al. demonstrate evidence of a parabolic relationship between number of firms in the market and the quality-of-care firms produce, as they predict in their theoretical framework. In order for their prediction of a parabolic relationship between the number of firms in a nursing home market and the quality of care provided by individual nursing homes to be corroborated by their model, β_1 must be positive and β_2 must be negative. This is the case in their primary specification and in their robustness checks by $p < 0.01$ level of confidence. The coefficients in their primary specification are 4.51 ($SE = 1.00$) and -1.54 ($SE = 0.39$) for the unsquared and squared market concentration terms respectively. In other words, the first coefficient indicates that an additional (non-squared) nursing home within a 5-mile radius of nursing home i is associated with an increase in quality deficiency citations. The negative sign on the squared term indicates that the direction of this relationship may reverse for greater magnitudes of firms in the market, such that at some point adding nursing homes to the market is associated with fewer quality of care deficiencies.

Their results are indicative that consumer expectations and market structure both play a role in the narrative surrounding the quality of care in the nursing home market. Reported shortcomings in quality may not be immediately relevant in consumers' evaluation of a nursing home's care, which could cause disincentives for nursing homes to maintain higher quality levels in favor of higher immediate payoff. The number of options for care in the market may either help or hinder the transfer of information to consumers. A moderate number of firms in the market make evaluating the relative quality levels between facilities difficult but still keep prices too low to strongly incentivize quality. The extreme ends of spectrum of the number of firms may sponsor stronger incentives for higher quality of care due to nursing homes' need to differentiate themselves from others and due to higher prices increasing the affordability of providing quality care.

A.8 Theory Section Proofs

Deriving x_1^* in the regulated case

Each new line is a step in the simplification process.

$$\lambda = \lambda$$

$$\gamma_1 - 2q_1\left(\frac{1}{x_1}\right) - \frac{q_2\sigma}{x_2x_1} - \cdots - \frac{q_n\sigma}{x_nx_1} = r\left(\frac{1}{EbL_1^aK^{b-1}}\right)$$

$$-2q_1\left(\frac{1}{x_1}\right) - \frac{q_2\sigma}{x_2x_1} - \cdots - \frac{q_n\sigma}{x_nx_1} = r\left(\frac{1}{EbL_1^aK^{b-1}}\right) - \gamma_1$$

$$2q_1\left(\frac{1}{x_1}\right) + \frac{q_2\sigma}{x_2x_1} + \cdots + \frac{q_n\sigma}{x_nx_1} = -r\left(\frac{1}{EbL_1^aK^{b-1}}\right) + \gamma_1$$

$$\frac{1}{x_1}\left(2q_1 + \frac{q_2\sigma}{x_2} + \cdots + \frac{q_n\sigma}{x_n}\right) = \frac{-r}{EbL_1^aK^{b-1}} + \gamma_1$$

$$2q_1 + \frac{q_2\sigma}{x_2} + \cdots + \frac{q_n\sigma}{x_n} = x_1\left(\frac{-r}{EbL_1^aK^{b-1}} + \gamma_1\right)$$

$$\frac{2q_1 + \frac{q_2\sigma}{x_2} + \cdots + \frac{q_n\sigma}{x_n}}{\frac{-r}{EbL_1^aK^{b-1}} + \gamma_1} = x_1$$

Isolating q_1^* in the private-payer case

Each new line is a step in the simplification process.

$$p_1 = \gamma_1 - \frac{1}{x_1}q_1 - \frac{\sigma}{x_1}\left(\frac{q_2}{x_2} + \cdots + \frac{q_n}{x_n}\right)$$

$$p_1 - \gamma_1 + \frac{\sigma}{x_1}\left(\frac{q_2}{x_2} + \cdots + \frac{q_n}{x_n}\right) = -\frac{1}{x_1}q_1$$

$$p_1 - \gamma_1 + \frac{q_2\sigma}{x_1x_2} + \cdots + \frac{q_n\sigma}{x_1x_n} = -\frac{1}{x_1}q_1$$

$$-p_1 + \gamma_1 - \frac{q_2\sigma}{x_1x_2} - \cdots - \frac{q_n\sigma}{x_1x_n} = \frac{1}{x_1}q_1$$

$$-p_1x_1 + \gamma_1x_1 - \frac{q_2\sigma x_1}{x_1x_2} - \cdots - \frac{q_n\sigma x_1}{x_1x_n} = q_1$$

$$-p_1x_1 + \gamma_1x_1 - \frac{q_2\sigma}{x_2} - \cdots - \frac{q_n\sigma}{x_n} = q_1$$

Isolating x_1^* in the private-payer case

Each new line is a step in the simplification process.

$$\lambda = \lambda$$

$$2p_1 - \gamma_1 + \frac{\sigma q_2}{x_1 x_2} + \cdots + \frac{\sigma q_n}{x_1 x_n} = r\left(\frac{1}{EbL_1^a K^{b-1}}\right)$$

$$2p_1 - \gamma_1 + \frac{1}{x_1}\left(\frac{\sigma q_2}{x_2} + \cdots + \frac{\sigma q_n}{x_n}\right) = \frac{r}{EbL_1^a K^{b-1}}$$

$$\frac{1}{x_1}\left(\frac{\sigma q_2}{x_2} + \cdots + \frac{\sigma q_n}{x_n}\right) = \frac{r}{EbL_1^a K^{b-1}} - 2p_1 + \gamma_1$$

$$\frac{\sigma q_2}{x_2} + \cdots + \frac{\sigma q_n}{x_n} = x_1\left(\frac{r}{EbL_1^a K^{b-1}} - 2p_1 + \gamma_1\right)$$

$$\frac{\frac{\sigma q_2}{x_2} + \cdots + \frac{\sigma q_n}{x_n}}{\frac{r}{EbL_1^a K^{b-1}} - 2p_1 + \gamma_1} = x_1$$

A.9. Do-Files

Below are pasted versions of the do-files used in empirical analysis. The first six do-files were used for cleaning the data. The seventh was used for merging the data into a cohesive set. The eighth do-file was use for the actual estimations. All do-files and estimations were edited in StataSE.

A.9.1 do_clean_county_cases

```
//Tessa Ireton
//7 December 2020

clear all
set more off

import delimited using "/Users/tjnireton/Desktop/IS_DataFiles/statewide_cases.csv"

//COMPRESS DATES TO WEEKLY AVERAGE
	//First, convert date string variables into datetimes
		//First, replace "-" in date strings with "."
		gen dateform = subinstr(date,"/",".",2)
		//next, turn that variable into a string variable with proper format
		gen day = date(dateform, "MD20Y" )
		//Then, turn the string-date variable into a datetime type
		format day %td
```

//then, drop all days before one week before 5/24/2020 (earliest week-end in nursing home COVID data)

drop if day < date("05242020","MDY")

//Third, generate a new variable, endwk, that is an average of the newcountconfirmed over the week, first week-end 2020-05-24

//First, generate new variable that stores what the week-end date of each day is (sunday)

gen endwk = day - dow(day) +7

//put endwk in datetime format

format endwk %td

//drop unneeded variables

drop totalcountconfirmed totalcountdeaths newcountdeaths date dateform day

//collapse the newcountconfirmed variable to be an average of the newcountconfirmeds over the week, the obs with the same county and the same endwk date:

collapse (sum) newcountconfirmed , by(county endwk)

//newcountconfirmed is the new case count in that counting in the week ending with date == endwk

//drop unneeded or unexplainable categories: Out Of Country, Unassigned

drop if county == "Out Of Country" | county == "Unassigned"

//ensure continuity on Do_Merge

rename county COUNTY

rename newcountconfirmed cnty_newcountconfirmed

replace COUNTY = subinstr(COUNTY ," County","",1)

//remove out-of-bounds dates. //DATE BOUNDS: 31may2020 to 22nov2020

drop if endwk < date("05312020","MDY")

drop if endwk > date("11222020","MDY")

//EXPORT AND SAVE AS .dta

save "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/statewide_cases.dta" , replace

//This .dta will be called by the analysis do file.

A.9.2 do_clean_elderly_pop

//Tessa Ireton

//9 December 2020

clear all

set more off

log using "/Users/tjnireton/Desktop/IS_DataFiles/Dos_Clean/Log_clean_county_elderlypop" , replace

import delimited using "/Users/tjnireton/Desktop/IS_DataFiles/cc-est2019-agesex-06.csv"

```
//Drop old years. Year 12 represents that the estimates are for 7/1/2019
	keep if year == 12

//calculate elderly population by dividing the #population of people age 65 and older by the total population *100

gen cnty_elderlypercent = (age65plus_tot / popestimate)*100

//drop uneeded variables
	keep ctyname cnty_elderlypercent

//ensure continuity on Do_Merge
	rename ctyname COUNTY
	replace COUNTY = subinstr(COUNTY ," County","",1)

//EXPORT AND SAVE AS .dta
	save "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/cc-est2019-agesex-06.dta" ,
replace //This .dta will be called by the analysis do file.

log close
```

A.9.3 do_clean_county_ethnicity

```
//Tessa Ireton
//9 December 2020

clear all
set more off

import delimited using "/Users/tjnireton/Desktop/IS_DataFiles/cc-est2019-alldata-06.csv"

//get rid of uneeded ears
	keep if year == 12
	keep if agegrp == 0

//calculate percent black : bac_male = Black or African American alone or in combination
male, bac_female = Black or African American alone or in combination female
	gen popAAper = ((bac_male + bac_female) / tot_pop )*100

//get rid of unneeded variables
	keep ctyname tot_pop popAAper

//ensure continuity on Do_Merge
	rename ctyname COUNTY
	rename tot_pop cnty_totpop
```

rename popAAper cnty_popAAper
replace COUNTY = subinstr(COUNTY ," County","",1)

//EXPORT AND SAVE AS .dta
save "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/cc-est2019-alldata-06.dta" , replace //This .dta will be called by the analysis do file.

A.9.4 do_clean_county_poverty

//Tessa Ireton
//9 December 2020

clear all
set more off

import excel using "/Users/tjnireton/Desktop/IS_DataFiles/est18all.xls", sheet(est18ALL) cellrange(A4:AE3198) firstrow

//Drop uneeded observations. I only need observations from California, and I don't need california totals
drop if PostalCode != "CA"
drop if Name == "California"

//drop uneeded variables. only need county and PovertyPercentAllAges
keep Name PovertyPercentAllAges MedianHouseholdIncome

//rename to emphasize the date of data
rename PovertyPercentAllAges cnty_PovertyPercent2018

//Destring the poverty percents
destring cnty_PovertyPercent2018 , replace
destring MedianHouseholdIncome , replace

//ensure continuity on Do_Merge
rename Name COUNTY
rename MedianHouseholdIncome cnty_MedianHouseholdIncome2018
replace COUNTY = subinstr(COUNTY ," County","",1)

//EXPORT AND SAVE AS .dta
save "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/est18all.dta" , replace //This .dta will be called by the analysis do file.

A.9.5 do_clean_NH_cases

//Tessa Ireton
//10 December 2020

```
clear all
set more off

import delimited using "/Users/tjnireton/Desktop/IS_DataFiles/COVID-
19_Nursing_Home_Dataset.csv"

//drop unneeded observations
	drop if submitteddata == "N"
	drop if providerstate != "CA"

//drop unneeded variables
	keep ïweekending providername providercity provideraddress providerzipcode
passedqualityassurancecheck totalnumberofoccupiedbeds abletotestorobtainresourcestotes
shortageofnursingstaff shortageofclinicalstaff shortageofaides shortageofotherstaff
oneweeksupplyofn95masks oneweeksupplyofsurgicalmasks oneweeksupplyofhandsanitizer
weeklyresidentconfirmedcovid19ca county

//rename some of these variables
	rename ïweekending wkendday
	rename passedqualityassurancecheck QApass
	rename totalnumberofoccupiedbeds totoccup_beds
	rename abletotestorobtainresourcestotes test_able
	rename shortageofnursingstaff RN_short
	rename shortageofaides aides_short
	rename shortageofclinicalstaff clinical_short
	rename shortageofotherstaff otherstaff_short
	rename oneweeksupplyofn95masks n95_wksupp
	rename oneweeksupplyofhandsanitizer sanitizer_wksupp
	rename oneweeksupplyofsurgicalmasks surgicalmasks_wksupp
	rename weeklyresidentconfirmedcovid19ca wkly_cases_per1000res

//destring variables
	destring totoccup_beds , replace
	destring wkly_cases_per1000res , generate(wklynewcases_1000res) force
	drop wkly_cases_per1000res

//string variables
	tostring providerzipcode , replace

//dummies: turn "Y" into 1 numeric, and "N" into 0 numeric
	//testable dummy
		replace test_able = "1" if test_able == "Y"
		replace test_able = "0" if test_able == "N"
		replace test_able = "." if test_able == ""
		destring test_able , generate(testable_dum) force
```

```
		drop test_able
		//==1 if the nursing home is able to test all residents or obtain the resources to do so within the next seven days
	//RN_short dummy
		replace RN_short = "1" if RN_short == "Y"
		replace RN_short = "0" if RN_short == "N"
		replace RN_short = "." if RN_short == ""
		destring RN_short , generate(RN_short_dum) force
		drop RN_short
		//==1 if there is a shortage of nursing staff
	//clinical_short dummy
		replace clinical_short = "1" if clinical_short == "Y"
		replace clinical_short = "0" if clinical_short == "N"
		replace clinical_short = "." if clinical_short == ""
		destring clinical_short , generate(clinical_short_dum) force
		drop clinical_short
		//1 if there is a shortage of clinical staff
	//aides_short dummy
		replace aides_short = "1" if aides_short == "Y"
		replace aides_short = "0" if aides_short == "N"
		replace aides_short = "." if aides_short == ""
		destring aides_short , generate(aides_short_dum) force
		drop aides_short
		//==1 if there is a shortage of aides staff
	//otherstaff_short dummy
		replace otherstaff_short = "1" if otherstaff_short == "Y"
		replace otherstaff_short = "0" if otherstaff_short == "N"
		replace otherstaff_short = "." if otherstaff_short == ""
		destring otherstaff_short , generate(otherstaff_short_dum) force
		drop otherstaff_short
		//==1 if there is a shortge of other staff
	//n95_wksupp dummy
		replace n95_wksupp = "1" if n95_wksupp == "Y"
		replace n95_wksupp = "0" if n95_wksupp == "N"
		replace n95_wksupp = "." if n95_wksupp == ""
		destring n95_wksupp , generate(n95_wksupp_dum) force
		drop n95_wksupp
		//==1 if the nh has a 7 day supply of n95 masks
	//surgicalmasks_wksupp dummy
		replace surgicalmasks_wksupp = "1" if surgicalmasks_wksupp == "Y"
		replace surgicalmasks_wksupp = "0" if surgicalmasks_wksupp == "N"
		replace surgicalmasks_wksupp = "." if surgicalmasks_wksupp == ""
		destring surgicalmasks_wksupp , generate(surgicalmasks_wksupp_dum) force
		drop surgicalmasks_wksupp
		//==1 if nh has a week supply of surgical masks
	//sanitizer_wksupp dummy
```

```
replace sanitizer_wksupp = "1" if sanitizer_wksupp == "Y"
replace sanitizer_wksupp = "0" if sanitizer_wksupp == "N"
replace sanitizer_wksupp = "." if sanitizer_wksupp == ""
destring sanitizer_wksupp , generate(sanitizer_wksupp_dum) force
drop sanitizer_wksupp
//==1 if the nh has a week supply of hand sanitizer
//QApass
replace QApass = "1" if QApass == "Y"
replace QApass = "0" if QApass == "N"
replace QApass = "." if QApass == ""
destring QApass , generate(QApass_dum) force
drop QApass
//==1 if the nh passed the quality assurance check

//label dummies
//RN_short_dum
label define shortstaff_label 0 "no shortage" 1 "shortage"
label values RN_short_dum shortstaff_label
//clinical_short_dum
label values clinical_short_dum shortstaff_label
//aides_short_dum
label values aides_short_dum shortstaff_label
//otherstaff_short_dum
label values otherstaff_short_dum shortstaff_label
//n95_wksupp_dum
label define wksupply_label 0 "not week supply" 1 "week supply"
label values n95_wksupp_dum wksupply_label
//surgicalmasks_wksupp_dum
label values surgicalmasks_wksupp_dum wksupply_label
//sanitizer_wksupp_dum
label values sanitizer_wksupp_dum wksupply_label
//QApass
label define QApass_label 0 "failed" 1 "passed"
label values QApass_dum QApass_label
//testable_dum
label define testable_label 0 "not able" 1 "able"
label values testable_dum testable_label

//convert week-end dates to datetime format
//First, convert date string variables into datetimes
//First, replace "-" in date strings with "."
gen dateform = subinstr(wkendday,"/",".",2)
//next, turn that variable into a string variable with proper format
gen week_end = date(wkendday, "MDY" )
//Then, turn the string-date variable into a datetime type
format week_end %td
```

```
            //drop string date variable
            drop wkendday
            //drop unneeded
            drop dateform

//rename for ease of merge
      rename week_end endwk
      rename county COUNTY
      sort providername endwk

//remove ampersands
      replace providername = subinstr(providername,"&","AND",10)

//remove out-of-bounds dates. //DATE BOUNDS: 31may2020 to 22nov2020
      drop if endwk < date("05312020","MDY")
      drop if endwk > date("11222020","MDY")

///INSURE UNIFORMITY IN FACILITY NAMES BETWEEN DATASETS
//remove leading spaces and trailing spaces on providername, merge variable
      replace providername = strtrim(providername)
//remove ampersands
      replace providername = subinstr(providername,"&","AND",10)
//remove hyphens
      replace providername = subinstr(providername, " - ", "-", 10)
      replace providername = subinstr(providername, "-", " ", 10)
//remove "D/P SNF"
      replace providername = subinstr(providername, "D/P SNF", "", 10)
      replace providername = subinstr(providername, "DP/SNF", "", 10)
//remove ", LP"
      replace providername = subinstr(providername, ", LP", "", 10)
//remove inc
      replace providername = subinstr(providername, ", INC.", "", 10)
      replace providername = subinstr(providername, " INC.", "", 10)
      replace providername = subinstr(providername, " INC", "", 10)
      replace providername = subinstr(providername, ", INC", "", 10)
//replace rehab with rehabilitation
      replace providername = subinstr(providername, "REHAB ", "REHABILITATION ",
10)
      replace providername = subinstr(providername, "REHAB. ", "REHABILITATION ",
10)
//conv to convalescent
      replace providername = subinstr(providername, "CONV ", "CONVALESCENT ",
10)
      replace providername = subinstr(providername, "CONV. ", "CONVALESCENT ",
10)
//centre, CNTR to center
```

```
	replace providername = subinstr(providername, "CENTRE", "CENTER", 10)
	replace providername = subinstr(providername, "CNTR", "CENTER", 10)
// get rid of LLC
	replace providername = subinstr(providername, ", LLC", "", 10)
	replace providername = subinstr(providername, " LLC", "", 10)
	replace providername = subinstr(providername, ", L.L.C.", "", 10)
	replace providername = subinstr(providername, " L.L.C.", "", 10)
//HCC vs Health Care Center
	replace providername = subinstr(providername, " HCC", " HEALTH CARE CENTER", 10)
//nrs ctr to nursing center
	replace providername = subinstr(providername, " NRS CT", " NURSING CENTER", 10)
//nrsg to nursing
	replace providername = subinstr(providername, "NRSG ", "NURSING ", 10)
//subacute for sub acute
	replace providername = subinstr(providername, "SUB ACUTE", "SUBACUTE", 10)
//hlth care change to HEALH CARE
	replace providername = subinstr(providername, "HLTH CARE", "HEALTH CARE", 10)
//HEALTH CARE instead of HEALTHCARE
	replace providername = subinstr(providername, "HEALTHCARE", "HEALTH CARE", 10)
//Skilled nursing for SKLD NURSING
	replace providername = subinstr(providername, "SKLD NURS", "SKILLED NURSING", 10)
//hospital instead of hosp.
	replace providername = subinstr(providername, "HOSP.", "HOSPITAL", 10)
//st. for st
	replace providername = subinstr(providername, "ST. ", "ST ", 10)
//remove apostrophes
	replace providername = subinstr(providername, "'", "", 10)
//specific typos:
	replace providername = "ALAMEDA CARE CENTER BURBANK" if providername == "ALAMEDA CARE CENTER"
	replace providername = "ALAMITOS BELMONT REHABILITATION HOSPITAL" if providername =="ALAMITOS BELMONT REHABILITATION HOSPITA"
	replace providername = subinstr(providername, "MED CTR", "MEDICAL CENTER", 10)
	replace providername = subinstr(providername, "CTR", "CENTER", 10)
	replace providername = subinstr(providername, "  ", " ", 10)
	replace providername = subinstr(providername, "HOSPITALL", "HOSPITAL", 10)
	replace providername = "CHINO VALLEY HEALTH CARE CENTER" if providername == "CHINO VALLEY HEALTH CARE CENTE"
	replace providername = "AMERICAN RIVER CARE CENTER" if providername == "AMERICAN RIVER CENTER"
```

replace providername = "ARCADIA HEALTH CARE CENTER" if providername == "ARCADIA CARE CENTER"

replace providername = "BAYBERRY SKILLED NURSING AND REHABILITATION CENTER" if providername == "BAYBERRY SKILLED NURSING AND HEALTH CARE CENTER"

replace providername = "BEL TOOREN VILLA CONVALESCENT HOSPITAL" if providername == "BEL TOOREN VILLA CONVALESCENT"

replace providername = "BERKLEY VALLEY CONVALESCENT HOSPITAL" if providername == "BERKLEY VALLEY CONV HOSPITAL"

replace providername = "BROOKDALE RANCH MIRAGE" if providername == "BROOKDALE RANCHO MIRAGE"

replace providername = "BROOKDALE RIVERWALK SNF CA" if providername == "BROOKDALE RIVERWALK SNF (CA)"

replace providername = "BUENA VISTA CARE CENTER SANTA BARBARA" if providername == "BUENA VISTA CARE CENTER" & COUNTY == "Santa Barbara"

replace providername = "BUENA VISTA CARE CENTER ANAHEIM" if providername == "BUENA VISTA CARE CENTER" & COUNTY == "Orange"

replace providername = "BURBANK HEALTH CARE AND REHABILITATION CENTER" if providername == "BURBANK HEALTH CARE AND REHAB"

replace providername = "CASA DE LAS CAMPANAS" if providername == "CASA DE LAS CAMPANAS HEALTH CENTER"

replace providername = "COMMUNITY CARE CENTER DUARTE" if providername == "COMMUNITY CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "COUNTRY CREST POST ACUTE" if providername == "COUNTRY CREST" & COUNTY == "Butte"

replace providername = "COUNTRY OAKS CARE CENTER SANTA MARIA" if providername == "COUNTRY OAKS CARE CENTER" & COUNTY == "Santa Barbara"

replace providername = "COUNTRY VILLA BELMONT HEIGHTS HEALTH CARE CENTER" if providername == "COUNTRY VILLA BELMONT HEIGHTS" & COUNTY == "Los Angeles"

replace providername = "COUNTRY VILLA CLAREMONT HEALTH CARE CENTER" if providername == "COUNTRY VILLA CLAREMONT HEALTH CENTER" & COUNTY == "Los Angeles"

replace providername = "COUNTRY VILLA HACIENDA HEALTH CARE CENTER" if providername == "COUNTRY VILLA HEALTH CARE" & COUNTY == "San Bernardino"

replace providername = "COUNTRY VILLA HACIENDA HEALTH CARE CENTER" if providername == "COUNTRY VILLA HEALTH CARE" & COUNTY == "San Bernardino"

replace providername = "COUNTRY VILLA BELMONT HEIGHTS HEALTH CARE CENTER" if providername == "COUNTRY VILLA BELMONT HEIGHTS" & COUNTY == "Los Angeles"

replace providername = "COUNTRY VILLA CLAREMONT HEALTH CARE CENTER" if providername == "COUNTRY VILLA CLAREMONT HEALTH CENTER" & COUNTY == "Los Angeles" replace providername = "COUNTRY VILLA NORTH CONVALESCENT CENTER" if providername == "COUNTRY VILLA NORTH" & COUNTY == "Los Angeles"

replace providername = "COUNTRY VILLA SHERATON NURSING AND REHABILITATION CENTER" if providername == "COUNTRY VILLA SHERATON" & COUNTY == "Los Angeles"

replace providername = "COUNTRY VILLA WESTWOOD CONVALESCENT CENTER" if providername == "COUNTRY VILLA WESTWOOD" & COUNTY == "Los Angeles"

replace providername = "COURTYARD CARE CENTER SIGNAL HILL" if providername == "COURTYARD CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "COVENANT VILLAGE OF TURLOCK CARE CENTER" if providername == "COVENANT VILLAGE CARE CENTER" & COUNTY == "Stanislaus"

replace providername = "CRESTWOOD TREATMENT CENTER FREMONT" if providername == "CRESTWOOD TREATMENT CENTER" & COUNTY == "Alameda"

replace providername = "DEL AMO GARDENS CONVALESCENT" if providername == "DEL AMO GARDENS CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "DELTA NURSING AND REHABILITATION HOSPITAL" if providername == "DELTA NURSING AND REHABILITATION CENTER" & COUNTY == "Tulare"

replace providername = "EASTLAND SUBACUTE AND REHABILITATION CENTER" if providername == "EASTLAND SUBACUTE AND REHABILITATION" & COUNTY == "Los Angeles"

replace providername = "EL ENCANTO HEALTH CARE AND HABILITATION CENTER" if providername == "EL ENCANTO HEALTH CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "ENGLISH OAKS CONVALESCENT HOSPITAL AND REHABILITATION CENTER" if providername == "ENGLISH OAKS CONVALESCENT AND REHABILITATION HOSPITA" & COUNTY == "Stanislaus"

replace providername = "GEM TRANSITIONAL CARE CENTER" if providername == "GEM TRANSITIONAL" & COUNTY == "Los Angeles"

replace providername = "GRANADA HILLS CONVALESCENT HOSPITAL" if providername == "GRANADA HILLS CONVALESCENT" & COUNTY == "Los Angeles"

replace providername = "GRANT CUESTA SUBACUTE AND REHABILITATION HOSPITAL" if providername == "GRANT CUESTA SUBACUTE AND REHABILITATION CENTER" & COUNTY == "Santa Clara"

replace providername = "GREEN ACRES HEALTH CARE CENTER" if providername == "GREEN ACRES LODGE" & COUNTY == "Los Angeles"

replace providername = "HARBOR POST ACUTE CARE CENTER" if providername == "HARBOR CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "KATHERINE HEALTH CARE CENTER" if providername == "KATHERINE HEALTH CARE" & COUNTY == "Monterey"

replace providername = "KENNEDY POST ACUTE CARE CENTER" if providername == "KENNEDY CARE CENTER" & COUNTY == "Los Angeles" replace providername = "LEGACY POST ACUTE REHABILITATION CENTER" if providername == "LEGACY POST ACUTE REHABILITATION" & COUNTY == "San Bernardino"

replace providername = "LINDSAY GARDENS NURSING AND REHABILITATION" if providername == "LINDSAY GARDENS AND REHABILITATION" & COUNTY == "Tulare"

replace providername = "LITTLE SISTERS OF THE POOR SAN PEDRO" if providername == "LITTLE SISTERS OF THE POOR" & COUNTY == "Tulare"

replace providername = "LONGWOOD MANOR CONVALESCENT HOSPITAL" if providername == "LONGWOOD MANOR CONV.HOSPITAL" & COUNTY == "Los Angeles"

replace providername = subinstr(providername, "MANOR CARE", "MANORCARE", 10)

replace providername = "MANORCARE HEALTH SERVICES ROSSMOOR" if providername == "MANOR CARE NURSING CENTER OF WALNUT CREEK" & COUNTY == "Contra Costa"

replace providername = "MANORCARE HEALTH SERVICES CITRUS HEIGHTS" if providername == "MANORCARE HEALTH SERVICES (CITRUS HEIGHTS)" & COUNTY == "Sacramento"

replace providername = "MANORCARE HEALTH SERVICES FOUNTAIN VALLEY" if providername == "MANORCARE HEALTH SERVICES (FOUNTAIN VALLEY)" & COUNTY == "Orange"

replace providername = "MANORCARE HEALTH SERVICES SUNNYVALE" if providername == "MANORCARE HEALTH SERVICES (SUNNYVALE)" & COUNTY == "Santa Clara"

replace providername = "MARLORA POST ACUTE REHABILITATION HOSPITAL" if providername == "MARLORA POST ACUTE REHABILITATION HOSP" & COUNTY == "Los Angeles"

replace providername = "MARQUIS CARE OF SHASTA" if providername == "MARQUIS CARE AT SHASTA" & COUNTY == "Shasta"

replace providername = "MARY HEALTH OF THE SICK CONVALESCENT AND NURSING HOSPITAL" if providername == "MARY HEALTH OF THE SICK CONVALESCENT AND NURSING HOS" & COUNTY == "Ventura"

replace providername = "MONTEREY PARK CONVALESCENT HOSPITAL" if providername == "MONTEREY PARK CONVALESCENT HOSP" & COUNTY == "Los Angeles"

replace providername = "MOUNTAIN VIEW CONVALESCENT HOSPITAL" if providername == "MOUNTAIN VIEW CONVALESCENT HOSP" & COUNTY == "Los Angeles"

replace providername = "NEW BETHANY SKILLED NURSING" if providername == "NEW BETHANY" & COUNTY == "Merced"

replace providername = "NEW VISTA POST ACUTE CARE CENTER" if providername == "NEW VISTA POST ACUTE CARE CENTER WEST L.A." & COUNTY == "Los Angeles" replace providername = "PALM VILLAGE RETIREMENT COMMUNITY HEALTH CARE CENTER" if providername == "PALM VILLAGE RETIREMENT COMM." & COUNTY == "Fresno"

replace providername = "PARADISE VALLEY HEALTH CARE CENTER" if providername == "PARADISE VALLEY HEALTH CARE" & COUNTY == "San Diego"

replace providername = "PLAYA DEL REY CARE AND REHABILITATION CENTER" if providername == "PLAYA DEL REY CENTER" & COUNTY == "Los Angeles"

replace providername = "RAMONA REHABILITATION AND POST ACUTE CARE CENTER" if providername == "RAMONA REHABILITATION AND POST ACUTE CENTER" & COUNTY == "Riverside"

replace providername = "RANCHO MIRAGE HEALTH AND REHABILITATION CENTER" if providername == "RANCHO MIRAGE HEALTH AND REHABILATION CENTER" & COUNTY == "Riverside"

replace providername = "RIVERSIDE CONVALESCENT HOSPITAL CHICO" if providername == "RIVERSIDE CONVALESCENT HOSPITAL" & COUNTY == "Butte"

replace providername = "SAN JOAQUIN NURSING CENTER AND REHABILITATION CENT" if providername == "SAN JOAQUIN NURSING CENTER" & COUNTY == "Kern"

replace providername = "SHERMAN OAKS HEALTH AND REHABILITATION CENTER" if providername == "SHERMAN OAKS HEALTH AND REHAB" & COUNTY == "Los Angeles"

replace providername = "SHIELDS RICHMOND NURSING CENTER" if providername == "SHIELDS/RICHMOND NURSING CENTER" & COUNTY == "Contra Costa"

replace providername = "SIERRA VISTA REHABILITATION CENTER" if providername == "SIERRA VISTA" & COUNTY == "San Bernardino"

replace providername = "SKYLINE HEALTH CARE CENTER LOS ANGELES" if providername == "SKYLINE HEALTH CARE CENTER LA" & COUNTY == "Los Angeles"

replace providername = "SMITH RANCH SKILLED NURSING AND REHABILITATION CENTER" if providername == "SMITH RANCH SKILLED NURSING AND REHABILITATION CENTE" & COUNTY == "Marin"

replace providername = "SPRING HILL MANOR CONVALESCENT HOSPITAL" if providername == "SPRING HILL MANOR" & COUNTY == "Nevada"

replace providername = "ST JOHN OF GOD RETIREMENT AND CARE CENTER" if providername == "ST JOHN OF GOD RETIREMENT" & COUNTY == "Los Angeles"

replace providername = "SUN MAR NURSING CENTER ANAHEIM" if providername == "SUN MAR NURSING CENTER" & COUNTY == "Orange"

replace providername = "SUNNYVALE POST ACUTE CENTER" if providername == "SUNNYVALE POST ACUTE CARE" & COUNTY == "Santa Clara"

replace providername = "SUNSET MANOR CONVALESCENT HOSPITAL" if providername == "SUNSET MANOR CONVALESCENT HOSP" & COUNTY == "Los Angeles" replace providername = "SYLMAR HEALTH AND REHABILITATION CENTER" if providername == "SYLMAR HLTH REHABILITATION CENTER" & COUNTY == "Los Angeles"

replace providername = "TEMPLE PARK CONVALESCENT HOSPITAL" if providername == "TEMPLE PARK CONVALESCENT HOSP" & COUNTY == "Los Angeles"

replace providername = "RIVERSIDE BEHAVIORAL HEALTH CARE CENTER" if providername == "RIVERSIDE HEALTH CARE CENTER" & COUNTY == "Riverside"

replace providername = "THE CALIFORNIAN PASADENA" if providername == "CALIFORNIAN PASADENA CONVALESCENT HOSP" & COUNTY == "Los Angeles"

replace providername = "COVINGTON CARE CENTER THE" if providername == "THE COVINGTON CARE CENTER" & COUNTY == "Orange"

replace providername = "DOROTHY AND JOSEPH GOLDBERG HEALTH CARE CENTER" if providername == "THE DOROTHY AND JOSEPH GOLDBERG HEALTH CARE CENTER" & COUNTY == "San Diego"

replace providername = "THE MEADOWS OF NAPA VALLEY" if providername == "MEADOWS OF NAPA VALLEY THE" & COUNTY == "Napa"

replace providername = "THE REHABILITATION CENTER OF BEVERLY HILLS" if providername == "REHABILITATION CENTER OF BEVERLY HILLS" & COUNTY == "Los Angeles"

replace providername = "THE ROWLAND SKILLED NURSING FACILITY" if providername == "THE ROWLAND" & COUNTY == "Los Angeles"

replace providername = "THE TAMALPAIS" if providername == "TAMALPAIS" & COUNTY == "Marin"

replace providername = "THE TERRACES OF LOS GATOS" if providername == "TERRACES OF LOS GATOS THE" & COUNTY == "Santa Clara"

replace providername = "TOPANGA TERRACE CONVALESCENT CENTER" if providername == "TOPANGA TERRACE" & COUNTY == "Los Angeles"

replace providername = "TOWN AND COUNTRY MANOR" if providername == "TOWN AND COUNTRY" & COUNTY == "Orange"

replace providername = "VACAVILLE CONVALESCENT AND REHABILITATION CENTER" if providername == "VACAVILLE CONVALESCENT AND REHAB" & COUNTY == "Solano"

replace providername = "VALLEY CONVALESCENT HOSPITAL BAKERSFIELD" if providername == "VALLEY CONVALESCENT HOSPITAL" & COUNTY == "Kern"

replace providername = "VALLEY CONVALESCENT HOSPITAL WATSONVILLE" if providername == "VALLEY CONVALESCENT HOSPITAL" & COUNTY == "Santa Cruz"

replace providername = "VICTORIA HEALTH CARE AND REHABILITATION CENTER" if providername == "VICTORIA HEALTH CARE CENTER" & COUNTY == "Orange"

replace providername = "VIEW HEIGHTS CONVALESCENT HOSPITAL" if providername == "VIEW HEIGHTS CONVALESCENT HOSP" & COUNTY == "Los Angeles"

```
	replace providername = "VIEW PARK CONVALESCENT CENTER" if providername == "VIEW PARK CONVALESCENT HOSP" & COUNTY == "Los Angeles"
	replace providername = "VILLA MARIN RETIREMENT RESIDENCES" if providername == "VILLA MARIN" & COUNTY == "Marin"
	replace providername = "VILLA SCALABRINI SPECIAL CARE UNIT" if providername == "VILLA SCALABRINI SPECIAL CARE" & COUNTY == "Los Angeles"
	replace providername = "VILLA VALENCIA HEALTH CARE CENTER" if providername == "VILLA VALENCIA" & COUNTY == "Orange"
	replace providername = "WALNUT CREEK SKILLED NURSING AND REHABILITATION CENTER" if providername == "WALNUT CREEK SKILLED NURSING AND REHABILITATION CENT" & COUNTY == "Contra Costa"
	replace providername = "WEST HAVEN HEALTH CARE CENTER" if providername == "WEST HAVEN HEALTH CARE" & COUNTY == "Los Angeles"
	replace providername = "WINDSOR GARDENS HEALTH CARE CENTER OF THE VALLEY" if providername == "WINDSOR GARDENS HEALTH CARE OF THE VALLEY" & COUNTY == "Los Angeles"
	replace providername = "WINDSOR MANOR REHABILITATION CENTER OF CONCORD" if providername == "WINDSOR MANOR REHABILITATION CENTER" & COUNTY == "Contra Costa"
	replace providername = "WINDSOR TERRACE HEALTH CARE CENTER" if providername == "WINDSOR TERRACE HEALTH CARE" & COUNTY == "Los Angeles"
//remove lingering commas
	replace providername = subinstr(providername, ",", "", 10)
//again leading spaces and trailing spaces on providername, merge variable
	replace providername = strtrim(providername)

//EXPORT AND SAVE AS .dta
	save "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/COVID-19_Nursing_Home_Dataset.dta" , replace //This .dta will be called by the analysis do file.
```

A.9.6 do_clean_NH_characteristics

```
//Tessa Ireton
//7 December 2020

clear all
set more off

import excel using "/Users/tjnireton/Desktop/IS_DataFiles/lafd1219-sub-initial_forodp.xlsx", firstrow

//drop useless observations (referring to 2012 Data File Documentation document)
```

//some facilities face different reporting standards and are not required to submit the full amount of documentation to OSHPD. These are marked as "non-comparable", as opposed to "comparable" datapoints that draw from the fully-submitted paperwork

drop if COMPARABLE != "Comparable"

//Not all facilities in this set adhere to this IS' definition of "Nursing Home." This IS will use CLHFs, SNFs, and SNF/RESs

keep if LIC_CAT == "SNF" | LIC_CAT == "SNF/RES" | LIC_CAT == "CLHF"

//I only want reports that cover the entire year

drop if DAY_PER != 365

//keep only needed variables

keep FAC_NO FAC_NAME ADDRESS CITY ZIP_CODE LIC_CAT TYPE_CNTRL COUNTY PRDHR_GNP PRDHR_RN PRDHR_LVN DAY_TOTL DAY_MCAL GR_RT_MCAL GR_AN_MCAL_IP GR_AN_MCAL_IP GR_AN_MCAL_OP GR_RT_TOTL GR_AN_TOTL

//PRDHR_GNP "Productive Hours Geriatric Nurse Practitioners"

//PRDHR_RN "Productive Hours Registered Nurses"

//PRDHR_LVN "Productive Hours Licensed Vocational Nurses"

//DAY_TOTL "Patient (Census) Days total"

//DAY_MCAL "Patient (Census) Days Medi-Cal"

//GR_RT_MCAL "Gross Routine Services Revenue Medi-Cal"

//GR_AN_MCAL_IP "Gross Ancillary Services Revenue Medi-Cal Inpatient"

//GR_AN_MCAL_OP "Gross Ancillary Services Revenue Medi-Cal Outpatient"

//GR_RT_TOTL "Gross Routine Services Revenue Total"

//GR_AN_TOTL "Gross Ancillary Services Revenue Total" //I sum these to get total services revenue.

//replace "."s in numerical variables with

//PROBLEM

egen tot_services_revenue = rowtotal(GR_RT_TOTL GR_AN_TOTL) , missing

//if I try to use +, it only puts a missing variable for tot_services_revenue instead of counting the "." as 0 ("egen var = rowtotal(varlist), missing" is supposed to, but my stata says that "rowtotal()" is an unknown function)

egen medi_services_revenue = rowtotal(GR_AN_MCAL_IP GR_AN_MCAL_OP) , missing

//same problem as above

//then I want to calculate the proportion of services revenue that is medicaid revenue (medi-Cal).

gen Medical_pro_services_revenue = medi_services_revenue / tot_services_revenue

```
//make a variable that indicates the number of nursing homes in nursing home i's county
	egen competition = count(FAC_NO), by (COUNTY)
	//so each NH is not counted as its own competator:
	replace competition = competition - 1

//make a variable that indicates the Hirschman-Herfindahl index of the nursing home.
//formula: the HH index = adding the squared proportion of each nursing home’s portion of
the total patient census days in the market.
	//make a variable that has the total market share in the county (total Patient Days)
	egen Total_MS = total(DAY_TOTL) , by (COUNTY)
	//Calculate each NH's squared_MS:
	gen squared_MS = (DAY_TOTL/Total_MS)^2
	//Calculate each Market's HHI
	egen HHI = total(squared_MS) , by (COUNTY)
	drop Total_MS squared_MS

//Calculate different types of productive care hours per patient day:
	gen PRDHR_GNP_ppd = PRDHR_GNP/DAY_TOTL
	gen PRDHR_RN_ppd = PRDHR_RN/DAY_TOTL
	gen PRDHR_LVN_ppd = PRDHR_LVN/DAY_TOTL
	drop PRDHR_GNP PRDHR_RN PRDHR_LVN

//Change variable names for ease of merge
	rename FAC_NAME providername
	rename ADDRESS provideraddress
	rename ZIP_CODE providerzipcode
	rename CITY providercity
	sort providername

///INSURE UNIFORMITY IN FACILITY NAMES BETWEEN DATASETS
//remove leading spaces and trailing spaces on providername, merge variable
	replace providername = strtrim(providername)
//remove ampersands
	replace providername = subinstr(providername,"&","AND",10)
//remove hyphens
	replace providername = subinstr(providername, " - ", "-", 10)
	replace providername = subinstr(providername, "-", " ", 10)
//remove "D/P SNF"
	replace providername = subinstr(providername, "D/P SNF", "", 10)
	replace providername = subinstr(providername, "DP/SNF", "", 10)
//remove ", LP"
	replace providername = subinstr(providername, ", LP", "", 10)
//remove inc
	replace providername = subinstr(providername, ", INC.", "", 10)
	replace providername = subinstr(providername, " INC.", "", 10)
```

```
	replace providername = subinstr(providername, " INC", "", 10)
	replace providername = subinstr(providername, ", INC", "", 10)
//replace rehab with rehabilitation
	replace providername = subinstr(providername, "REHAB ", "REHABILITATION ",
10)
	replace providername = subinstr(providername, "REHAB. ", "REHABILITATION ",
10)
//conv to convalescent
	replace providername = subinstr(providername, "CONV ", "CONVALESCENT ",
10)
	replace providername = subinstr(providername, "CONV. ", "CONVALESCENT ",
10)
//centre, CNTR to center
	replace providername = subinstr(providername, "CENTRE", "CENTER", 10)
	replace providername = subinstr(providername, "CNTR", "CENTER", 10)
// get rid of LLC
	replace providername = subinstr(providername, ", LLC", "", 10)
	replace providername = subinstr(providername, " LLC", "", 10)
	replace providername = subinstr(providername, ", L.L.C.", "", 10)
	replace providername = subinstr(providername, " L.L.C.", "", 10)
//HCC vs Health Care Center
	replace providername = subinstr(providername, " HCC", " HEALTH CARE
CENTER", 10)
//nrs ctr to nursing center
	replace providername = subinstr(providername, " NRS CT", " NURSING CENTER",
10)
//nrsg to nursing
	replace providername = subinstr(providername, "NRSG ", "NURSING ", 10)
//subacute for sub acute
	replace providername = subinstr(providername, "SUB ACUTE", "SUBACUTE", 10)
//hlth care change to HEALH CARE
	replace providername = subinstr(providername, "HLTH CARE", "HEALTH CARE",
10)
//HEALTH CARE instead of HEALTHCARE
	replace providername = subinstr(providername, "HEALTHCARE", "HEALTH
CARE", 10)
//Skilled nursing for SKLD NURSING
	replace providername = subinstr(providername, "SKLD NURS", "SKILLED
NURSING", 10)
//hospital instead of hosp.
	replace providername = subinstr(providername, "HOSP.", "HOSPITAL", 10)
//st. for st
	replace providername = subinstr(providername, "ST. ", "ST ", 10)
//remove apostrophes
	replace providername = subinstr(providername, "'", "", 10)
//specific typos:
```

replace providername = "ALAMEDA CARE CENTER BURBANK" if providername == "ALAMEDA CARE CENTER"

replace providername = "ALAMITOS BELMONT REHABILITATION HOSPITAL" if providername =="ALAMITOS BELMONT REHABILITATION HOSPITA"

replace providername = subinstr(providername, "MED CTR", "MEDICAL CENTER", 10)

replace providername = subinstr(providername, "CTR", "CENTER", 10)

replace providername = subinstr(providername, " ", " ", 10)

replace providername = subinstr(providername, "HOSPITALL", "HOSPITAL", 10)

replace providername = "CHINO VALLEY HEALTH CARE CENTER" if providername == "CHINO VALLEY HEALTH CARE CENTE"

replace providername = "AMERICAN RIVER CARE CENTER" if providername == "AMERICAN RIVER CENTER"

replace providername = "ARCADIA HEALTH CARE CENTER" if providername == "ARCADIA CARE CENTER"

replace providername = "BAYBERRY SKILLED NURSING AND REHABILITATION CENTER" if providername == "BAYBERRY SKILLED NURSING AND HEALTH CARE CENTER"

replace providername = "BEL TOOREN VILLA CONVALESCENT HOSPITAL" if providername == "BEL TOOREN VILLA CONVALESCENT"

replace providername = "BERKLEY VALLEY CONVALESCENT HOSPITAL" if providername == "BERKLEY VALLEY CONV HOSPITAL"

replace providername = "BROOKDALE RANCH MIRAGE" if providername == "BROOKDALE RANCHO MIRAGE"

replace providername = "BROOKDALE RIVERWALK SNF CA" if providername == "BROOKDALE RIVERWALK SNF (CA)"

replace providername = "BUENA VISTA CARE CENTER SANTA BARBARA" if providername == "BUENA VISTA CARE CENTER" & COUNTY == "Santa Barbara"

replace providername = "BUENA VISTA CARE CENTER ANAHEIM" if providername == "BUENA VISTA CARE CENTER" & COUNTY == "Orange"

replace providername = "BURBANK HEALTH CARE AND REHABILITATION CENTER" if providername == "BURBANK HEALTH CARE AND REHAB"

replace providername = "CASA DE LAS CAMPANAS" if providername == "CASA DE LAS CAMPANAS HEALTH CENTER"

replace providername = "COMMUNITY CARE CENTER DUARTE" if providername == "COMMUNITY CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "COUNTRY CREST POST ACUTE" if providername == "COUNTRY CREST" & COUNTY == "Butte"

replace providername = "COUNTRY OAKS CARE CENTER SANTA MARIA" if providername == "COUNTRY OAKS CARE CENTER" & COUNTY == "Santa Barbara"

replace providername = "COUNTRY VILLA BELMONT HEIGHTS HEALTH CARE CENTER" if providername == "COUNTRY VILLA BELMONT HEIGHTS" & COUNTY == "Los Angeles"

replace providername = "COUNTRY VILLA CLAREMONT HEALTH CARE CENTER" if providername == "COUNTRY VILLA CLAREMONT HEALTH CENTER" & COUNTY == "Los Angeles" replace providername = "COUNTRY VILLA HACIENDA HEALTH CARE CENTER" if providername == "COUNTRY VILLA HEALTH CARE" & COUNTY == "San Bernardino"

replace providername = "COUNTRY VILLA HACIENDA HEALTH CARE CENTER" if providername == "COUNTRY VILLA HEALTH CARE" & COUNTY == "San Bernardino"

replace providername = "COUNTRY VILLA BELMONT HEIGHTS HEALTH CARE CENTER" if providername == "COUNTRY VILLA BELMONT HEIGHTS" & COUNTY == "Los Angeles"

replace providername = "COUNTRY VILLA CLAREMONT HEALTH CARE CENTER" if providername == "COUNTRY VILLA CLAREMONT HEALTH CENTER" & COUNTY == "Los Angeles"

replace providername = "COUNTRY VILLA NORTH CONVALESCENT CENTER" if providername == "COUNTRY VILLA NORTH" & COUNTY == "Los Angeles"

replace providername = "COUNTRY VILLA SHERATON NURSING AND REHABILITATION CENTER" if providername == "COUNTRY VILLA SHERATON" & COUNTY == "Los Angeles"

replace providername = "COUNTRY VILLA WESTWOOD CONVALESCENT CENTER" if providername == "COUNTRY VILLA WESTWOOD" & COUNTY == "Los Angeles"

replace providername = "COURTYARD CARE CENTER SIGNAL HILL" if providername == "COURTYARD CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "COVENANT VILLAGE OF TURLOCK CARE CENTER" if providername == "COVENANT VILLAGE CARE CENTER" & COUNTY == "Stanislaus"

replace providername = "CRESTWOOD TREATMENT CENTER FREMONT" if providername == "CRESTWOOD TREATMENT CENTER" & COUNTY == "Alameda"

replace providername = "DEL AMO GARDENS CONVALESCENT" if providername == "DEL AMO GARDENS CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "DELTA NURSING AND REHABILITATION HOSPITAL" if providername == "DELTA NURSING AND REHABILITATION CENTER" & COUNTY == "Tulare"

replace providername = "EASTLAND SUBACUTE AND REHABILITATION CENTER" if providername == "EASTLAND SUBACUTE AND REHABILITATION" & COUNTY == "Los Angeles"

replace providername = "EL ENCANTO HEALTH CARE AND HABILITATION CENTER" if providername == "EL ENCANTO HEALTH CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "ENGLISH OAKS CONVALESCENT HOSPITAL AND REHABILITATION CENTER" if providername == "ENGLISH OAKS CONVALESCENT AND REHABILITATION HOSPITA" & COUNTY == "Stanislaus"

replace providername = "GEM TRANSITIONAL CARE CENTER" if providername == "GEM TRANSITIONAL" & COUNTY == "Los Angeles" replace providername = "GRANADA HILLS CONVALESCENT HOSPITAL" if providername == "GRANADA HILLS CONVALESCENT" & COUNTY == "Los Angeles"

replace providername = "GRANT CUESTA SUBACUTE AND REHABILITATION HOSPITAL" if providername == "GRANT CUESTA SUBACUTE AND REHABILITATION CENTER" & COUNTY == "Santa Clara"

replace providername = "GREEN ACRES HEALTH CARE CENTER" if providername == "GREEN ACRES LODGE" & COUNTY == "Los Angeles"

replace providername = "HARBOR POST ACUTE CARE CENTER" if providername == "HARBOR CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "KATHERINE HEALTH CARE CENTER" if providername == "KATHERINE HEALTH CARE" & COUNTY == "Monterey"

replace providername = "KENNEDY POST ACUTE CARE CENTER" if providername == "KENNEDY CARE CENTER" & COUNTY == "Los Angeles"

replace providername = "LEGACY POST ACUTE REHABILITATION CENTER" if providername == "LEGACY POST ACUTE REHABILITATION" & COUNTY == "San Bernardino"

replace providername = "LINDSAY GARDENS NURSING AND REHABILITATION" if providername == "LINDSAY GARDENS AND REHABILITATION" & COUNTY == "Tulare"

replace providername = "LITTLE SISTERS OF THE POOR SAN PEDRO" if providername == "LITTLE SISTERS OF THE POOR" & COUNTY == "Tulare"

replace providername = "LONGWOOD MANOR CONVALESCENT HOSPITAL" if providername == "LONGWOOD MANOR CONV.HOSPITAL" & COUNTY == "Los Angeles"

replace providername = subinstr(providername, "MANOR CARE", "MANORCARE", 10)

replace providername = "MANORCARE HEALTH SERVICES ROSSMOOR" if providername == "MANOR CARE NURSING CENTER OF WALNUT CREEK" & COUNTY == "Contra Costa"

replace providername = "MANORCARE HEALTH SERVICES CITRUS HEIGHTS" if providername == "MANORCARE HEALTH SERVICES (CITRUS HEIGHTS)" & COUNTY == "Sacramento"

replace providername = "MANORCARE HEALTH SERVICES FOUNTAIN VALLEY" if providername == "MANORCARE HEALTH SERVICES (FOUNTAIN VALLEY)" & COUNTY == "Orange"

replace providername = "MANORCARE HEALTH SERVICES SUNNYVALE" if providername == "MANORCARE HEALTH SERVICES (SUNNYVALE)" & COUNTY == "Santa Clara"

replace providername = "MARLORA POST ACUTE REHABILITATION HOSPITAL" if providername == "MARLORA POST ACUTE REHABILITATION HOSP" & COUNTY == "Los Angeles"

replace providername = "MARQUIS CARE OF SHASTA" if providername == "MARQUIS CARE AT SHASTA" & COUNTY == "Shasta"

replace providername = "MARY HEALTH OF THE SICK CONVALESCENT AND NURSING HOSPITAL" if providername == "MARY HEALTH OF THE SICK CONVALESCENT AND NURSING HOS" & COUNTY == "Ventura" replace providername = "MONTEREY PARK CONVALESCENT HOSPITAL" if providername == "MONTEREY PARK CONVALESCENT HOSP" & COUNTY == "Los Angeles"

replace providername = "MOUNTAIN VIEW CONVALESCENT HOSPITAL" if providername == "MOUNTAIN VIEW CONVALESCENT HOSP" & COUNTY == "Los Angeles"

replace providername = "NEW BETHANY SKILLED NURSING" if providername == "NEW BETHANY" & COUNTY == "Merced"

replace providername = "NEW VISTA POST ACUTE CARE CENTER" if providername == "NEW VISTA POST ACUTE CARE CENTER WEST L.A." & COUNTY == "Los Angeles"

replace providername = "PALM VILLAGE RETIREMENT COMMUNITY HEALTH CARE CENTER" if providername == "PALM VILLAGE RETIREMENT COMM." & COUNTY == "Fresno"

replace providername = "PARADISE VALLEY HEALTH CARE CENTER" if providername == "PARADISE VALLEY HEALTH CARE" & COUNTY == "San Diego"

replace providername = "PLAYA DEL REY CARE AND REHABILITATION CENTER" if providername == "PLAYA DEL REY CENTER" & COUNTY == "Los Angeles"

replace providername = "RAMONA REHABILITATION AND POST ACUTE CARE CENTER" if providername == "RAMONA REHABILITATION AND POST ACUTE CENTER" & COUNTY == "Riverside"

replace providername = "RANCHO MIRAGE HEALTH AND REHABILITATION CENTER" if providername == "RANCHO MIRAGE HEALTH AND REHABILATION CENTER" & COUNTY == "Riverside"

replace providername = "RIVERSIDE CONVALESCENT HOSPITAL CHICO" if providername == "RIVERSIDE CONVALESCENT HOSPITAL" & COUNTY == "Butte"

replace providername = "SAN JOAQUIN NURSING CENTER AND REHABILITATION CENT" if providername == "SAN JOAQUIN NURSING CENTER" & COUNTY == "Kern"

replace providername = "SHERMAN OAKS HEALTH AND REHABILITATION CENTER" if providername == "SHERMAN OAKS HEALTH AND REHAB" & COUNTY == "Los Angeles"

replace providername = "SHIELDS RICHMOND NURSING CENTER" if providername == "SHIELDS/RICHMOND NURSING CENTER" & COUNTY == "Contra Costa"

replace providername = "SIERRA VISTA REHABILITATION CENTER" if providername == "SIERRA VISTA" & COUNTY == "San Bernardino"

replace providername = "SKYLINE HEALTH CARE CENTER LOS ANGELES" if providername == "SKYLINE HEALTH CARE CENTER LA" & COUNTY == "Los Angeles"

replace providername = "SMITH RANCH SKILLED NURSING AND REHABILITATION CENTER" if providername == "SMITH RANCH SKILLED NURSING AND REHABILITATION CENTE" & COUNTY == "Marin"

replace providername = "SPRING HILL MANOR CONVALESCENT HOSPITAL" if providername == "SPRING HILL MANOR" & COUNTY == "Nevada" replace providername = "ST JOHN OF GOD RETIREMENT AND CARE CENTER" if providername == "ST JOHN OF GOD RETIREMENT" & COUNTY == "Los Angeles"

replace providername = "SUN MAR NURSING CENTER ANAHEIM" if providername == "SUN MAR NURSING CENTER" & COUNTY == "Orange"

replace providername = "SUNNYVALE POST ACUTE CENTER" if providername == "SUNNYVALE POST ACUTE CARE" & COUNTY == "Santa Clara"

replace providername = "SUNSET MANOR CONVALESCENT HOSPITAL" if providername == "SUNSET MANOR CONVALESCENT HOSP" & COUNTY == "Los Angeles"

replace providername = "SYLMAR HEALTH AND REHABILITATION CENTER" if providername == "SYLMAR HLTH REHABILITATION CENTER" & COUNTY == "Los Angeles"

replace providername = "TEMPLE PARK CONVALESCENT HOSPITAL" if providername == "TEMPLE PARK CONVALESCENT HOSP" & COUNTY == "Los Angeles"

replace providername = "RIVERSIDE BEHAVIORAL HEALTH CARE CENTER" if providername == "RIVERSIDE HEALTH CARE CENTER" & COUNTY == "Riverside"

replace providername = "THE CALIFORNIAN PASADENA" if providername == "CALIFORNIAN PASADENA CONVALESCENT HOSP" & COUNTY == "Los Angeles"

replace providername = "COVINGTON CARE CENTER THE" if providername == "THE COVINGTON CARE CENTER" & COUNTY == "Orange"

replace providername = "DOROTHY AND JOSEPH GOLDBERG HEALTH CARE CENTER" if providername == "THE DOROTHY AND JOSEPH GOLDBERG HEALTH CARE CENTER" & COUNTY == "San Diego"

replace providername = "THE MEADOWS OF NAPA VALLEY" if providername == "MEADOWS OF NAPA VALLEY THE" & COUNTY == "Napa"

replace providername = "THE REHABILITATION CENTER OF BEVERLY HILLS" if providername == "REHABILITATION CENTER OF BEVERLY HILLS" & COUNTY == "Los Angeles"

replace providername = "THE ROWLAND SKILLED NURSING FACILITY" if providername == "THE ROWLAND" & COUNTY == "Los Angeles"

replace providername = "THE TAMALPAIS" if providername == "TAMALPAIS" & COUNTY == "Marin"

replace providername = "THE TERRACES OF LOS GATOS" if providername == "TERRACES OF LOS GATOS THE" & COUNTY == "Santa Clara"

replace providername = "TOPANGA TERRACE CONVALESCENT CENTER" if providername == "TOPANGA TERRACE" & COUNTY == "Los Angeles"

replace providername = "TOWN AND COUNTRY MANOR" if providername == "TOWN AND COUNTRY" & COUNTY == "Orange"

replace providername = "VACAVILLE CONVALESCENT AND REHABILITATION CENTER" if providername == "VACAVILLE CONVALESCENT AND REHAB" & COUNTY == "Solano"

replace providername = "VALLEY CONVALESCENT HOSPITAL BAKERSFIELD" if providername == "VALLEY CONVALESCENT HOSPITAL" & COUNTY == "Kern"

replace providername = "VALLEY CONVALESCENT HOSPITAL WATSONVILLE" if providername == "VALLEY CONVALESCENT HOSPITAL" & COUNTY == "Santa Cruz"
replace providername = "VICTORIA HEALTH CARE AND REHABILITATION CENTER" if providername == "VICTORIA HEALTH CARE CENTER" & COUNTY == "Orange"
replace providername = "VIEW HEIGHTS CONVALESCENT HOSPITAL" if providername == "VIEW HEIGHTS CONVALESCENT HOSP" & COUNTY == "Los Angeles"
replace providername = "VIEW PARK CONVALESCENT CENTER" if providername == "VIEW PARK CONVALESCENT HOSP" & COUNTY == "Los Angeles"
replace providername = "VILLA MARIN RETIREMENT RESIDENCES" if providername == "VILLA MARIN" & COUNTY == "Marin"
replace providername = "VILLA SCALABRINI SPECIAL CARE UNIT" if providername == "VILLA SCALABRINI SPECIAL CARE" & COUNTY == "Los Angeles"
replace providername = "VILLA VALENCIA HEALTH CARE CENTER" if providername == "VILLA VALENCIA" & COUNTY == "Orange"
replace providername = "WALNUT CREEK SKILLED NURSING AND REHABILITATION CENTER" if providername == "WALNUT CREEK SKILLED NURSING AND REHABILITATION CENT" & COUNTY == "Contra Costa"
replace providername = "WEST HAVEN HEALTH CARE CENTER" if providername == "WEST HAVEN HEALTH CARE" & COUNTY == "Los Angeles"
replace providername = "WINDSOR GARDENS HEALTH CARE CENTER OF THE VALLEY" if providername == "WINDSOR GARDENS HEALTH CARE OF THE VALLEY" & COUNTY == "Los Angeles"
replace providername = "WINDSOR MANOR REHABILITATION CENTER OF CONCORD" if providername == "WINDSOR MANOR REHABILITATION CENTER" & COUNTY == "Contra Costa"
replace providername = "WINDSOR TERRACE HEALTH CARE CENTER" if providername == "WINDSOR TERRACE HEALTH CARE" & COUNTY == "Los Angeles"
//remove lingering commas
replace providername = subinstr(providername, ",", "", 10)
//again leading spaces and trailing spaces on providername, merge variable
replace providername = strtrim(providername)

sort providername

//EXPORT AND SAVE AS .dta
save "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/lafd1219-sub-initial_forodp.dta" , replace //This .dta will be called by the analysis do file.

A.9. 7 do_merge

//MERGE DO FILE
//Tessa Ireton

```
//15 December 2020

clear all
set more off

use "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/lafd1219-sub-initial_forodp.dta" , clear

//MERGE county elderlypop dataset
	merge m:m COUNTY using "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/cc-est2019-agesex-06.dta"
	//many to many

	//drop counties that did not have any nursing homes, so they could not be matched
		drop if COUNTY == "Alpine" || COUNTY == "Mariposa" || COUNTY == "Modoc" || COUNTY == "Mono" || COUNTY == "Plumas" || COUNTY == "San Benito" || COUNTY == "Sierra" || COUNTY == "Trinity"

	//drop _merge
		drop _merge

//MERGE county_ethnicity dataset
	merge m:m COUNTY using "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/cc-est2019-alldata-06.dta"

	//drop counties that did not have any nursing homes, so they could not be matched
		drop if COUNTY == "Alpine" || COUNTY == "Mariposa" || COUNTY == "Modoc" || COUNTY == "Mono" || COUNTY == "Plumas" || COUNTY == "San Benito" || COUNTY == "Sierra" || COUNTY == "Trinity"

	//drop _merge
		drop _merge

//MERGE county median income/poverty dataset
	merge m:m COUNTY using "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/est18all.dta"
	//many to many

	//drop counties that did not have any nursing homes, so they could not be matched
		drop if COUNTY == "Alpine" || COUNTY == "Mariposa" || COUNTY == "Modoc" || COUNTY == "Mono" || COUNTY == "Plumas" || COUNTY == "San Benito" || COUNTY == "Sierra" || COUNTY == "Trinity"

	//drop _merge
		drop _merge

//DATE BOUNDS: 31may2020 to 22nov2020
```

```
//MERGE weekly county cases dataset
	joinby COUNTY using
"/Users/tjnireton/Desktop/IS_DataFiles/DTAs/statewide_cases.dta"
	//sort so I can see the different new weekly county cases for each nursing home
	sort providername endwk

	//drop county data that is not in the NH characteristics set; there are no nursing homes
in these counties.
	drop if COUNTY == "Alpine" || COUNTY == "Mariposa" || COUNTY == "Modoc"
|| COUNTY == "Mono" || COUNTY == "Plumas" || COUNTY == "San Benito" || COUNTY
== "Sierra" || COUNTY == "Trinity"

//MERGE weekly new NH cases
	merge m:m providername endwk COUNTY using
"/Users/tjnireton/Desktop/IS_DataFiles/DTAs/COVID-19_Nursing_Home_Dataset.dta"
	//many to many

sort providername endwk
order providername COUNTY endwk _merge provideraddress
//With the dataset ordered as such, I went through facilities and insured facility names were
the same between the providernames in both datasets. The commands to insure uniformity
appear in the Do_Clean_NH Characteristics and Do_clean _NH_cases do-files.

//Drop unmatched observations
	drop if _merge != 3

//drop unneeded variables for regressions
	drop _merge provideraddress FAC_NO providercity providerzipcode LIC_CAT

//TYPE_CNTRL variable: turn to number type and label "Investor Owned" if var == 1, 0 else
(not for-profit)
	replace TYPE_CNTRL = "1" if TYPE_CNTRL == "Investor Owned"
	replace TYPE_CNTRL = "0" if TYPE_CNTRL == "Not-for-Profit"
	replace TYPE_CNTRL = "0" if TYPE_CNTRL == "Church Related"
	replace TYPE_CNTRL = "0" if TYPE_CNTRL == "District"
	destring TYPE_CNTRL , generate(ForProfit_dum) force
		drop TYPE_CNTRL
	//make new control type label, 0 if not investor owned, 1 if investor owned
		label define orgtype_label 0 "not investor owned" 1 "investor owned"
		label values ForProfit_dum orgtype_label

//Assign each facility an observation/facility/NH number
	egen NHnum = group(providername)

//to keep some anonymity of facilities in the dataset, drop providername
	drop providername
```

```
//assign numbers to the counties
	egen cnty_num = group(COUNTY)
	order NHnum COUNTY cnty_num endwk

//EXPORT AND SAVE AS .dta
	save "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/combined_set.dta" , replace
//This .dta will be called by the analysis do file.
```

A.9.8 Do_analysis1

```
//Tessa Ireton 11 Jan 2020
//Analysis do_file

clear all
set more off
use "/Users/tjnireton/Desktop/IS_DataFiles/DTAs/combined_set.dta"

//install outreg2
	ssc install outreg2
	ssc install asdoc
	ssc install xttest3

//deal with possibility of duplicates ruining xtreg
	duplicates report NHnum endwk
	//duplicates list NHnum endwk
	duplicates tag NHnum endwk, gen(duplic)
	drop if duplic == 1
	drop duplic
xtset NHnum endwk
// drop observations that did not pass CMS' data quality assurance check
	drop if QApass_dum == 0
	drop QApass_dum

//Calculating weekly HHI
	//HHI formula again: Hirschman-Herfindahl Index. In this model, HHI_weekly is the
sum of the squared proportion of each nursing home’s portion of the total current bed
occupancy share in the county during the week. (one HHI per market)
	egen Total_wkly_MS = total(totoccup_beds) , by (cnty_num endwk)
	//Calculate each NH's weekly squared_MS:
	gen squared_wkly_MS = (totoccup_beds/Total_wkly_MS)^2
	//Calculate each Market's weekly HHI
	egen HHI_wkly = total(squared_wkly_MS) , by (COUNTY endwk)
	drop Total_wkly_MS squared_wkly_MS
```

//Summary Statistics

//time-invariant variables - non dummy
asdoc tabstat PRDHR_GNP_ppd PRDHR_RN_ppd PRDHR_LVN_ppd ForProfit_dum DAY_TOTL medi_services_revenue Medical_pro_services_revenue competition HHI cnty_elderlypercent cnty_totpop cnty_popAAper cnty_MedianHouseholdIncome2018 , stat(min max mean sd median) replace

//time-variant - non dummy
asdoc tabstat HHI_wkly totoccup_beds wklynewcases_1000res cnty_newcountconfirmed , stat(min max mean sd median) append

//time-variant - dummy
asdoc tabstat testable_dum RN_short_dum clinical_short_dum aides_short_dum otherstaff_short_dum n95_wksupp_dum surgicalmasks_wksupp_dum sanitizer_wksupp_dum , stat(mean) append

//print formatted summaries
outreg2 NHnum if endwk == d(31may2020) using "/Users/tjnireton/Desktop/IS_DataFiles/outregs/SummaryStats.doc", replace sum(detail)

//Estimations

//(1)FE with time-moving variables only. Does not control for in the surrounding area, nor proportion of revenue from medicaid. For the purpose of robustness test.

xtreg wklynewcases_1000res HHI_wkly cnty_newcountconfirmed testable_dum RN_short_dum n95_wksupp_dum totoccup_beds clinical_short_dum aides_short_dum otherstaff_short_dum surgicalmasks_wksupp_dum sanitizer_wksupp_dum , fe robust

outreg2 using "/Users/tjnireton/Desktop/IS_DataFiles/outregs/OutputTables1.doc" , replace ctitle(Estimation of New COVID Cases, Time-Variant Variables Only, FE) adjr2

//If conducting hausman test with non-robust SE regression.
//if "robust" removed
//estimates store fixed

//(1)RE with with time-moving variables only. Compare to FE results from above. Does not control for market structure nor demographics in the surrounding area, nor proportion of revenue from medicaid. For the purpose of robustness test.

xtreg wklynewcases_1000res HHI_wkly cnty_newcountconfirmed testable_dum RN_short_dum n95_wksupp_dum totoccup_beds clinical_short_dum aides_short_dum otherstaff_short_dum surgicalmasks_wksupp_dum sanitizer_wksupp_dum , re robust outreg2 using "/Users/tjnireton/Desktop/IS_DataFiles/outregs/OutputTables1.doc" , append ctitle(Estimation of New COVID Cases, Time-Variant Variables Only, RE) r2

//If conducting hausman test with non-robust SE regression.
//if "robust" removed
//estimates store random
//hausman fixed random

//THE ESTIMATIONS BELOW ADDRESS THE PRIMARY RESEARCH QUESTION

//(3)random effects. Main specification. COMPETITION
xtreg wklynewcases_1000res competition ForProfit_dum DAY_TOTL Medical_pro_services_revenue cnty_elderlypercent cnty_totpop cnty_popAAper cnty_MedianHouseholdIncome2018 cnty_newcountconfirmed testable_dum RN_short_dum clinical_short_dum aides_short_dum otherstaff_short_dum n95_wksupp_dum , re robust
outreg2 using "/Users/tjnireton/Desktop/IS_DataFiles/outregs/OutputTables3.doc" , replace ctitle(Estimation of New COVID Cases, Main Specification, RE) r2

//(3)random effects. Main specification. BUT WITH HHI
xtreg wklynewcases_1000res HHI ForProfit_dum DAY_TOTL Medical_pro_services_revenue cnty_elderlypercent cnty_totpop cnty_popAAper cnty_MedianHouseholdIncome2018 cnty_newcountconfirmed testable_dum RN_short_dum clinical_short_dum aides_short_dum otherstaff_short_dum n95_wksupp_dum , re robust
outreg2 using "/Users/tjnireton/Desktop/IS_DataFiles/outregs/OutputTables3.doc" , append ctitle(Estimation of New COVID Cases, HHI Rather Than Competition, RE) r2

//(3)random effects. Main specification, BUT WITH 2019 CARE LABOR HOURS INSTEAD OF WEEKLY CARE SHORTAGE DUMMIES
xtreg wklynewcases_1000res competition ForProfit_dum DAY_TOTL Medical_pro_services_revenue cnty_elderlypercent cnty_totpop cnty_popAAper cnty_MedianHouseholdIncome2018 cnty_newcountconfirmed testable_dum PRDHR_GNP_ppd PRDHR_RN_ppd PRDHR_LVN_ppd n95_wksupp_dum , re robust
outreg2 using "/Users/tjnireton/Desktop/IS_DataFiles/outregs/OutputTables3.doc" , append ctitle(Estimation of New COVID Cases, 2019 Labor Hours rather than Weekly Labor Shortage Dummies, RE) r2

//(3)random effects. Main specification, BUT WITH HHI_wkly
xtreg wklynewcases_1000res HHI_wkly ForProfit_dum DAY_TOTL Medical_pro_services_revenue cnty_elderlypercent cnty_totpop cnty_popAAper

cnty_MedianHouseholdIncome2018 cnty_newcountconfirmed testable_dum RN_short_dum clinical_short_dum aides_short_dum otherstaff_short_dum n95_wksupp_dum , re robust
outreg2 using "/Users/tjnireton/Desktop/IS_DataFiles/outregs/OutputTables3.doc" , append ctitle(Estimation of New COVID Cases, Weekly HHI, RE) r2

//NEW SPEC. THIS IS THE ONE HT I USED
//1 HT spec accounting for endog requirement, endogenous variables in endog()
xthtaylor wklynewcases_1000res HHI_wkly ForProfit_dum DAY_TOTL Medical_pro_services_revenue cnty_elderlypercent cnty_totpop cnty_popAAper cnty_MedianHouseholdIncome2018 cnty_newcountconfirmed testable_dum RN_short_dum clinical_short_dum aides_short_dum otherstaff_short_dum n95_wksupp_dum Medical_pro_services_revenue , endog(ForProfit_dum DAY_TOTL HHI_wkly RN_short_dum clinical_short_dum aides_short_dum otherstaff_short_dum n95_wksupp_dum) vce(robust)

outreg2 using "/Users/tjnireton/Desktop/IS_DataFiles/outregs/OutputTables10.doc" , replace ctitle(Estimation of New COVID Cases, weekly HHI is Independent Variable of Interest, Hausman-Taylor Estimation) r2

//test for the appropriateness of the following variables as instruments in HT
correlate ForProfit_dum DAY_TOTL cnty_newcountconfirmed testable_dum Medical_pro_services_revenue cnty_elderlypercent cnty_totpop cnty_popAAper cnty_MedianHouseholdIncome2018

//Looking for Heteroskedasticity and Multicollinearity
//Heteroskedasticity
//choose a specification for a generalized least squares estimation
xtgls wklynewcases_1000res HHI_wkly ForProfit_dum DAY_TOTL Medical_pro_services_revenue cnty_elderlypercent cnty_totpop cnty_popAAper cnty_MedianHouseholdIncome2018 cnty_newcountconfirmed testable_dum RN_short_dum clinical_short_dum aides_short_dum otherstaff_short_dum n95_wksupp_dum
outreg2 using "/Users/tjnireton/Desktop/IS_DataFiles/outregs/Heteroskedasticity.doc" , replace ctitle(Generalized Least Squares Estimation) r2

//run postestimation test for heteroskedasticty
xttest3
//Modified Wald test for groupwise heteroskedasticity in cross-sectional time-series FGLS regression model
//gives chi^2 of 3.7e+10
//prob>chi2 = 0.0000
//model assumes homoskedasticity. So a p value of 0 rejects null hypothesis of homoskedasticity.

```
//looking for Multicollinearity
	//look at pair-wise correllation
		asdoc pwcorr wklynewcases_1000res HHI_wkly ForProfit_dum DAY_TOTL Medical_pro_services_revenue cnty_elderlypercent cnty_totpop cnty_popAAper cnty_MedianHouseholdIncome2018 cnty_newcountconfirmed testable_dum RN_short_dum clinical_short_dum aides_short_dum otherstaff_short_dum n95_wksupp_dum, nonum replace
	//Run OLS regression to calculate VIF of Weekly HHI by hand.
		regress wklynewcases_1000res HHI_wkly ForProfit_dum DAY_TOTL Medical_pro_services_revenue cnty_elderlypercent cnty_totpop cnty_popAAper cnty_MedianHouseholdIncome2018 cnty_newcountconfirmed testable_dum RN_short_dum clinical_short_dum aides_short_dum otherstaff_short_dum n95_wksupp_dum
	outreg2 using "/Users/tjnireton/Desktop/IS_DataFiles/outregs/OLSforVIF.doc" , replace ctitle(OLS) r2
		//keep R2
		estat vif
```

Works Cited

Aadland, D. (2009). Hausman Taylor model. University of Wyoming College of Business Department of Economics and Finance. 1–7. https://www.aadecon.com/classes/econ5360/notes1.pdf .

Abrams, H. R., Loomer, L., Gandhi, A., & Grabowski, D. C. (2020). Characteristics of U.S. Nursing Homes with COVID-19 Cases. *Journal of the American Geriatrics Society*, *68*(8), 1653–1656. https://doi.org/10.1111/jgs.16661

Adams, Katie & Anderson, Maia (25 February 2021.) States ranked by percentage of COVID-19 vaccines administered: Feb. 25. *Becker's Hospital Review.* Retrieved February 25, 2021, from https://www.beckershospitalreview.com/public-health/states-ranked-by-percentage-of-covid-19-vaccines-administered.html

Albouy, D. (2005). The neoclassical firm. University of California Berkeley. https://eml.berkeley.edu/~webfac/card/e101a_s05/neoclassicalfirm.pdf

Barnett, M. L., & Grabowski, D. C. (2020). Nursing Homes Are Ground Zero for COVID-19 Pandemic. *JAMA Health Forum*, *1*(3), e200369. https://doi.org/10.1001/jamahealthforum.2020.0369

Bishop, C. E. (1988). Competition in the Market for Nursing Home Care. *Journal of Health Politics, Policy and Law*, *13*(2), 341–360. https://doi.org/10.1215/03616878-13-2-341

Bowers, B. J., Fibich, B., & Jacobson, N. (2001). Care-as-Service, Care-as-Relating, Care-as-Comfort: Understanding Nursing Home Residents' Definitions of Quality. *The Gerontologist*, *41*(4), 539–545. https://doi.org/10.1093/geront/41.4.539

Bui, D. P. (2020). Association Between CMS Quality Ratings and COVID-19 Outbreaks in Nursing Homes—West Virginia, March 17–June 11, 2020. *MMWR. Morbidity and Mortality Weekly Report*, *69*. https://doi.org/10.15585/mmwr.mm6937a5

California COVID-19 Response (2021). *COVID-19 Cases: California Open Data*. (25 January 2021). Retrieved 7 December 2020, from https://data.ca.gov/dataset/covid-19-cases/resource/926fd08f-cc91-4828-af38-bd45de97f8c3

California Health and Human Services Open Data Portal. (3 May 2018). *Long-term Care Facility Integrated Disclosure and Medi-Cal Cost Report Data & Pivot Tables, Documentation: Long-term Care Facility Annual Financial Data.* (3 May 2018). Retrieved January 25, 2021, from https://data.chhs.ca.gov/dataset/long-term-care-facility-disclosure-report-data/resource/5d8c2a00-88b4-43cb-8f16-b4a72e35212c

California Health and Human Services Open Data Portal. (22 January 2019). *Long-term Care Facility Integrated Disclosure and Medi-Cal Cost Report Data & Pivot Tables, Documentation: Complete Data Set, Long-Term Care Annual Financial Data.* Retrieved January 25, 2021, from https://data.chhs.ca.gov/dataset/long-term-care-facility-disclosure-report-data/resource/857957ed-9059-4033-98e6-0e8d47d6fe50

California Health and Human Services (CHHS) Open Data Portal. (4 December 2020). *Long-term Care Facility Integrated Disclosure and Medi-Cal Cost Report Data & Pivot Tables, 2019 Selected File (As Submitted): Long-Term Care Annual Financial Data (December 2020).* Retrieved January 25, 2021, from https://data.chhs.ca.gov/dataset/long-term-care-facility-disclosure-report-data/resource/9f1e4a28-37d7-419e-b304-24c6b79238d6

Campbell, S. M., Roland, M. o., & Buetow, S. A. (2000). Defining Quality of Care. *Social Science & Medicine*, *51*, 1611–1625. file:///Users/tjnireton/Downloads/1-s2.0-S0277953600000575-main.pdf.

CHHS California Department of Public Health. (22 September 2020). *Updated COVID-19 Testing Guidance*. Retrieved January 31, 2021, from https://www.cdph.ca.gov/Programs/CID/DCDC/Pages/COVID-19/Updated-COVID-19-Testing-Guidance.aspx

Conlen, M., Mervosh, S., & Ivory, D. (2021, February 25). Nursing Homes, Once Hotspots, Far Outpace U.S. in Covid Declines. The New York Times. https://www.nytimes.com/interactive/2021/02/25/us/nursing-home-covid-vaccine.html

Cowles Research Group (2020). CASPER & OSCAR. Retrieved November 24, 2020, from https://www.longtermcareinfo.com/data/casper-and-oscar.php

Chou, S.-Y. (2002). Asymmetric information, ownership and quality of care: An empirical analysis of nursing homes. *Journal of Health Economics*, *21*(2), 293–311. https://doi.org/10.1016/S0167-6296(01)00123-0

Chesteen, S., Helgheim, B., Randall, T., & Wardell, D. (2005). Comparing quality of care in non-profit and for-profit nursing homes: A process perspective. *Journal of Operations Management*, *23*(2), 229–242.

CHHS. (2020, February 20). Licensed and Certified Healthcare Facility Listing—Lookup Table-Licensed and Certified Healthcare Facility Listing Lookup Table—California Health and Human Services Open Data Portal. CHHS. https://data.chhs.ca.gov/dataset/healthcare-facility-locations/resource/1e29e3d3-2e35-48e5-a903-de6cfb2eaf2a?inner_span%3DTrue

CHHS. (2020). Long-term Care Facility Integrated Disclosure and Medi-Cal Cost Report Data & Pivot Tables—2018 Selected File—Long-Term Care Annual Financial Data—(Sept 2019 Extract). California Health and Human Services Open Data Portal. Retrieved December 1, 2020, from https://data.chhs.ca.gov/dataset/long-term-care-facility-disclosure-report-data/resource/0856b659-b8dc-4f2b-a2a6-8417c3ab0935.

CMS. (2020). COVID-19 Archived Data. https://data.cms.gov/stories/s/COVID-19-Nursing-Home-Data-Archived-Datasets/dyvx-s2qr/. Last updated data table was uploaded November 30, 2020 https://data.cms.gov/Special-Programs-Initiatives-COVID-19-Nursing-Home/COVID-19-Nursing-Home-Dataset/s2uc-8wxp.

CMS. (November 2020). *COVID-19 Nursing Home Data—Archived Datasets: Week ending in 11/22/2020.* Retrieved from https://data.cms.gov/stories/s/COVID-19-Nursing-Home-Data-Archived-Datasets/dyvx-s2qr/

CMS. (2021). *Find Healthcare Providers: Compare Care Near You*. Medicare.Gov. https://www.medicare.gov/care-compare/?providerType=NursingHome&redirect=true#search

CMS. (January 2021). *COVID-19 Nursing Home Dataset: Data Dictionary and Dataset.CMS.gov*. (2020). Retrieved from https://data.cms.gov/Special-Programs-Initiatives-COVID-19-Nursing-Home/COVID-19-Nursing-Home-Dataset/s2uc-8wxp

COVID-19 Nursing Home Data. (2020). Centers for Medicare and Medicaid Services. Retrieved October 19, 2020, from https://data.cms.gov/stories/s/COVID-19-Nursing-Home-Data/bkwz-xpvg/

COVID-19 Response News Alert: CMS Issues Key Protective Mask Guidance for Healthcare Workers. CMS. (2020, March 10). https://www.cms.gov/newsroom/press-

releases/covid-19-response-news-alert-cms-issues-key-protective-mask-guidance-healthcare-workers
Davis, M. A. (1993). Nursing home ownership revisited: Market, cost and quality relationships. *Medical Care, 31*(11), 1062–1068. https://www.jstor.org/stable/3766304?seq=5#metadata_info_tab_contents.
Director, Quality, Safety & Oversight Group. (9 September 2020). *Center for Clinical Standards and Quality/Quality, Safety & Oversight Group Memorandum*. Department of Health & Human Services, Centers for Medicare & Medicaid Services. https://www.cms.gov/files/document/qso-20-14-nh-revised.pdf
Domi, M., Leitson, D., Gifford, D., & Sreenivas, K. (2021). Nursing Home Resident and Staff COVID-19 Cases After the First Vaccination Clinic. The Center for Health Policy Evaluation in Long-Term Care. https://www.ahcancal.org/Data-and-Research/Center-for-HPE/Documents/CHPE-Report-Vaccine-Effectiveness-Feb2021.pdf
Fallon, A., Dukelow, T., Kennelly, S. P., & O'Neill, D. (2020). COVID-19 in nursing homes. *QJM: An International Journal of Medicine*, *113*(6), 391–392. https://doi.org/10.1093/qjmed/hcaa136
FastStats—Nursing Home Care. (2020, August 10). National Center for Health Statistics, Centers for Disease Control and Prevention. https://www.cdc.gov/nchs/fastats/nursing-home-care.htm
Find and compare Nursing Homes: Nursing Home Compare. (2020). CMS Retrieved October 11, 2020, from https://www.medicare.gov/nursinghomecompare/search.html.
Friedberg, L., Hou, W., Sun, W., & Webb, A. (2014). Long Term Care: How Big a Risk? *Center for Retirement Research*, *14–18*(November 2014).
General Accounting Office. (1999). *Nursing Homes: Additional Steps Needed to Strengthen Enforcement of Federal Quality Standards*. General Accounting Office. https://www.gao.gov/assets/230/227015.pdf
Centers for Medicare and Medicaid Services (CMS). (2020). Nursing Home COVID-19 Data Quality: Frequently Asked Questions. file:///Users/tjnireton/Downloads/NH_COVID-19_data_FAQ.pdf
Gitmez, A. (2017). 14.770-Fall 2017: Recitation 6 Notes. MIT Economics Department. http://economics.mit.edu/files/15905
Gharpure, R., et al. (1 February 2021). Early COVID-19 First-Dose Vaccination Coverage Among Residents and Staff Members of Skilled Nursing Facilities Participating in the Pharmacy Partnership for Long-Term Care Program—United States, December 2020–January 2021. MMWR. Morbidity and Mortality Weekly Report, 70. https://doi.org/10.15585/mmwr.mm7005e2
Glied, S., & Smith, P. C. (Eds.). (2011). *The Oxford Handbook of Health Economics*. Oxford University Press.
Government Accountability Office. (2015). *CMS Should Continue to Improve Data and Oversight*. Government Accountability Office. https://www.gao.gov/assets/680/673480.pdf.
Government Accountability Office. (2018). *Continued Improvements Needed in CMS' Data and Oversight*. Government Accountability Office. https://www.gao.gov/assets/700/694324.pdf.

General Accounting Office. (1999). *Nursing Homes: Additional Steps Needed to Strengthen Enforcement of Federal Quality Standards*. General Accounting Office. https://www.gao.gov/assets/230/227015.pdf

Grabowski, D. C., & Castle, N. G. (2004). Nursing Homes with Persistent High and Low Quality. *Medical Care Research and Review*, *61*(1), 89–115. https://doi.org/10.1177/1077558703260122

Grabowski, D. C., & Hirth, R. A. (2003). Competitive spillovers across non-profit and for-profit nursing homes. *Journal of Health Economics*, *22*(1), 1–22. https://www.sciencedirect.com/science/article/pii/S0167629601001230.

Grabowski, D. C., & Norton, E. C. (2012). Nursing Home Quality of Care. In A. M. Jones (Ed.), *The Elgar Companion to Health Economics* (Second). Edward Elgar.

Grabowski, D. C., & Town, R. J. (2011). Does Information Matter? Competition, Quality, and the Impact of Nursing Home Report Cards. *Health Services Research*, *46*(6pt1), 1698–1719. https://doi.org/10.1111/j.1475-6773.2011.01298.x

Gulley, O. D., & Santerre, R. E. (2004). Market Power of California Nursing Homes. University of Connecticut and Bentley College.

Hackmann, M. 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. https://doi.org/10.1257/aer.20151057

Harrington, C., Ross, L., Chapman, S., Halifax, E., Spurlock, B., & Bakerjian, D. (2020). Nurse Staffing and Coronavirus Infections in California Nursing Homes. *Policy, Politics, & Nursing Practice*, *21*(3), 174–186. https://doi.org/10.1177/1527154420938707

He, M., Li, Y., & Fang, F. (July 2020). *Is There a Link between Nursing Home Reported Quality and COVID-19 Cases? Evidence from California Skilled Nursing Facilities | Elsevier Enhanced Reader*. JAMDA. https://doi.org/10.1016/j.jamda.2020.06.016

Hirth, R. A. (1999). Consumer information and competition between nonprofit and for-profit nursing homes. *Journal of Health Economics*, *18*(2), 219–240. https://www.sciencedirect.com/science/article/pii/S0167629698000356.

Huang, H. (2015). Introduction to Game Theory Lecture 5: Repeated Games. Duke University and the U.S. Intercollegiate Social Science Research Alliance. https://sites.duke.edu/niou/files/2011/05/Lecture-5-Repeated-Games.pdf.

John E., D. (2020). *Infection Control Deficiencies Were Widespread and Persistent in Nursing Homes Prior to COVID-19 Pandemic*. U.S. Government Accountability Office. https://www.gao.gov/assets/710/707069.pdf

Khorasani, E., Karimi, S., Keyvanara, M., Etemadi, M., & Khorasani, F. (2015). Identification of the Healthcare Services with Potential Induced Demand. *International Journal of Hospital Research*, *4*(2), 47–54.

Lin, H. (2015). Quality Choice and Market Structure: A Dynamic Analysis of Nursing Home Oligopolies. *International Economic Review*, *56*(4), 1261–1290. https://doi.org/10.1111/iere.12137

Linn, M. W., Gurel, L., & Linn, B. S. (1977). Patient outcome as a measure of quality of nursing home care. *American Journal of Public Health*, *67*(4), 337–344. https://doi.org/10.2105/AJPH.67.4.337

Lu, S. F., Serfes, K., & Wu, B. (2017). Does Competition Improve Quality? The Case of Nursing Homes Where Public and Private Payers Coexist. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2963566

Lu, W. (2014). Economic Determinants of Quality of Care in Nursing Homes. Wayne State University. 101.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). Chapter 5: production. In *microeconomic theory* (pp. 127–166). Oxford University Press.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). Chapter 12: market power. In *microeconomic theory* (pp. 383–435). Oxford University Press.

McMichael, T. M., Currie, D. W., Clark, S., Pogosjans, S., Kay, M., Schwartz, N. G., Lewis, J., Baer, A., Kawakami, V., Lukoff, M. D., Ferro, J., Brostrom-Smith, C., Rea, T. D., Sayre, M. R., Riedo, F. X., Russell, D., Hiatt, B., Montgomery, P., Rao, A. K., … Duchin, J. S. (2020). Epidemiology of Covid-19 in a Long-Term Care Facility in King County, Washington. *New England Journal of Medicine*, *382*(21), 2005–2011. https://doi.org/10.1056/NEJMoa2005412

Mollica, R. (2009). *Home and Community-Based Long-Term Care: Recommendations to Improve Access for Californians*. California Health and Human Services Agency. http://www.canhr.org/reports/HomeCommunityBasedCare2009.pdf

Machado, M. (2012). 4.1. Product Differentiation. Industrial Organization at Universidad Carlos III. http://www.eco.uc3m.es/~mmachado/Teaching/OI-I-MEI/slides/4.1.Product_differentiation.pdf.

Muck, J. (2018). Econometrics of Panel Data. Warsaw School of Economics, Department of Quantitative Economics, 21–29. http://web.sgh.waw.pl/~jmuck/EoPD/Meeting5_2018.pdf .

Mukamel, D. B., & Spector, W. D. (2003). Quality Report Cards and Nursing Home Quality. *The Gerontologist*, *43*(suppl_2), 58–66. https://doi.org/10.1093/geront/43.suppl_2.58

Munroe, D. J. (1990). The influence of registered nurse staffing on the quality of nursing home care. *Research in Nursing & Health*, *13*(4), 263–270. https://doi.org/10.1002/nur.4770130409.

National Center for Health Statistics. Centers for Disease Control National Study of Long-Term Care Providers. Appendix I, National Study of Long-Term Care Providers (NSLTCP). Data Finder—Health, United States—Products. (2020, March 31). https://www.cdc.gov/nchs/hus/contents2018.htm. Figure 17. https://www.cdc.gov/nchs/data/hus/2018/fig17.pdf

Non-Profit vs. For-Profit Nursing Homes: Is there a Difference in Care? (2012, March 15). Center for Medicare Advocacy. https://medicareadvocacy.org/non-profit-vs-for-profit-nursing-homes-is-there-a-difference-in-care/

"Nursing Home Quality Initiative". (2020). CMS. Retrieved November 11, 2020, from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits

Ouslander, J. G., & Grabowski, D. C. (2020). COVID-19 in Nursing Homes: Calming the Perfect Storm. *Journal of the American Geriatrics Society*, *68*(10), 2153–2162. https://doi.org/10.1111/jgs.16784

Payson, S. (1994). The Meaning of Quality. In *Quality Measurement in Economics: New Perspectives of the Evolution of Goods and Services* (pp. 1–33). Edward Elgar.

Perloff, J. M. (2017). chapter 6: firms and production. In *microeconomics: theory and applications with calculus* (Fourth). Pearson.
Perloff, J. M. (2017). Chapter 14: oligopoly and monopolistic competition. In *microeconomics: theory and applications with calculus* (Fourth). Pearson.
Residential Facilities, Assisted Living, and Nursing Homes. (2020) National Institute on Aging, National Institutes of Health. Retrieved October 15, 2020, from http://www.nia.nih.gov/health/residential-facilities-assisted-living-and-nursing-homes
Scanlon, W. J. (1980). A theory of the nursing home market. *Inquiry*, *17*(1), 25–41. JSTOR.
See R. T. G. (2005). Competition with Horizontal and Vertical Differentiation. *National University of Singapore Department of Economics*, 200. https://core.ac.uk/download/pdf/48628726.pdf.
Shah, M., Sachdeva, M., & Dodiuk-Gad, R. P. (2020). COVID-19 and racial disparities. *Journal of the American Academy of Dermatology*, *83*(1), e35. https://doi.org/10.1016/j.jaad.2020.04.046
Skilled Nursing Facilities—Care Options. (2020). California Department of Aging. https://www.aging.ca.gov/Care_Options/Skilled_Nursing_Facilities/
Studenmund, A. H. (2017). *Using Econometrics: A Practical Guide* (Seventh). Pearson.
Torres-Reyna, O. (2007). *Panel Data Analysis: Fixed and Random Effects using Stata*. Princeton University Data & Statistical Services. http://www.princeton.edu/~otorres/Panel101.pdf
Tulchinsky, T., & Varavikova, E. (2014). Family Health. In *The New Public Health* (pp. 372–377). Elsevier.
US Census Bureau. (2019). County Population by Characteristics: 2010-2019. The United States Census Bureau. Retrieved from https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-detail.html
US Census Bureau. (December 2019). *SAIPE State and County Estimates for 2018*. The United States Census Bureau. Retrieved January 25, 2021, from https://www.census.gov/data/datasets/2018/demo/saipe/2018-state-and-county.html
US Census Bureau. (2020). *Methodology for the United States Population Estimates: Vintage 2019*. https://www2.census.gov/programs-surveys/popest/technical-documentation/methodology/2010-2019/natstcopr-methv2.pdf
US Census Bureau. (June 2020). *Data Dictionary: Annual County and Puerto Rico Municipio Resident Population Estimates by Selected Age Groups and Sex: April 1, 2010 to July 1, 2019.* https://www2.census.gov/programs-surveys/popest/technical-documentation/file-layouts/2010-2019/cc-est2019-agesex.pdf
Varian, H. R. (1993). Chapter 26: Oligopoly. In *Intermediate Microeconomics: A Modern Approach* (Third, pp. 447–468). W. W. Norton & Company, Inc.
Wang, J., Yang, W., Pan, L., Ji, J. S., Shen, J., Zhao, K., Ying, B., Wang, X., Zhang, L., Wang, L., & Shi, X. (2020, November). *Prevention and control of COVID-19 in nursing homes, orphanages, and prisons | Elsevier Enhanced Reader*. Environmental Pollution. https://doi.org/10.1016/j.envpol.2020.115161
White, E. M., Kosar, C. M., Feifer, R. A., Blackman, C., Gravenstein, S., Ouslander, J., & Mor, V. (2020). Variation in SARS-CoV-2 Prevalence in U.S. Skilled Nursing Facilities. *Journal of the American Geriatrics Society*, *68*(10), 2167–2173. https://doi.org/10.1111/jgs.16752

xthtaylor: Hausman-Taylor estimator for error-components models. (2021). from The Stata Manual. STATA.com. https://www.stata.com/manuals13/xtxthtaylor.pdf

Zinn, J. S. (1994). Market Competition and the Quality of Nursing Home Care. *Journal of Health Politics, Policy and Law*, *19*(3), 555–582. https://doi.org/10.1215/03616878-19-3-555.

More Than One-Third of U.S. Coronavirus Deaths Are Linked to Nursing Homes. (2020, June 27). *The New York Times*. https://www.nytimes.com/interactive/2020/us/coronavirus-nursing-homes.html