# Vertical Integration in the U.S. Health Care Market: An Empirical Analysis of Hospital-Insurer Consolidation

Ayse Sera Diebel*

**Abstract**

This paper examines the effects of vertical integration between hospitals and insurers on market outcomes and welfare. To assess the impact on market outcomes, I estimate reduced-form regressions that demonstrate vertically-integrated entities engage in market foreclosure (both upstream and downstream), and offer lower-premium higher-quality health plans. To determine the overall impact on welfare, I estimate a structural model, use the estimates in a policy experiment that prohibits vertical integration, and calculate the change in consumer and producer surplus. The structural model captures consumer behavior by estimating discrete choice models of hospital and insurer demand, while it captures the hospital-insurer contracting process by estimating a bargaining model that reveals consolidation increases vertically-integrated hospitals' bargaining power. My findings illustrate that vertical integration acts as an entry barrier to the downstream market due to the cost advantages vertically-integrated entities achieve through upstream foreclosure. Additionally, I find that vertical integration benefits consumers at the expense of the producers. Banning exclusionary restrictions gets rid of market foreclosure and gives access to vertically-integrated entities. As a result, insurers offer wider hospital networks but increase their premiums, which harms consumers. Producers, on the other hand, are better off in the absence of vertical integration as many hospitals and insurers enjoy higher profits driven by higher market shares and increased premiums.

JEL: I11, I13, L13, L14, L42

Keywords: vertical integration, foreclosure, bargaining, hospitals, health insurance

*The Graduate Center at the City University of New York, Ph.D. Program in Economics, 365 Fifth Avenue, New York, NY 10016, ayagli@gradcenter.cuny.edu. I am grateful to my advisors David A. Jaeger, Michael Grossman, and Partha Deb for their advice, guidance, and support. I acknowledge financial support from CUNY Doctoral Student Research Grant.

## 1 Introduction

The health care industry in the U.S. is becoming more consolidated. The number of highly-concentrated markets has been on the rise, which raises anti-trust concerns. This trend is expected to continue since market environments are becoming more challenging for small providers and insurers (Ginsburg (2016)). The empirical literature to date has focused on horizontal consolidation in the form of hospital mergers and insurer mergers, however the effects of vertical integration in the market for health care remain to be discovered. Aiming to fill this void, in this paper, I empirically investigate the impact of hospital-insurer consolidation on market outcomes and welfare using newly-available data.

Theory does not provide a definitive answer as to whether vertical integration is welfare-improving or not (Tirole (1988), Perry (1989), Joskow (2008)). Vertical integration leads to efficiency gains through improved coordination and reduced costs; but it also raises anti-competitive concerns due to increased market foreclosure, market power, and barriers to entry. The overall impact on welfare depends on which effect is dominant. In the case of hospital-insurer consolidation, several factors contribute to gains in efficiency. Increased quality stems from clinical integration of hospitals and insurers. Better coordination between the two entities facilitates tracking patients' medical history and enables the health system to take preventive measures accordingly. Cost savings are realized through three main channels. Vertically-integrated (VI) entities have reduced transaction costs as they no longer face contracting or negotiation costs. Cost of providing care is also reduced as improved coordination of care prevents wasteful procedures. Finally, elimination of double marginalization contributes to cost savings of VI insurers as they no longer pay a markup, and instead access the input at its marginal cost.

On the other hand, hospital-insurer consolidation raises anti-competitive concerns. VI entities aiming to gain monopoly power have an incentive to engage in market foreclosure by raising their rivals' costs and denying them access to their hospitals and health plans. This results in increased market power of VI entities as well as increased hospital prices and premiums in the market, leading to a decline in welfare. Presence of VI entities in the market also acts as a barrier to entry in the downstream market. Unlike other insurers, VI insurers have cost advantages as they have access to their own hospitals at lower prices and as they are able to avoid the sunk costs associated with building a health plan (such as building a provider network, finding compatible billing systems with hospitals etc.). As a result, they are able to offer lower premiums. This deters entry of insurers who do not have the same cost advantages, harming insurer competition.

Empirical papers that analyze the impacts of vertical integration also find conflicting evidence. Chipty (2001) finds vertical integration leads to market foreclosure as well as reduced prices; and concludes the efficiency-inducing effects dominate. Lee (2013) shows vertical integration harms consumers as it limits choices available to them. Ho (2009a) predicts that the entry of VI insurers into the market increases consumer welfare, while Crawford et al. (2017) assert that the impact on welfare depends on the extent of exclusion and foreclosure. This paper contributes to the empirical literature by investigating vertical integration in the health care industry and analyzing its relationship to foreclosure, prices, quality, bargaining power, and welfare.

There are two main components of my paper. First, I illustrate that vertical integration is associated with market foreclosure (both upstream and downstream), lower premiums, and higher quality by conducting reduced-form analysis. Second, I show that vertical integration benefits consumers but harms majority of producers by estimating a full structural model. In doing so, I combine estimates from discrete choice models of hospital and insurer demand with hospital cost function estimates, and use them to estimate a bargaining model. These estimated models allow me to capture the changes for all participants in the market when simulating the counterfactual scenario that removes vertical integration from the market. My findings indicate that removal of vertical integration from the market would decrease consumer surplus by $9.5 billion and increase producer surplus by $44 billion a year. I also find evidence that VI hospitals have more bargaining power in the negotiation process, and there are entry barriers in the downstream market.

The passage of the Patient Protection and Affordable Care Act (ACA) in 2010 encouraged consolidation in health care, and as a result, the growth rate of number of VI plans has increased from 1% to 3%.[1] Three regulations passed under the ACA, in particular, make the study of hospital-insurer consolidation valuable. Two pertain to incentivization of formation of VI entities, while one emphasizes the relevance for the future given the current policy environment. First, the ACA provided incentives to both providers and insurers in the form of bonus payments with the purpose of improving quality of care and reducing costs for Medicare beneficiaries. These incentivizations led to formation of provider groups called Accountable Care Organizations (ACOs). Many ACOs vertically integrated with insurers that offer private Medicare plans (called Medicare Advantage (MA) plans) to mutually benefit from the current regulation. Bonus payments are given

[1] Figures represent author's calculations from Atlantic Information Services (AIS) Directory of Health Plans and Weiss Ratings Guide to Health Insurers. Khanna et al. (2016) report the same growth rate using data collected by McKinsey & Company.

to ACOs and MA plans based on quality measures. To the extent that integration achieves higher quality, providers and insurers have an incentive to integrate to guarantee the receipt of the bonus payments. Moreover, ACOs face penalties if their spending is above a certain threshold level, which puts them at financial risk. Therefore, provider groups that merge with insurers are at an advantage (Frakt et al. (2013)). While vertical integration of this form took place to accommodate the migration to value-based payment systems under Medicare, these VI entities also started offering plans in the commercial market. The expansion of these firms into the private market likely took place because firms already absorbed the sunk cost of integrating. Another ACA item that incentivized vertical integration is the medical loss ratio (MLR) regulation that requires insurers to spend at least 85% of their premium revenues on medical claims and quality improvement for large-group fully-insured enrollees.[2] This made meeting the MLR criterion easier for VI insurers as they could make internal transfer payments to their health care facilities and label profits as costs (Dafny (2015)).

The last relevant policy change pertains to the formation of insurance exchange markets under the ACA. In line with the above incentivizations, VI plans constituted 52% of all entities in the market by 2016, with commercial plans -including the ones offered on the exchanges- being their biggest growth area.[3] By 2016, 22% of the population enrolled in health insurance through public exchanges were enrollees of VI plans.[4] Implications of their success are notably important given the recent anti-competitive concerns about exchanges. The number of VI insurers as a share of all the entities operating in exchanges nationwide has risen from 23.6% in 2014 to 25.3% in 2017. Moreover, VI entities represented 20% of all the entities operating in monopoly or duopoly exchange markets in 2014, while this number increased to 24% in 2017.[5] These figures strengthen the relevance of my finding that foreclosure is present in the market since vertically integrated entities engage in market foreclosure with the purpose of gaining more market power (Tirole (1988)). Lastly, even though the traditional insurers have been struggling in exchanges, VI insurers managed to survive in these markets. Among 34 VI insurers that entered exchanges in 2014, 20 were active in the marketplace every year until 2017, implying a 4-year survival rate of 59%. Their prominence in concentrated markets combined with their high survival rate call into question whether consumer welfare will be harmed if VI entities become the dominant players in their markets.

[2]For small-group and individual enrollees, this minimum is 80%. Self-insured plans are not subject to this regulation.

[3]See https://aishealth.com/node/47256.

[4]Enrollment in VI plans constituted a substantial portion of the overall insured population in many states. In 2015, around 36 million people were enrolled in integrated plans across all lines of business nationwide. See Figure 1 and Figure 2 in the appendix. Source of all calculations and figures is AIS Directory of Health Plans.

[5]All figures represent author's calculations from AIS and Centers for Medicare and Medicaid Services (CMS) data. The figures for 2017 represent the lower bounds as I only observe VI insurers until the beginning of 2016, and hence do not consider the vertical consolidations occurred between 2016 and 2017.

In order to assess the overall impact on welfare, I separately model every component of the market and use these models to predict the changes in actions of consumers, hospitals, and insurers under counterfactual simulations. I begin by using static discrete choice setup to model patients' choices for hospitals. In this setting, individuals maximize utility when choosing hospitals taking into account hospital characteristics as well as their own characteristics such as sex, age, diagnosis. Hospital demand model is estimated using data on the universe of discharges from every hospital in seven states. Identification in this model comes from variation in patients' choice sets of hospitals across markets. Next, using the parameter estimates from the hospital demand model, I calculate the expected utility the average patient gets from a network of hospitals offered by an insurer. Then, I use this expected utility measure as an input in my insurer demand estimation. Insurer demand specification follows Berry, Levinsohn, and Pakes (1995) and takes into account unobserved insurer characteristics and heterogeneity in individual preferences towards certain insurer characteristics. This model is identified by the variation in insurer characteristics within a market as well as the variation in choice sets of individuals across markets. Ultimately, I estimate a bargaining model that combines demand side estimates with the results from hospital cost function estimation. Hospital-insurer bargaining is modeled in a Nash bargaining framework developed by Brooks et al. (1997) and later extended by Lewis and Pflum (2015).

Following the estimation of the full structural model, I analyze the counterfactual environment where exclusionary vertical restraints are removed from the market. Under this counterfactual scenario, I remove vertical integration from the market altogether. As a result, previously-VI entities become separate hospitals and insurers. This changes consumer demand for hospitals, leads to formation of new insurer networks through the bargaining model, and results in new premiums. I find that removal of vertical integration from the market lowers consumer welfare as it generates higher premiums on average due to broader insurer networks. While larger hospital choice sets benefit consumers, this increase in utility is offset by the disutility from higher premiums. The overall producer surplus in the market goes up, although this effect is a combination of some hospitals and insurers losing profits while others are gaining. The main loss is endured by previously-VI insurers who no longer can enjoy the dedicated market share they had through exclusivity, and by previously-non-VI hospitals who now have more competitors in the market. Previously-non-VI insurers mostly gain as they start including previously-VI hospitals in their networks. Majority of these hospitals also gain as they increase their shares. Finally, I also analyze the impact of removal of vertical integration from the incumbent on the entrant in the downstream market.[6] I find that the entrant is better off as it now also has the same cost advantages as the VI entities. This result points to the presence of entry barriers in the downstream market, which harms downstream competition.

The rest of the paper is organized as follows. Section 2 discusses industry background and related literature. Section 3 describes the data. Section 4 presents results from the reduced-form analysis. Section 5 lays out the structural model. Section 6 discusses estimation details and results. Section 7 simulates the counterfactual scenario and analyzes the change in welfare. Section 8 concludes.

## 2 Industry Background and Related Literature

As in every vertically-separated market, there are three main players in the market for health care: providers (hospitals and physicians) who are the upstream entities that produce the services, insurers who are the downstream entities that bundle the upstream products, and consumers who purchase these bundles (health plans) to have access to the network of providers offered by the insurer. The relationship between these players is vertical in nature, implying every player belongs in a different tier. Given this structure, the market for health care is prone to two kinds of consolidation: horizontal and vertical. Horizontal consolidation occurs when firms in the same tier go through mergers and acquisitions. Literature to date mostly focused on this kind of consolidation in the form of hospital mergers and insurer mergers. Previous research on hospital mergers finds hospital consolidation has no impact on quality of care (Ho and Hamilton (2000), Capps (2005)) but leads to increased hospital prices (Gaynor and Town (2012)).[7] Similarly, papers on insurer mergers point to the strong relationship between insurer consolidation and higher premiums. Dafny et al. (2012) study Aetna-Prudential merger and find premiums are rising more quickly in more concentrated insurance markets. Trish and Herring (2015) also find premiums are higher in insurance markets with higher levels of concentration. Lastly, some papers study the impact of hospital-physician integration on market outcomes.[8] Among these, Baker et al. (2014) find hospital-physician integration results in reduced health spending and increased quality of care, but increases hospital prices and providers' market power.

Research on vertical consolidation in the health care market, on the other hand, has been somewhat limited,

[6] In the data, all VI entities are incumbents and there is only one entrant to the insurer market which is non-VI.

[7] Gaynor and Town (2012) provide a survey of nine studies that analyze the relationship between hospital mergers and prices.

[8] Most of the literature refers to hospital-physician integration as vertical integration even though integration occurs within the same tier (among providers).

mainly due to unavailability of data. By definition, vertical integration occurs when a firm starts offering services that it was not traditionally offering by acquisition of a player that belongs in another tier. In the market for health care, this happens when providers and insurers merge or go through acquisitions to offer a health plan.[9] These plans, often called provider-sponsored health plans (PSHPs), are characterized by low premiums and narrow networks where enrollees can visit only the member hospitals of the parent organization or the few other hospitals it contracted with. The most commonly known provider-sponsored plans are Kaiser Permanente products where enrollees can visit only in-network Kaiser facilities while paying low premiums. Frakt et al. (2013) analyze PSHPs in the Medicare Advantage market prior to the passage of the ACA and find these plans charge higher premiums, controlling for quality. They also find such plans have higher quality ratings. Ho (2009a) studies hospital-insurer integration and finds that entry of PSHPs into new markets increases social surplus. I contribute to this strand of literature by assessing the welfare impacts of the presence of vertical integration in the market for health care. Mine is also the first paper to reveal that vertically-integrated health systems engage in both upstream and downstream foreclosure.

While literature from health care industry is limited, a number of empirical papers study vertical integration and exclusive contracting in other industries (Lee (2013) in videogames, Crawford et al. (2017) and Chipty (2001) in cable television, Vita (2000) in gasoline, Asker (2016) in beer). These empirical papers aim to assess the impact of vertical integration on welfare as the theory does not provide a definitive answer. Theoretical literature (Tirole (1988), Perry (1989), Joskow (2008)) emphasizes firms have incentives to vertically integrate due to both the efficiency gains and anti-competitive motives. Efficiency gains occur through elimination of double-marginalization[10] (Tirole (1988)), elimination of transaction costs[11], and alignment of interests[12] (Williamson (1971), Grossman and Hart (1986)). The anti-competitive motives are present as firms can gain market power by foreclosing their rivals from the market through increased entry costs and input costs (Hart and Tirole (1990), Ordover et al. (1990), Salop and Scheffman (1983)). Capturing these "efficiency" and "foreclosure" theories and assessing the impact on welfare requires estimation of structural

[9] Other examples of vertical integration in the health care market include hospitals acquiring or becoming partners with other firms to offer durable medical equipment, hospice care, rehabilitation services, long-term care, or home-health services. See Antitrust Health Care Handbook $4^{th}$ edition, page 263, *American Bar Association*.

[10] If both upstream and the downstream firms have market power, then they both charge a price above marginal cost (both add markups). This phenomenon is called double-marginalization. Tirole (1988) examines the extreme case where both the upstream and downstream markets are monopolized and concludes that the independent behavior of the two monopolies results in smaller profits compared to what they could achieve if they set prices to maximize joint profits. Therefore, vertical integration benefits firms by increasing profits, and benefits consumers by lowering prices.

[11] Transaction costs involve writing, monitoring, and enforcing the contracts.

[12] If the contracts are incomplete, divergent interests of the parties engaged in bilateral exchange can lead to opportunistic behavior and therefore, losses. Vertical integration aligns interests of upstream and downstream firms, hence remedies this problem.

models of consumer demand, and in some cases, bargaining.

Estimation of a bargaining model is especially important when analyzing vertical integration in health care as reimbursement rates (prices paid to hospitals by insurers) are determined in yearly negotiations in the absence of vertical integration. The empirical literature on hospital-insurer bargaining aims to uncover how the surplus in the market is split between insurers and hospitals depending on their market power in the negotiation process. Brooks et al. (1997) were the first to estimate hospital-insurer bargaining in a Nash bargaining framework. Their model was later extended by many researchers, including Gowrisankaran et al. (2015), Lewis and Pflum (2015), and Ho and Lee (2017a, 2017b). This literature finds that hospitals have higher bargaining power in the negotiation process, agents' bargaining power increases in industry concentration, and hospitals in systems are able to set higher prices and extract a larger share of the surplus generated by contracting. I contribute to the bargaining literature by showing vertical integration increases hospitals' bargaining power after controlling for industry concentration, system membership, and observable hospital characteristics.

The bargaining model in this paper is also used to form new networks under the counterfactual policy experiment. The empirical literature on hospital-insurer network formation has established that restricted hospital choice decreases consumer welfare (Ho (2006)) and individuals value broader insurer networks (Ericson and Starc (2015)). Ho and Lee (2017b) incorporate bargaining into this framework and analyze networks formed under insurers' profit maximization problem. My analysis also synthesizes bargaining model with network formation, but it takes into account hospital profits while forming new networks and insurer profits while setting new premiums. Consistent with the findings in this literature, I conclude that consumer welfare increases with broader insurer networks and decreases with higher premiums.

# 3 Data

This paper uses data from several sources. Patient characteristics and discharge reports come from 2014 State Inpatient Databases (SID) published by the Health Care Utilization Project (HCUP)[13], and from California Office of Statewide Health Planning and Development (OSHPD) 2014 Public Patient Discharge data.[14] Hospital characteristics come from American Hospital Association (AHA) Annual Survey of Hospi-

[13]HCUP is operated by Agency for Healthcare Research and Quality (AHRQ).

[14]California OSHPD data is used only in bargaining estimation and counterfactual simulations.

Table 1: Patient Characteristics

| | Mean | SD | Min | Max |
|---|---|---|---|---|
| Distance (miles) | 15.56 | 117.25 | 0.04 | 428.31 |
| Female | 0.66 | 0.47 | 0 | 1 |
| Age | 27.68 | 21.91 | 0 | 64 |

*Notes:* N=1,152,081 discharges.

tals 2014. I supplement the hospital dataset using hospital star ratings obtained from Centers for Medicare and Medicaid Services (CMS). I also use financial data on hospitals reported in OSHPD Financial Disclosure Reports 2008-2014.[15] Insurer characteristics come from Atlantic Information Services (AIS) Directory of Health Plans 2016[16] with premium and enrollment data being supplemented by the WEISS Ratings Guide to Health Insurers 2016. Insurer characteristics from AIS include vertical integration indicator[17], enrollment, and number of enrolled by sector (commercial risk, public risk etc.). WEISS provides investment ratings, enrollment, age, and premiums of insurers. Additional plan characteristics are taken from National Committee for Quality Assurance (NCQA) Report on Health Plan Rankings 2015-2016. These characteristics include the type of the insurance plan (HMO, PPO etc.), states served, NCQA accreditation, an overall quality score as well as quality measures of consumer satisfaction, prevention, and treatment. I also use U.S. Census data on population (by age and sex) and number of uninsured by state to supplement the dataset.

I use SID data from Arizona, Florida, Kentucky, New Jersey, New York, Rhode Island, and Washington. These states represent over 20% of the entire U.S. population, and cover 1,152,081 discharges from 753 hospitals in total.[18] SID reports patient ZIP code, diagnosis[19], treatment, age, sex, and charges. I use only in-state, non-emergency-room (non-ER) hospital visits in my analysis. I observe patients' ZIP codes and the hospitals they visited, therefore I calculate the distance between a patient's residence location and hospital location.[20] This data is summarized in Table 1. Average patient in my data travels 16 miles to get care at a hospital. Females constitute 66% of all discharges due to the large number of pregnancies and childbirths. This paper focuses only on the non-elderly population (ages between 0 and 64) as people above 65 are likely

[15]Hospital cost function estimation uses this dataset.

[16]This issue of the directory reports the vertical integration indicator, which is key to my analysis. Other data I use, such as SID or OSHPD, come from 2014 as this was the last year available at the time of this study.

[17]AIS defines a health plan as provider-sponsored if it is owned by, or owns, in whole or in part, a segment of its provider network. My analysis only focuses on vertical integration of hospitals and insurers, therefore I do not consider health plans offered by physician groups as vertically integrated. For every VI insurer, I use data collected from the insurer's website to determine which hospitals are VI.

[18]I exclude federal government hospitals (such as Air Force hospitals, Veterans Affairs hospitals etc.) and long-term care hospitals from my analysis.

[19]I use 25 Major Diagnostic Categories (MDCs) defined by CMS.

[20]Distance is calculated as the distance between two latitude and longitude points of the hospital (as reported by AHA) and the centroid of the patient's zip code (as reported by SID).

Table 2: Hospital Characteristics

| | Mean | SD | Min | Max |
|---|---|---|---|---|
| Teaching hospital | 0.08 | 0.27 | 0 | 1 |
| Beds | 228 | 246 | 6 | 2,478 |
| Admissions | 10,169 | 12,356 | 36 | 146,388 |
| Full-time physicians | 32 | 99 | 0 | 1,346 |
| Full-time nurses | 333 | 494 | 0 | 5,819 |
| Inpatient days | 56,357 | 70,548 | 131 | 715,156 |
| Nurses per bed | 1.30 | 0.71 | 0 | 6.93 |
| For-profit | 0.24 | 0.43 | 0 | 1 |
| Women's health center | 0.49 | 0.50 | 0 | 1 |
| Kidney transplant | 0.05 | 0.21 | 0 | 1 |
| MRI | 0.61 | 0.49 | 0 | 1 |
| Pain management | 0.49 | 0.50 | 0 | 1 |

*Notes:* N=753 hospitals.

to be enrolled in Medicare plans and I analyze insurers active in the commercial business only. Since all newborns are considered as new patients in this dataset, the average patient is younger than expected.

Table 2 provides a summary of select variables from the hospital dataset.[21] I report information on 753 hospitals that operate in the seven states mentioned above. I observe ownership type, teaching status, system membership, total inpatient days, total number of admissions, and services offered by each hospital among other variables.

Variables available at the insurer level are summarized in Table 3. More than a fifth of the plans reported in AIS are VI. Therefore, any impact they will have on the market is expected to be substantial. Average premium per plan is calculated by dividing the total premium revenue reported in Weiss Ratings Guide by the enrollment data reported in AIS. Average premium per patient per month ranges between $42 and $1140 with an average of $461. The range is large since all types of plans (low-premium HMOs, high-premium PPOs etc.) are present in the dataset. In addition to premiums, I observe the age of the plan, the number of physicians who participated in the insurer's network of providers, and the total number of enrollees. I also create an indicator variable for Blue Cross Blue Shield (BCBS) plans as these are the dominant players in many markets. The rest of the variables are created using NCQA reports on plan performance. This source reports type[22] of each plan, which I aggregate to two categories: HMO/POS and PPO/Indemnity. In my data, 56% of the plans are PPO/Indemnity. NCQA also reports a score that takes into account NCQA

[21]Full list of hospital characteristics used in the analysis can be found in Table A2.

[22]Plan types reported in NCQA are: Health Maintenance Organization (HMO), Point of Service (POS), Preferred Provider Organization (PPO), and indemnity.

Table 3: Insurer Characteristics

| | Mean | SD | Min | Max |
|---|---|---|---|---|
| Vertically integrated | 0.21 | 0.41 | 0 | 1 |
| Premiums ($) | 461 | 290 | 42 | 1,140 |
| Age | 28.83 | 19.18 | 0 | 165 |
| Physicians | 44,828 | 56,387 | 27 | 574,650 |
| Total enrollment | 467,271 | 817,277 | 209 | 8m |
| PPO/Indemnity | 0.56 | 0.50 | 0 | 1 |
| Consumer satisfaction | 3 | 0.59 | 1.5 | 5 |
| Treatment | 2.84 | 0.67 | 1 | 4.5 |
| Prevention | 2.92 | 0.73 | 1 | 4.5 |
| NCQA rating | 3.34 | 0.61 | 1.5 | 5 |
| NCQA accreditation | 0.81 | 0.39 | 0 | 1 |
| BCBS | 0.15 | 0.36 | 0 | 1 |

*Notes:* N=380 insurers.

accreditation standards, member satisfaction, and clinical measures. The average NCQA rating for a plan is 3.34 out of 5. For quality, I use three measures of plan performance: consumer satisfaction, treatment, and prevention that also range between 1 (lowest performance level) and 5 (highest performance level).[23]

Lastly, hospital networks offered by insurers are hand-collected from individual insurers' websites. This data is available for 989 insurers from 50 states and Washington DC. It is important to note that while 380 insurers reported in Table 3 are distinct, the 989 insurers for which I have networks are not. I determine the insurers competing in a state by using enrollment data from AIS. If an insurer has more than 200 enrollees and offers commercial plans in a state, then that insurer is included in that state's market. Therefore, the same insurer can be listed in several states; however, it offers different hospital networks in different states as I create networks using only in-state hospitals. VI hospitals are identified in 16 states[24] by collecting data from VI insurers' websites. 13% of all hospitals and 14% of all insurers in these states are VI. This data is used in the reduced-form analysis whenever identification of VI-hospitals is necessary.

[23]The clinical quality measures (treatment and prevention) are calculated using a subset of the Healthcare Effectiveness Data and Information Set (HEDIS) measures whereas consumer satisfaction measure comes from the HEDIS survey which is overseen by AHRQ. Consumer satisfaction measure covers patients' satisfaction with health plans (handling claims, customer service etc.), satisfaction with physicians (doctors' communication, care received etc.), and access of getting care in terms of ease and promptness. The treatment measure evaluates scores in subcategories such as asthma, diabetes, heart attack, and mental health. Finally, the prevention score assesses measures such as timeliness of prenatal check ups, breast cancer screening, and early immunizations.

[24]I have data on Alaska, Alabama, Arkansas, Arizona, California, Colorado, Delaware, Florida, Kentucky, Massachusetts, Maryland, Maine, New Jersey, New York, Rhode Island, and Washington. Percent of VI hospitals ranges between 0% to 39%, percent of VI insurers ranges between 0% to 38% in these states.

# 4 Reduced-Form Analysis

## 4.1 Foreclosure

I begin my analysis by examining the impact of vertical integration on market foreclosure. Theory suggests that vertical integration can lead to two forms of market foreclosure: upstream and downstream. Upstream foreclosure arises when downstream firms' access to an upstream supplier is reduced as a result of commercial practices. On the contrary, downstream foreclosure occurs when upstream suppliers' access to a downstream firm is limited (Tirole (1988)). In the context of provider-insurer integration, upstream foreclosure arises if VI hospitals are not included in rival insurers' (VI or non-VI) hospital networks. Downstream foreclosure, on the other hand, emerges if VI insurers are less likely to carry rival hospitals (VI or non-VI) in their networks.

VI entities engage in market foreclosure with the purpose of weakening competition in either upstream or downstream market (Tirole (1988), Bolton and Whinston (1991)). For example, the VI entity may engage in upstream foreclosure by charging higher prices to competing insurers with the purpose of increasing rivals' costs. This practice will harm downstream competition through one of the two channels. If the competing downstream insurers are unwilling to or cannot afford to include the VI hospital in their networks, then the part of the market represented by this hospital will no longer be available to them. If they choose to include the VI hospital at the higher price, then they will have higher costs and therefore higher premiums. The VI entity, on the other hand, will have absolute cost advantages and hence can offer lower premiums and attract more consumers, all else being equal. Similarly, downstream foreclosure practices might harm upstream competition as the VI insurer does not give rival hospitals access to its health plan. A familiar example is Kaiser Permanente that includes only its own health facilities in its network. In this case, the rival hospitals do not have access to the VI insurer's enrollees and upstream competition is harmed as the market represented by the VI insurer is no longer accessible to the competing hospitals.

In order to investigate whether upstream foreclosure is present in the market, I start my analysis at the hospital level. If VI hospitals engage in upstream foreclosure, access to them should be reduced, hence they should be included in fewer health plans compared to non-VI hospitals. To examine this, I estimate the following linear regression model at the hospital level:

$$Y_h = \beta_0 + \beta_1 VIhospital_h + \alpha X_h + D_s + \epsilon_h \tag{1}$$

Table 4: Foreclosure Estimation - Hospitals

| | (1) | | (2) | | (3) | |
|---|---|---|---|---|---|---|
| | Number of plans included | Percent of plans included | Number of VI-plans included | Percent of VI-plans included | Number of non-VI-plans included | Percent of non-VI-plans included |
| VI-hospital | -1.45** | -0.04* | 0.85*** | 0.15*** | -2.30*** | -0.08*** |
| | (0.72) | (0.03) | (0.08) | (0.02) | (0.69) | (0.03) |
| Physicians | -0.003* | -0.0001* | 0.0001 | -0.00002 | -0.003** | -0.0001* |
| | (0.002) | (0.00006) | (0.0002) | (0.00005) | (0.002) | (0.00007) |
| Star rating | -0.11 | -0.0003 | 0.002 | 0.01 | -0.11 | -0.001 |
| | (0.28) | (0.009) | (0.03) | (0.009) | (0.27) | (0.01) |
| ln(Expenditure per adm.) | 0.57** | 0.04*** | 0.08*** | 0.05*** | 0.50** | 0.04*** |
| | (0.23) | (0.008) | (0.03) | (0.007) | (0.21) | (0.009) |
| Teaching hospital | -0.004 | 0.03 | -0.07 | 0.006 | 0.07 | 0.03 |
| | (0.97) | (0.04) | (0.13) | (0.03) | (0.92) | (0.04) |
| System member | 0.78* | 0.02 | 0.003 | 0.002 | 0.77* | 0.03 |
| | (0.43) | (0.02) | (0.05) | (0.01) | (0.41) | (0.02) |
| Physician group | 2.48** | 0.08 | -0.09 | -0.01 | 2.57** | 0.09 |
| | (1.11) | (0.05) | (0.09) | (0.04) | (1.07) | (0.05) |
| Burn care | 1.93** | 0.08** | 0.14 | 0.05 | 1.80** | 0.09** |
| | (0.90) | (0.03) | (0.13) | (0.03) | (0.87) | (0.04) |
| Cardiology services | 0.27 | 0.01 | 0.14** | 0.04* | 0.13 | 0.01 |
| | (0.60) | (0.02) | (0.07) | (0.02) | (0.58) | (0.03) |
| Chemotherapy | 0.10 | 0.005 | 0.13** | 0.03* | -0.03 | 0.002 |
| | (0.56) | (0.02) | (0.07) | (0.02) | (0.54) | (0.02) |
| $R^2$ | 0.85 | 0.85 | 0.69 | 0.71 | 0.84 | 0.85 |

*Notes:* Results from ordinary least squares estimation. $N = 1285$ hospitals. Robust standard errors in parentheses. State fixed effects included. All specifications control for hospital bed size category, primary service category, control authority, and hospital services reported in Table A1. *** statistically significant at 1% level, ** statistically significant at 5% level, * statistically significant at 10% level.

where the dependent variable $Y_h$ is the number (and percent) of health plans hospital $h$ is included in within a state, $VIhospital_h$ is a dummy variable indicating whether hospital $h$ is VI or not, $X_h$ are other hospital characteristics such as star rating, teaching status, number of full-time physicians at a hospital etc., $D_s$ are state fixed effects, and $\epsilon_h$ is the error term. Results are reported in Panel 1 of Table 4.[25] The coefficient of interest, $\beta_1$, is negative and significant implying that VI hospitals, on average, are included in 4% fewer health plans compared to their non-VI counterparts. Therefore, one can conclude foreclosure is present in the market. This estimation, however, does not inform us as to whether upstream or downstream foreclosure is present in the market as $\beta_1$ reflects multiple mechanisms that are at work. In particular, $\beta_1$ combines the following effects of efficiency, upstream foreclosure, and downstream foreclosure:

1. Positive effect of VI hospital $h^*$ being included in its own VI insurer's network due to efficiency

[25] Table A1 reports results from full specification with the entire set of hospital characteristics included.

2. Negative effect of VI hospital $h^*$'s upstream foreclosure of competing VI insurers

3. Negative effect of VI hospital $h^*$'s upstream foreclosure of competing non-VI insurers

4. Negative effect of competing VI insurers' downstream foreclosure of hospital $h^*$

It is important to notice that three (1, 2, and 4) out of four mechanisms pertain to VI insurers. Consequently, regressing the number (and percent) of VI health plans hospital $h$ is included in on the same set of hospital attributes gives a combination of these three effects. Results from this estimation are reported in Panel 2 of Table 4. The positive estimate of $\beta_1$ implies that the efficiency effects dominate the foreclosure effects in the market when only VI insurers are considered. Finally, to assess whether upstream foreclosure is exercised by VI hospitals, I regress the number of non-VI health plans hospital $h$ is included in on the same set of hospital characteristics. Results from this specification are reported in Panel 3 of Table 4. The negative and significant estimate of $\beta_1$ reflects only mechanism 3 above, and suggests that upstream foreclosure is present in the market. VI hospitals are included in 8% fewer non-VI health plans compared to their non-VI counterparts. While VI hospitals might also foreclose rival VI insurers, it is not possible to capture this mechanism using regression analysis.

Next, I investigate the presence of downstream foreclosure in the market using the same setup where the unit of observation is insurer-market and estimate the following equation:

$$M_{js} = \gamma_0 + \gamma_1 VIinsurer_j + \zeta Z_j + D_s + \epsilon_{js} \tag{2}$$

where $M_{js}$ is the number (and percent) of hospitals covered in insurer $j$'s network in state $s$, $VIinsurer_j$ is a dummy variable indicating whether insurer $j$ is VI or not, $Z_j$ are other insurer characteristics such as not-for-profit status, PPO, BCBS etc., $D_s$ are state fixed effects, and $\epsilon_{js}$ is the error term. I use two-way clustering (at the state level and at the plan level) for the standard errors to account for the correlation in errors within a state and within an insurer identifier as there are some large insurers that operate in multiple states in my data. Results from this specification are reported in Panel 1 of Table 5. While the negative and significant estimate of the coefficient of interest, $\gamma_1$, suggests that foreclosure is present in the market, again it is not possible to disentangle the impact of efficiency, upstream foreclosure, and downstream foreclosure. In this case, $\gamma_1$ combines the following analogous effects:

1. Positive effect of VI insurer $j^*$ including its own VI hospitals in its network due to efficiency

2. Negative effect of VI insurer $j^*$'s downstream foreclosure of competing VI-hospitals

Table 5: Foreclosure Estimation - Insurers

| | (1) | | (2) | | (3) | |
|---|---|---|---|---|---|---|
| | Number of hospitals covered | Percent of hospitals covered | Number of VI-hospitals covered | Percent of VI-hospitals covered | Number of non-VI-hospitals covered | Percent of non-VI-hospitals covered |
| VI-insurer | -14.16* | -0.14* | 0.31 | 0.03 | -14.47** | -0.14* |
| | (7.84) | (0.08) | (1.40) | (0.09) | (7.02) | (0.08) |
| Age | 0.05 | 0.0006 | 0.003 | 0.0002 | 0.05 | 0.0006 |
| | (012) | (0.0008) | (0.01) | (0.0007) | (0.12) | (0.0008) |
| PPO | 7.36 | 0.008 | 1.51* | 0.03 | 5.85 | 0.009 |
| | (5.51) | (0.04) | (0.78) | (0.02) | (4.95) | (0.04) |
| Not-for-profit | 19.29** | 0.09*** | 2.06** | 0.07*** | 17.23** | 0.09*** |
| | (8.36) | (0.03) | (0.91) | (0.02) | (7.64) | (0.03) |
| NCQA rating | -3.37 | -0.04 | 0.13 | -0.004 | -3.50 | -0.04 |
| | (4.28) | (0.03) | (0.75) | (0.03) | (4.21) | (0.03) |
| NCQA Accr. | -1.03 | 0.002 | -1.17 | -0.04 | 0.14 | 0.007 |
| | (8.26) | (0.05) | (0.83) | (0.03) | (7.53) | (0.05) |
| BCBS | 10.16 | 0.06 | 1.50* | 0.07 | 8.66 | 0.05 |
| | (9.11) | (0.04) | (0.83) | (0.05) | (8.41) | (0.04) |
| Weiss rating | 2.15* | 0.005 | 0.14 | 0.008* | 2.01** | 0.005 |
| | (1.14) | (0.006) | (0.19) | (0.005) | (1.004) | (0.006) |
| Enrollment | 0.003 | 0.000007 | 0.0005 | 0.000007 | 0.003 | 0.0001 |
| | (0.006) | (0.00002) | (0.0003) | (0.00002) | (0.00002) | (0.00006) |
| $R^2$ | 0.61 | 0.38 | 0.71 | 0.64 | 0.62 | 0.39 |

*Notes:* Results from ordinary least squares estimation. $N = 339$ insurers from 16 markets. Robust, two-way clustered (at the plan level and at the state level) standard errors in parentheses. All specifications include state fixed effects. Enrollment is in thousands. *** statistically significant at 1% level, ** statistically significant at 5% level, * statistically significant at 10% level.

3. Negative effect of VI insurer $j^*$'s downstream foreclosure of non-VI-hospitals

4. Negative effect of competing VI hospitals' upstream foreclosure of insurer $j^*$

Following the same reasoning as above, I show that downstream foreclosure is present in the market by isolating mechanism 3 and estimating the analogue regression. Results are reported in Panel 3 of Table 5. When I use the number (and percent) of non-VI hospitals included in insurer $j$'s network as the dependent variable, I find that VI insurers are less likely to carry non-VI hospitals in their networks, suggesting they engage in downstream foreclosure. In particular, VI insurers carry 14% fewer non-VI hospitals in their networks, on average, compared to non-VI insurers. While VI insurers might also engage in downstream foreclosure by denying rival VI hospitals access to their health plans, this mechanism cannot be captured using regression analysis.

### 4.2 Network Size, Premiums, and Quality

**Network Size:**

The results from previous section on foreclosure suggest that VI insurers, on average, offer fewer hospitals in their networks. To investigate whether this result implies VI plan enrollees have access to fewer resources, I use the same sample and the same specification as equation (2) with two measures of network size as dependent variables: beds per a hundred enrollee population and physicians per a hundred enrollee population. Enrollment is not included as an explanatory variable as it is accounted for in the dependent variable. I cluster standard errors both at the state level and at the insurer level as errors might be correlated within a state and are expected to be correlated for plans that operate in several states. Results are represented in Table 6. In both specifications, while the coefficient of interest is negative, it is not statistically significantly different than zero. This result suggests that even though individuals enrolled in VI plans are constrained to narrower hospital networks, they do not necessarily have fewer resources at their use.

Table 6: Network Size Estimation

| | (1) Beds | (2) Physicians |
|---|---|---|
| VI-insurer | -0.76 | -0.81 |
| | (0.75) | (0.75) |
| Age | -0.005 | -0.005 |
| | (0.01) | (0.01) |
| PPO | 0.99* | 1.004** |
| | (0.51) | (0.51) |
| Not-for-profit | 1.30** | 1.27** |
| | (0.52) | (0.52) |
| NCQA rating | 1.07*** | 1.03*** |
| | (0.35) | (0.34) |
| NCQA accreditation | -0.47 | -0.57 |
| | (0.50) | (0.52) |
| BCBS | 1.24** | 1.33** |
| | (0.53) | (0.54) |
| Weiss rating | -0.02 | -0.03 |
| | (0.08) | (0.08) |
| $R^2$ | 0.20 | 0.20 |

*Notes:* Results from ordinary least squares estimation. N=339 insurers from 16 states. Robust, two-way clustered (at the plan level and at the state level) standard errors in parentheses. All specifications include state fixed effects. *** statistically significant at 1% level, ** statistically significant at 5% level, * statistically significant at 10% level.

**Premiums:**

Next, I investigate whether VI insurers charge higher or lower premiums compared to their non-VI counterparts. If the VI insurers pass along the cost savings to their enrollees, then one would expect the VI entity to charge lower premiums. If, on the other hand, they are able to gain and exercise market power through foreclosure, they will be able to charge higher premiums to their enrollees.

To my knowledge, the only paper to date that examined the impact of provider-insurer integration on premiums is Frakt et al. (2013). The authors find VI insurers charge $28 more per month in the Medicare Advantage market. However, their data is from 2009, hence it does not reflect the effects of the ACO bonus program that might have given these plans cost advantages in the post-ACA era. Recent papers in the literature find that insurers charge higher premiums if the insurance market concentration is high (Dafny et al. (2012)), the upstream industry is consolidated (Town et al. (2006)), or insurers have lower bargaining power while negotiating with providers (Trish and Herring (2015)).

I estimate the following linear model to examine the vertical integration-premium relationship:

$$lnPremium_j = \beta_0 + \beta_1 VIinsurer_j + \alpha Z_j + \gamma M_j + D_s + \epsilon_j \tag{3}$$

where $Premium_j$ is insurer $j$'s premium per enrollee per month, $Z_j$ are insurer characteristics, $M_j$ are market characteristics, $D_s$ are state fixed effects, and $\epsilon_j$ is the error term. Since the unit of observation is at the insurer level, I aggregate all market-level variables to insurer-level variables using insurer $j$'s state enrollments as weights for the states insurer $j$ is active in.[26] I also include state fixed effects for every state the insurer operates in, therefore more than one state fixed effect can be assigned to an observation.

Results are reported in Table 7. Specifications reported in first two columns use aggregated market variables mentioned above, and only differ in the market characteristics included in the estimation.[27] Specification in the last column uses state-level market characteristics (instead of weighted and aggregated) for the domicile state of the insurer (as reported by Weiss) as well as state fixed effects for a single domicile state. Standard errors in this specification are clustered at the state level.

[26]This is a common practice in the literature. See Song et al. (2012), Frakt et al. (2013), and Trish and Herring (2015) among others.

[27]I also estimated the baseline specification using additional market controls (such as percent Hispanic, percent elderly, percent non-profit insurers, hospital discharges) and obtained similar results.

Table 7: Premium Estimation

| | ln Premium (per person per month) | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| VI-insurer | -0.23** | -0.25*** | -0.32*** |
| | (0.09) | (0.09) | (0.11) |
| Age | -0.004** | -0.004* | -0.006** |
| | (0.002) | (0.002) | (0.002) |
| PPO | 0.33*** | 0.32*** | 0.27*** |
| | (0.07) | (0.07) | (0.08) |
| BCBS | -0.45*** | -0.46*** | -0.45*** |
| | (0.10) | (0.10) | (0.11) |
| Weiss rating | 0.04*** | 0.03*** | 0.03** |
| | (0.01) | (0.01) | (0.01) |
| Percent integrated | -0.15 | 0.56 | -1.77*** |
| | (0.68) | (0.75) | (0.18) |
| Insurer HHI | -0.005* | -0.004 | 0.03* |
| | (0.003) | (0.003) | (0.02) |
| Hospital HHI | 0.0005 | -0.002 | 0.02** |
| | (0.003) | (0.003) | (0.006) |
| ln(Per capita income) | 0.50 | 0.04 | 1.64*** |
| | (0.52) | (0.59) | (0.21) |
| ln(Medicare hospital payments) | 0.37 | 0.48 | 1.83*** |
| | (0.70) | (0.72) | (0.32) |
| ln(Input price) | 1.97** | 1.87** | -4.47*** |
| | (0.76) | (0.76) | (1.60) |
| Unemployment rate | -0.43 | -0.02 | -0.02 |
| | (0.33) | (0.39) | (0.05) |
| Percent Black | | -1.42* | -4.69** |
| | | (0.73) | (2.01) |
| Percent Uninsured | | 3.21 | -12.42*** |
| | | (3.56) | (3.54) |
| $R^2$ | 0.38 | 0.39 | 0.34 |

*Notes:* Results from ordinary least squares estimation. N=380 insurers. Robust standard errors in parentheses. State fixed effects included. *** statistically significant at 1% level, ** statistically significant at 5% level, * statistically significant at 10% level.

The main finding is that VI insurers charge lower premiums after controlling for insurer and market characteristics, and this result is robust across specifications. The magnitude of reduction in premiums compared to non-VI insurers is between 20.5% to 27.4% based on the specification.[28] The coefficient estimates for other variables are as expected and in line with the previous findings in the literature. PPOs that are characterized by larger networks charge higher premiums, while BCBS insurers have lower premiums likely because they exercise their bargaining power in negotiations with hospitals, extract lower prices, and pass on the savings to their enrollees. Percent of the health plans that are integrated does not influence the premiums when all the states the insurer operates in are taken into account. I use the commonly accepted Herfindahl-Hirschman Index (HHI) as a measure of market concentration in both insurer and hospital markets.[29] Insurer HHI is positive and significant in the last specification, implying insurers charge higher premiums when the market gets closer to a monopoly. Hospital HHI coefficient is positive where significant, as anticipated. As hospital HHI increases, hospital markets get more concentrated, and as a result, hospital prices increase. This implies increased costs for insurers, which translates into higher premiums.

Since I have only one year of data and no source of exogenous variation within this year, I cannot assign a causal interpretation to my findings. However, the main result that VI insurers charge lower premiums on average is robust across specifications. It is also robust to inclusion of other market level controls and different definitions of HHI as discussed above. This result implies that VI insurers are indeed passing along the savings from reduced costs to their enrollees. This is also in line with their strategies to offer low-premium high-quality plans and attract large patient pools.[30]

**Quality:**

The last piece of my reduced-form analysis examines the relationship between quality and vertical integration. Frakt et al. (2013) find that VI insurers in the Medicare Advantage market offer higher quality plans. I conduct a similar analysis for the plans operating in the commercial market, and estimate the following equation at the plan level:

$$Quality_j = \beta_0 + \beta_1 VIinsurer_j + \alpha Z_j + \gamma M_j + D_s + \epsilon_j \tag{4}$$

[28] The percent impact on raw premiums is calculated as $(e^{\beta_1} - 1) \times 100$.

[29] HHI is calculated as the sum of squared market shares of individual entities in the market, scaled by 100. Therefore, HHI ranges between [0, 100] with 100 representing the highest concentration possible. Both HHI measures are calculated at the state level. Hospital HHI is calculated using inpatient days. I also estimated the premium equation using hospital HHI calculated based on total admissions, total number of beds, and total Medicare discharges, and obtained similar results. Insurer HHI is calculated using state enrollment as reported by AIS.

[30] Source: “The Rapid Rise of Provider-Sponsored Health Plans: Two Case Studies” webinar by AIS.

I use the four quality measures as reported by NCQA: an overall quality measure (NCQA Rating), quality of prevention, quality of treatment, and quality of consumer satisfaction. Again, the market level variables are aggregated to insurer level using the states the insurer is active in, and state fixed effects are included for the states the insurer operates in.

The results reported in Table 8 imply that, on average, VI insurers offer higher quality plans in terms of all measures except consumer satisfaction. This superior quality of VI plans compared to their non-VI counterparts likely arises from the clinical integration within the VI entity that enables both the provider-side and insurer-side to better track the patients, and take preventive measures or treat them accordingly.

To conclude, the reduced-form analysis shows that VI entities engage in both upstream and downstream foreclosure; VI insurers offer fewer hospitals in their networks but not necessarily fewer beds or physicians; and they charge lower premiums and provide better quality in multiple dimensions. While consumer welfare might be harmed as a result of limited choice due to foreclosure, it might also improve due to lower premiums and better quality. Similarly, VI entities might attract more consumers as a result of lower premiums and better quality, but they might also be losing consumers due to limited choice. Non-VI hospitals and insurers might be harmed by vertical integration as they do not have access to VI entities hence do not have the same cost advantages, but they might also be preferred by consumers as the networks they offer are not necessarily limited. Consequently, the overall impact of vertical integration on welfare remains to be ambiguous. I estimate the structural model outlined in the next section and measure the overall impact on consumer and producer welfare under a counterfactual scenario that removes vertical integration from the market.

## 5 Structural Analysis: Model and Methodology

The structural model, used to assess the overall impact on welfare, consists of three main components. First, I use discrete choice models of demand to explain how consumers choose hospitals to visit and choose insurance plans to enroll in. Demand side closely follows the setup developed by Capps et al. (2003) and extended by Ho (2006). In the second stage, I estimate a cost function to examine how hospital costs are determined. The final component of my model combines parameter estimates from demand models and cost function to estimate a bargaining model. The hospital-insurer negotiations are modeled within a Nash

Table 8: Quality Estimation

| | (1) NCQA Rating | (2) Prevention | (3) Treatment | (4) Consumer Satisfaction |
|---|---|---|---|---|
| VI-insurer | 0.39*** | 0.43*** | 0.52*** | 0.12 |
| | (0.08) | (0.10) | (0.09) | (0.09) |
| Age | 0.001 | 0.003 | 0.002 | -0.003* |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| PPO | -0.21*** | -0.22*** | -0.21*** | -0.10 |
| | (0.06) | (0.08) | (0.07) | (0.07) |
| BCBS | 0.41*** | 0.56*** | 0.46*** | 0.41*** |
| | (0.09) | (0.11) | (0.10) | (0.10) |
| Weiss rating | -0.006 | -0.02 | -0.01 | 0.01 |
| | (0.01) | (0.01) | (0.01 | (0.01) |
| Percent integrated | 0.26 | 1.04 | 0.38 | -1.66** |
| | (0.67) | (0.80) | (0.72) | (0.75) |
| Insurer HHI | -0.004 | -0.002 | -0.004 | -0.004 |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| Hospital HHI | 0.004* | 0.006*** | 0.002 | -0.001 |
| | (0.002) | (0.003) | (0.002) | (0.002) |
| ln(Per capita income) | 1.45*** | 1.66*** | 1.26** | -0.46 |
| | (0.52) | (0.63) | (0.56) | (0.58) |
| ln(Medicare hospital payments) | 0.10 | 0.38 | -1.35** | 1.80** |
| | (0.64) | (0.76) | (0.69) | (0.71) |
| ln(Input price) | 0.14 | -0.25 | -0.92 | 1.45* |
| | (0.68) | (0.81) | (0.73) | (0.75) |
| Unemployment rate | 0.29 | -0.04 | 0.69* | -1.26*** |
| | (0.35) | (0.42) | (0.38) | (0.39) |
| Percent black | 0.55 | 1.13 | -0.28 | 2.75*** |
| | (0.65) | (0.78) | (0.70) | (0.72) |
| Percent uninsured | -1.28 | 0.08 | -1.38 | 0.93 |
| | (3.16) | (3.80) | (3.42) | (3.53) |
| $R^2$ | 0.48 | 0.49 | 0.40 | 0.31 |

*Notes:* Results from ordinary least squares estimation. N=380 insurers from 50 U.S. states. Robust standard errors in parentheses. State fixed effects included for each state the insurer operates in. *** statistically significant at 1% level, ** statistically significant at 5% level, * statistically significant at 10% level.

bargaining framework following Lewis and Pflum (2015). I use the bargaining model to investigate how vertical integration affects bargaining power, and to determine the new insurer networks under counterfactual simulations.

## 5.1 Estimation of the Demand Side

The vertical relationship between the three agents in the health care market calls for two separate demand models: demand for hospitals by consumers and demand for insurers by consumers. First, I estimate the hospital demand model using a conditional logit framework.[31] Next, I use the results from hospital demand estimation to construct an "expected utility" measure and use this as an insurer characteristic in health plan demand estimation. This measure captures the predicted utility an average patient gets from a network of hospitals offered by an insurer. Finally, I estimate health plan demand using the contraction mapping algorithm developed by Berry, Levinsohn, and Pakes (1995) (henceforth BLP) taking into account unobserved plan characteristics as well as heterogeneity in consumer preferences towards certain insurer characteristics.

### 5.1.1 Hospital Demand Model

Let utility of patient $i$ from visiting hospital $h$ in market $m$ given diagnosis $l$ be:

$$u_{ihlm} = u(X_{hm}, V_{ilm}|\lambda, \theta) \tag{5}$$

where $(\lambda, \theta)$ are parameters to be estimated, and $X_{hm}$ is a vector of observed hospital characteristics. $V_{ilm} = [D_{im}, C_{ilm}]$ is a vector of observed consumer characteristics where $D_{im}$ represents demographic characteristics such as sex, age, location and $C_{ilm}$ represents clinical attributes such as diagnosis. Patients choose hospitals to maximize utility, so if patient $i$ with diagnosis $l$ chooses hospital $h$, then the following inequality must hold for all other hospitals $h'$ in the market, where the market subscript $m$ will be suppressed for notational ease:

$$u_{ihl} = u(X_h, V_{il}|\lambda, \theta) \geq u_{ih'l} = u(X_{h'}, V_{il}|\lambda, \theta) \tag{6}$$

In particular, let the utility specification be:

$$u_{ihl} = \theta X_h + \lambda_1 X_h D_i + \lambda_2 X_h C_{il} - \gamma(V_i) OPC(C_{il}) + \epsilon_{ihl} \tag{7}$$

[31] I use the standard conditional logit model proposed in McFadden (1974).

where $OPC(C_{il})$ is the out-of-pocket cost for patient $i$ with diagnosis $C_{il}$ and $\gamma$ converts money into utils based on patient characteristics $V_{il}$ and informs us about the price sensitivity of patients.[32] I assume that the independently and identically distributed error term $\epsilon_{ihl}$ captures idiosyncratic tastes and has an Extreme Value Type 1 distribution. Then, the hospital share equation can be written as:

$$s_h(\mathcal{M}) = \frac{e^{u(X_h,V_{il}|\lambda,\theta)}}{\sum\limits_{k\in\mathcal{M}} e^{u(X_h,V_{il}|\lambda,\theta)}} \tag{8}$$

where $\mathcal{M}$ is the set of hospitals that are in the same choice set as hospital $h$. There is no outside option because I only observe sick patients who are hospitalized. Since I observe the actual shares, I use maximum likelihood to obtain the parameter estimates $\hat{\lambda}$ and $\hat{\theta}$. Identification in this model comes from the variation in patients' hospital choice sets across markets. Unlike the health plan demand model (presented next), this model does not account for unobserved quality of hospitals. I have very rich hospital characteristics data, therefore I assume that the characteristics I use in estimation capture the quality of hospitals.

### 5.1.2 Expected Utility and Willingness-to-Pay

The hospital demand model presented above is used to create two measures: expected utility from an insurer's network of hospitals (used in insurer demand estimation) and willingness-to-pay (WTP) of an individual for a hospital to be included in his choice set (used in bargaining estimation).[33] The common element to both calculations is patient $i$'s interim utility from having access to a set of hospitals $\mathcal{M}$:

$$V(\mathcal{M}|X_h,V_{il}) = E\Big[\max_{h\in\mathcal{M}} \hat{u}(X_h,V_{il}|\hat{\lambda},\hat{\theta})\Big] = ln\Big[\sum_{h\in\mathcal{M}} e^{\hat{u}(X_h,V_{il}|\hat{\lambda},\hat{\theta})}\Big] \tag{9}$$

The expected utility a type $q$ patient[34] gets from a network of hospitals $\mathcal{M}$ offered by insurer $j$ is given by:

$$EU_{qj}(\mathcal{M}) = \sum_l p_{ql}\ V(\mathcal{M}|X_h,V_{ql}) = \sum_l p_{ql}\ ln\Bigg(\sum_{h\in\mathcal{M}} e^{\hat{u}(X_h,V_{ql}|\hat{\lambda},\hat{\theta})}\Bigg) \tag{10}$$

where $p_{ql}$ is the probability that patient type $q$ is hospitalized with diagnosis $l$. That is, first I calculate a different interim utility for each diagnosis each patient type $q$ might have. Then, I take a weighted sum across these values where the weights are $p_{ql}$ to calculate the expected utility patient type $q$ receives from

[32] In practice, a patient's hospital choice depends on out-of-pocket expenditures at a hospital. However, since I do not observe these costs, while estimating the hospital demand model, I make the common assumption (as in Capps et al. (2003), Ho (2006), Lewis and Pflum (2015)) that they are constant across hospitals for a given patient and therefore do not affect hospital choice. Nonetheless, I still estimate the price sensitivity parameter $\gamma$ jointly with the parameters of the bargaining model.

[33] The construction of these measures follows Ho (2006) and Capps et al. (2003), respectively.

[34] Patient types are defined by age-sex-ZIP code cells, where age brackets are 0-17, 18-34, 35-44, 45-54, 55-64.

having access to insurer $j$'s network of hospitals $\mathcal{M}$. The expected utility measure is calculated for type $q$ patient (instead of individual $i$) because I do not observe consumers' choices of health plans. Health plan demand estimation uses aggregate shares data to obtain the parameter estimates.

On the contrary, I calculate the WTP measure based on individual $i$'s interim utility because I observe hospital choice at the individual level. The contribution of hospital $h$ to patient $i$'s interim utility from the network of hospitals $\mathcal{M}$ can be calculated as:

$$\Delta_h V(\mathcal{M}|X_h, V_{il}) = V(\mathcal{M}|X_h, V_{il}) - V(\mathcal{M} \setminus h|X_h, V_{il}) = ln\Big(\frac{1}{1 - s_h(\mathcal{M})}\Big) \tag{11}$$

The total ex-ante WTP for inclusion of hospital $h$ in the network of hospitals $\mathcal{M}$ is then given by integrating (11) across a cumulative distribution of patient characteristics and diagnosis $F(V_{il})$:

$$\Delta W_j(\mathcal{M}) = N_j \int\limits_V \frac{1}{\gamma} ln\Big(\frac{1}{1 - s_h(\mathcal{M})}\Big) dF(V_{il}) \tag{12}$$

where $N_j$ is the number of patients enrolled with insurer $j$ that visit a hospital.

#### 5.1.3 Insurer Demand

I estimate consumer demand for insurers using a discrete choice setting that accounts for unobserved individual characteristics as well as the expected utility a patient gets from a network of hospitals. I start by estimating a benchmark logit model that closely follows the specification in Berry (1994) and then move onto BLP estimation that takes into account heterogeneity in individual preferences towards insurer characteristics as an additional layer.

**Logit Model:**

Let utility individual $i$ gets from plan $j$ in market $r$ be:

$$w_{ijr} = \sum_k x_{jkr}\beta_k + \xi_{jr} + \epsilon_{ijr} \tag{13}$$

where $x_{jkr}$ is the $k^{th}$ observed plan characteristic of plan $j$ and $\xi_j$ represents the unobserved plan characteristic (such as patients' perception about quality, status, service, reputation, past experience etc.). For simplicity, I drop the market subscripts in the rest of the analysis. Therefore, the utility function can be written as:

$$w_{ij} = \sum_k x_{jk}\beta_k + \xi_j + \epsilon_{ij} = \delta_j(x_j, \xi_j, \beta) + \epsilon_{ij} \tag{14}$$

where $\delta_j$ represents the mean utility level from plan $j$. The unobserved characteristics are assumed to be mean independent of $x_j$'s and also independent across markets. The error term $\epsilon_{ij}$ is independently and identically distributed across consumers and plans and assumed to have an Extreme Value Type 1 distribution. Normalizing the mean utility from the outside good to be zero (i.e. $\delta_o = 0$), the closed-form solution for the market share equation for product $j$ can be written as:

$$s_j = \frac{e^{\delta_j}}{1 + \sum_{g=1}^{G} e^{\delta_g}} \tag{15}$$

where $G$ is the number of plans in the market. The share of the outside good is given by:

$$s_o = \frac{1}{1 + \sum_{g=1}^{G} e^{\delta_g}} \tag{16}$$

Dividing equation (14) by equation (15) gives:

$$\frac{s_j}{s_o} = e^{\delta_j} \Longrightarrow ln(s_j) - ln(s_o) = \delta_j \tag{17}$$

Hence, I generate $\delta$'s using the market share data. Having obtained the dependent variable, I estimate the following equation to obtain the parameter estimates:

$$\delta_j = \sum_k x_{jk}\beta_k + \xi_j \tag{18}$$

Before moving on with estimation, the endogeneity problem caused by premiums needs to be addressed. The unobserved plan characteristic $\xi_j$ (the error term in equation (18)) is likely to be correlated with the plan's premium which is one of the observed plan characteristics. One would expect a high-quality, better-service plan to charge a higher premium. For this reason, I instrument for the premium variable. Traditional instruments used in the literature for price are cost shifters (these are difficult to find as they are usually correlated with $\xi$'s), characteristics of competing products in the same market, and prices of the same product in other markets (because a shock to marginal cost will be carried to prices in other markets). I use characteristics (other than premium) of other plans within the same market as instruments. These instruments and the relevant validity tests are further discussed in section 5. Given these instruments $Z$, I form the moment conditions as follows. First, I calculate the unobserved quality term $\xi_j$ as a function of model parameters:

$$\xi_j = \delta_j - \sum_k x_{jk}\beta_k = ln(s_j) - ln(s_o) - \sum_k x_{jk}\beta_k \tag{19}$$

The instruments should be orthogonal to this unobserved quality term, so I form the moment conditions as $E[\xi(\beta)'Z] = 0$. In applying iterative Generalized Method of Moments (GMM), I use the "optimal" weighting matrix $W$ which is the inverse of the variance of moment conditions. Therefore, the problem reduces to:

$$\min_\beta \xi(\beta)'ZWZ'\xi(\beta) \quad where \quad W = (E[Z'\xi\xi'Z])^{-1} \tag{20}$$

The analytical solution to this problem is:

$$\beta = (X'ZWZ'X)^{-1}(X'ZWZ'\delta) \tag{21}$$

The iterative estimation algorithm starts with $W = (Z'Z)^{-1}$ to get an initial estimate $\hat{\beta}$, and then I re-compute $W = (E[Z'\xi(\hat{\beta})\xi(\hat{\beta})'Z])^{-1}$ to get a new estimate of $\beta$.

Identification in the health plan demand model comes from the variation in consumers' choice sets across markets as well as the variation of health plan characteristics within a market. Markets are defined by states. Results from the health plan demand estimation are presented in Table 6.

**BLP:**

The major drawback of the previous model is that it does not generate realistic substitution patterns. In this setting, the cross-price elasticity between any two plans depends only on their market shares. Consider two health plans A and B whose market shares are the same. Let A be an HMO plan with low premiums, narrow hospital and physician networks, and low ratings; and B be a PPO plan with high premiums, large provider network, and top ratings. Assume there is another PPO plan C in the market with high premiums, large provider network, and high quality ratings. The cross-price elasticity of the previous model implies that if plan C increases its premiums, the demand for plan A and plan B will increase equally. This is unintuitive as we expect the cross-price effect to be larger for health plans that are similar in characteristics. The model presented by BLP solves this problem and generates realistic substitution patterns. With the BLP estimation outlined below, cross-price elasticities are larger for products that are closer together in terms of their characteristics.

Let utility of patient $i$ from insurer $j$ in market $r$ be:

$$w_{ijr} = \xi_{jr} + x_{jr}\phi + \beta_1 EU_{jr} + \beta_2 Prem_{jr} + \gamma_1 \nu_{i1} EU_{jr} + \gamma_2 \nu_{i2} Prem_{jr} + \eta_{ijr} = \delta_j + \mu(\nu_{i1}, \nu_{i2}) + \eta_{ijr} \quad (22)$$

where $\xi_j$ are unobserved insurer characteristics, $x_j$ are observed insurer characteristics, $Prem_j$ is insurer $j$'s premium, $\nu_i$ are random draws from a normal distribution and represent unobserved individual preferences, and $\eta_{ij}$ are idiosyncratic shocks to consumer tastes that are assumed to be independently and identically distributed with Extreme Value Type 1. The expected utility measure presented in the previous subsection is aggregated to insurer level by taking a weighted sum across patient types where weights are population shares of type $q$ individuals obtained from Census data. $\delta_j$ is the mean utility level that a patient gets from plan $j$. It is the presence of the interaction terms $\mu$ that allows me to capture the heterogeneity of preferences. In this setting, consumers with similar characteristics prefer similar products. Therefore, if an insurer is removed from the choice set, consumers will substitute to other insurers that are similar in terms of characteristics and this generates more realistic substitution patterns.

Identification in this model comes from the variation in patients' plan choice sets across markets. To address the endogeneity issue, I again instrument for premiums using the BLP-type instruments mentioned above. The outside good is defined as having no insurance and its share is calculated using the Census data. In this setting, the share equation for plan $j$ cannot be solved analytically. As in BLP, I use simulation techniques to obtain the predicted shares:

$$\hat{s_{jr}}(\phi, \gamma, \beta) = \frac{1}{ns} \sum_{i=1}^{ns} \frac{e^{(\xi_{jr} + x_{jm}\phi + \beta_1 EU_{jr} + \beta_2 Prem_{jr} + \gamma_1 \nu_{i1} EU_{jr} + \gamma_2 \nu_{i2} Prem_{jr})}}{1 + \sum_{k \in P} e^{(\xi_{kr} + x_{kr}\phi + \beta_1 EU_{kr} + \beta_2 Prem_{kr} + \gamma_1 \nu_{i1} EU_{kr} + \gamma_2 \nu_{i2} Prem_{kr})}} \quad (23)$$

where $ns$ is the number of random draws (1000 in my estimation), and $P$ is the set of plans in the market. That is, I calculate a different share with each distinct draw of the unobserved individual preference term $\nu_i$, and then obtain the predicted share as an average of these simulated shares across draws. Dropping the market subscript and simplifying notation, one can write the predicted shares as:

$$\hat{s_j}^{ns} = \frac{1}{ns} \sum_i \frac{e^{\delta_j + \mu(x_j, \nu_i)}}{1 + \sum_j e^{\delta_j + \mu(x_j, \nu_i)}} \quad (24)$$

Given the equation for predicted shares, I use the contraction mapping algorithm introduced by BLP to obtain $\delta$, the mean utility level vector. This algorithm aims to match the predicted shares $\hat{s}$ to the observed true shares $s$ using the following equation:

$$\delta^h = \delta^{h-1} + ln(s) - ln(\hat{s}) \tag{25}$$

I begin by evaluating the right-hand side at an initial guess of parameters and $\delta$, obtain a new $\delta$, put it back into the right-hand side and repeat this until convergence is reached. Once I obtain $\delta$, I rewrite the unobserved plan characteristics as $\xi_j = \delta_j - x_j\phi$. Therefore, I form the moment conditions as $E[\xi' Z] = 0$ and estimate via GMM.

### 5.2 Hospital Cost Function

I adapt the trans-log specification common in the literature (Fournier and Mitchell (1997), Capps et al. (2010), Lewis and Pflum (2015)) that is used to estimate cost functions of multiproduct firms and hospitals. According to this specification, hospital $h$'s costs at time $t$ are given by:

$$\begin{aligned} ln(Cost_{ht}) = \beta_0 \ + \ \beta_1 ln(Y_{ht}) \ + \ \beta_2 ln(Y_{ht}) \times ln(Y_{ht}) \ + \ \beta_3 ln(W_{ht}) \ + \ \beta_4 ln(W_{ht}) \times ln(W_{ht}) \\ + \ \beta_5 ln(Y_{ht}) \times ln(W_{ht}) \ + \ \kappa_{ht} \ + \ t \ + \ \epsilon_{ht} \end{aligned} \tag{26}$$

where $Y_h$ are hospital outputs, $W_h$ are hospital inputs, $\kappa$ are hospital fixed effects, and $t$ is a time trend. The error term $\epsilon_{ht}$ is clustered at the hospital level, therefore it is allowed to be correlated for a hospital across years, but errors are assumed to be distributed independently across hospitals. The dependent variable is the natural logarithm of the total operating costs at a hospital. The hospital output vector, $Y_h$, consists of inpatient days and outpatient visits for private insurance, Medicare, and other payer types (such as Medi-Cal, workers' compensation, county indigent programs, self-pay etc.). The hospital input vector, $W_h$, includes size measures for the hospital (such as number of beds, fixed assets, total number of hours for registered nurses etc.), governance structure, for-profit status, rural status, vertical integration status, teaching status, and full-line forcing (FLF) status.[35] The marginal costs of both components are allowed to vary by other inputs and outputs through the inclusion of the interaction terms. The results from hospital cost function estimation are presented in Table A3 in the appendix.

[35] A hospital system is defined as an FLF system if all of its member hospitals are included in at least one insurer's network. FLF status variable indicates whether the hospital belongs to an FLF system or not.

### 5.3 Bargaining

The model used for bargaining estimation closely follows the specification in Lewis and Pflum (2015).[36] I use the Nash bargaining framework where the two agents negotiate to split the surplus they jointly generate by successfully contracting. The outcome of the bargaining game is a contract equilibrium as in Crémer and Riordan (1987) that relies on the following assumptions:

1. All hospital-insurer pairs negotiate contracts simultaneously.
2. All hospital-insurer pairs negotiate under the anticipation that all other hospital-insurer pairs will successfully negotiate contracts.
3. The bargaining outcome between a hospital and an insurer does not influence the bargaining outcome of these parties with other insurers and hospitals.
4. When a hospital is removed from a choice set, patients re-allocate themselves to other hospitals in the same choice set.

Given these assumptions, the objective function of the Nash bargaining game is:

$$\max_{p_{hj}}[\Pi_h(\mathcal{H}) - \Pi_h(\mathcal{H}\backslash j)]^{\alpha_h}[\Pi_j(\mathcal{M}) - \Pi_j(\mathcal{M}\backslash h)]^{1-\alpha_h} \tag{27}$$

where $\Pi_h(\mathcal{H})$ are profits of hospital $h$ when it contracts with a set of insurers $\mathcal{H}$, $\Pi_h(\mathcal{H}\backslash j)$ are profits of hospital $h$ when it contracts with the same set of insurers except insurer $j$, $\Pi_j(\mathcal{M})$ are profits of insurer $j$ when it contracts with a set of hospitals $\mathcal{M}$, $\Pi_j(\mathcal{M}\backslash h)$ are profits of insurer $j$ when it contracts with the same set of hospitals except hospital $h$, $p_{hj}$ are the negotiated prices between hospital $h$ and insurer $j$, $\alpha_h$ represents the bargaining power of hospital $h$, while $1-\alpha_h$ represents the bargaining power of insurer $j$.

The objective function can be expressed as $\max_{p_{hj}}[\Delta_j\Pi_h]^{\alpha_h}[\Delta_h\Pi_j]^{1-\alpha_h}$ where:

$$\Delta_j\Pi_h = \Pi_h(\mathcal{H}) - \Pi_h(\mathcal{H}\backslash j) = p_{hj}D_h(\mathcal{M}) - \Delta C_h(D_h(\mathcal{M})) \tag{28}$$

$$\Delta_h\Pi_j = \Pi_j(\mathcal{M}) - \Pi_j(\mathcal{M}\backslash h) = \Delta W_j(\mathcal{M}) - \Delta R(p_{\mathcal{M}j}) \tag{29}$$

[36] Different from their setup, I estimate the model at the insurer-hospital and insurer-system level using data on insurer enrollment instead of aggregating the estimating equation to the hospital level. I explain data and variable construction for each specification in detail in the next section.

In the above expressions, the change in hospital $h$'s profits, $\Delta_j \Pi_h$, is the difference between the additional revenues it will generate ($p_{hj} D_h(\mathcal{M})$) and the additional costs it will bear ($\Delta C_h(D_h(\mathcal{M}))$) by contracting with insurer $j$, as a result of the expected change in hospital visits that emerges from insurer $j$'s enrollees ($D_h(\mathcal{M})$). The change in insurer $j$'s profits by successfully negotiating a contract with hospital $h$, $\Delta_h \Pi_j$, is the difference between the change in WTP of its enrollees ($\Delta W_j(\mathcal{M})$) to have access to hospital $h$ and the change in insurer $j$'s total reimbursements ($\Delta R(p_{\mathcal{M}j})$) when hospital $h$ is included in its network, compared to the reimbursements when hospital $h$ is not available as an option and the enrollees visit other hospitals instead. Formally, $\Delta R(p_{\mathcal{M}j}) = \sum_{k \in \mathcal{M}} p_{kj} D_k(\mathcal{M}) - \sum_{k \in \mathcal{M} \setminus h} p_{kj} D_k(\mathcal{M} \setminus h)$.

Inserting these expressions into the objective function and taking the first order condition leads to the estimating equation:

$$\Delta_j \Pi_h = \alpha_h [\Delta W_j(\mathcal{M}) - \Delta C_h(D_h(\mathcal{M})) - \Delta R_j(p_{\mathcal{M}j}) + p_{hj} D_h(\mathcal{M})] \tag{30}$$

All the elements of equation (30) can be calculated from the previous parts of the structural model. Change in WTP, $\Delta W_j(\mathcal{M})$, is calculated as in equation (11), using results from hospital demand estimation. Change in costs, $\Delta C_h(D_h(\mathcal{M}))$, is calculated by predicting the change in hospital days using the hospital demand function, and then using these in hospital cost function to obtain predicted change in costs. Change in reimbursements of insurer $j$, $\Delta R_j(p_{\mathcal{M}j})$, are calculated by using the hospital demand model to predict how patients will re-allocate themselves to other hospitals if insurer $j$ fails to successfully contract with hospital $h$. Finally, change in hospital revenues, $p_{hj} D_h(\mathcal{M})$, are calculated by using the predicted change in hospital visits if a contract is reached, again using the estimates from the hospital demand model. The only unknown, and the parameter of interest, is therefore the bargaining power $\alpha_h$. I further parameterize $\alpha_h$ as follows in order to analyze how hospital bargaining power is determined:

$$\alpha_h = \alpha_0 + \beta H_h + \eta M_h + \epsilon_h \tag{31}$$

where $H_h$ are hospital characteristics such as FLF status, teaching status etc., and $M_h$ are market characteristics such as HHI measure of market concentration. Therefore, I take the following equation to estimation and use nonlinear least squares[37] to obtain the parameter estimates:

$$\Delta_j \Pi_h = (\alpha_0 + \beta H_h + \eta M_h)[\Delta W_j(\mathcal{M}) - \Delta C_h(D_h(\mathcal{M})) - \Delta R_j(p_{\mathcal{M}j}) + p_{hj} D_h(\mathcal{M})] \tag{32}$$

In the above equation, the term in brackets is the surplus generated by hospital $h$ and insurer $j$ successfully signing a contract. Identification in this model comes from the variation that identifies each individual component of surplus. In particular, parameters of the hospital demand model are identified by the variation in patients' choice sets across markets, while the parameters of the hospital cost function are identified by relating the variation in hospital costs to the variation in observable hospital input and output data. The parameters of the bargaining model are identified by relating the variation in change in hospital profits to predicted surplus that varies by the above-mentioned sources. Therefore, the parameters are mainly identified by variation in the data as opposed to the functional form.

# 6 Structural Analysis: Estimation Details and Results

## 6.1 Hospital Demand Estimation

The hospital choice model uses two data sources: patient characteristics come from SID for Arizona, Florida, Kentucky, New Jersey, New York, Rhode Island, and Washington while hospital characteristics come from AHA Survey. I estimate a conditional logit model where the utility specification[38] is given by:

$$u_{ihl} = \theta X_h + \lambda_1 X_h D_i + \lambda_2 X_h C_{il} + \epsilon_{ihl} \tag{33}$$

where $X_h$ is a vector of observed hospital characteristics, $D_i$ is a vector of demographic characteristics such as sex, age, location, and $C_{ilm}$ is a vector of diagnosis. One of the interaction terms $X_h D_i$ is the distance the patient travels to visit a hospital. In the model presented here, patients' choice sets are defined by ZIP codes. In particular, I put a hospital in a patient's choice set if another patient who lives in the same ZIP code visited that hospital.[39]

[37]Consumer price sensitivity parameter $\gamma$ is estimated via this equation and is a part of the WTP measure $\Delta W_j(\mathcal{M})$ as defined by Equation (11).

[38]While the original specification includes out-of-pocket costs (OPC), I exclude this term here as I do not observe it. The coefficient in front of OPC is estimated along with the bargaining parameters, see Table 11.

[39]I also estimated specifications where I constructed choice sets based on the Hospital Service Area (HSA) and Hospital Referral Region (HRR) of the hospital the patient visited and obtained similar results. In an ideal world, I would construct the choice sets based on the hospital network the patient's insurer offers. Unfortunately, I do not observe which individual is enrolled in which health plan in any of my datasets, therefore I cannot use this approach.

Table 9: Hospital Demand Estimation

| Variable | Odds Ratio | Variable | Odds Ratio |
|---|---|---|---|
| Distance (miles) | 0.97***<br>(0.0001) | Ultrasound | 1.42***<br>(0.02) |
| VI hospital | 1.05***<br>(0.005) | Orthopedic services | 1.02**<br>(0.01) |
| Distance x VI | 0.997***<br>(0.0001) | Birthing room x Female | 1.16***<br>(0.02) |
| Distance x Female | 0.987***<br>(0.0001) | Adult cardiac surgery<br>x Circulatory system | 3.67***<br>(0.09) |
| Distance x Age (18-34) | 0.98***<br>(0.0002) | Burn care x Burns | 42.35***<br>(5.43) |
| Nurses per bed | 1.24***<br>(0.003) | Neurological services<br>x Nervous system | 2.50***<br>(0.08) |
| Teaching hospital | 1.06***<br>(0.005) | Hemodialysis<br>x Kidney and urinary tract | 1.41***<br>(0.04) |
| Medical/surgical care | 1.29***<br>(0.02) | Oncology services<br>x Blood disorders | 2.14***<br>(0.24) |
| Chemotherapy | 1.18***<br>(0.008) | Obstetrics care<br>x Pregnancy and childbirth | 97.10***<br>(4.66) |
| Fertility clinic | 1.02***<br>(0.006) | Fertility clinic x Female | 0.96***<br>(0.006) |

*Notes:* Results from maximum likelihood estimation. N=1,152,081 hospital discharges from 7 states. See Table A2 for a full set of covariates included in the estimation. Standard errors in parentheses. *** statistically significant at 1% level, ** statistically significant at 5% level, * statistically significant at 10% level.

Table 9 presents a subset[40] of results from hospital demand estimation. Most hospital characteristics and services offered have positive coefficients that are highly significant. Same is true for the interaction terms. Consistent with the previous findings in the literature, I find that having to travel an extra mile to get treated at a hospital decreases the odds of that hospital being chosen by 3%.[41] Odds of being chosen is 1.05 times higher for a VI hospital compared to a non-VI one if both hospitals are at zero distance to the patient. This effect decreases with distance. For every standard deviation increase in distance, effect of VI hospital decreases by a factor of 0.997. In other words, the effect of a VI hospital that is one standard deviation away is 1.0468[42] which implies that a VI hospital is 4.68% more likely to be chosen over non-VI in this case. Finally, patients are more likely to visit hospitals that offer services that are related to their diagnosis. For example, the effect of an adult cardiac surgery unit for a patient diagnosed with a circulatory system disease is 3.7 times that of a patient who is not diagnosed with circulatory system disease, as anticipated.

### 6.2 Insurer Demand Estimation

The insurer demand model uses data at the national level. A market is defined as a state since health plans are observed to serve residents of specific states. An insurance plan is assumed to be a competitor in a market if it serves the residents of that state. The logit framework I use takes into account unobservable plan characteristics and is estimated via GMM. The utility function is of the form:

$$w_{ij} = \sum_k x_{jk}\beta_k + \xi_j + \epsilon_{ij} = \delta_j(x_j, \xi_j, \beta) + \epsilon_{ij} \tag{34}$$

where the observable insurer characteristics $x_j$ are VI insurer indicator, insurer premium per person per month (in \$100s), expected utility, age of the insurer, physicians per 100 population, Weiss rating, NCQA rating, NCQA accreditation, prevention quality measure, PPO indicator, BCBS indicator, and a large plan indicator.[43] In addition to these variables, the BLP specification includes interactions of expected utility and premiums with random draws to capture heterogeneity in individual preferences. Both specifications also include state fixed effects.

[40] For the full set of coefficient (not odds ratio) estimates, see Table A2.

[41] The figure reported is for a male patient aged between 55-64 who visits a non-VI hospital. The effect of distance decreases (becomes more negative) in various individual characteristics such as female and age.

[42] As calculated by $0.997 \times 1.05 = 1.0468$

[43] I define an insurer as large if it operates in multiple states. According to this definition, I mark Aetna, Anthem, BCBS, CIGNA, Humana, Kaiser Permanente, United Healthcare as large health insurers. Consumers perceptions about these plans are likely to be reflected in their preferences.

Since premiums are endogenous, I instrument for them using the average of characteristics of other plans ($x_n, n \neq j$) in the same market. These characteristics are Weiss rating, prevention, age, number of physicians, expected utility, and NCQA rating. These instruments satisfy the three traditional conditions of instrumental variables. They are relevant as they are correlated with premiums via competition and markups[44], they are uncorrelated with the error term, and they affect utility only through their impact on premiums. To further support the choice of the instruments, I analyze two statistics. The first stage results report a partial R-squared of 0.72 and an F-statistic of 17.46. These statistics suggest a large portion of the unexplained variation in premiums comes from the excluded instruments and the instruments are not weak since the F-statistic is greater than 10.[45]

To complete the estimation, the last element needed is the share of the outside good. Since I observe HMO/POS and PPO/indemnity plans in my data, I define the outside good as being uninsured. U.S. Census reports number of uninsured and state population by age group. Therefore, I calculate the share of the outside good, $s_0$, by dividing the number of nonelderly uninsured by nonelderly population of that state.

The parameter estimates are reported in Table 10. The first column reports results from Berry (1994) specification while the second column presents results from the BLP estimation. The coefficient in front of premiums is negative and significant as expected. Its magnitude implies an insurer-perspective elasticity of -2.11.[46] This suggests that a $10 increase in monthly premiums per enrollee decreases the demand for that insurer by 6%. The expected utility coefficient is positive and significant, implying people value the hospital network offered by an insurer while making their choices. The parameter estimate for VI insurer indicator is positive, however not statistically significantly different than zero in the BLP specification. A plausible explanation is that consumers evaluate characteristics such as premium, hospital network, physician network, or quality as opposed to ownership structure while selecting their health plans. Therefore, the favorable and unfavorable aspects of VI plans are likely captured by the remaining covariates in the estimation. As logit and BLP specifications give similar coefficients in terms of sign, magnitude, and significance; I use the estimates from the BLP model in counterfactual estimations as this model better captures individual heterogeneity in preferences and creates more realistic substitution patterns.

[44]This relationship is implied by the first order conditions in the supply side that leads to the pricing Equation (38).
[45]See Bound et al. (1995).
[46]This is the elasticity of demand with respect to premiums from insurer's perspective, as opposed to the elasticity from the consumer's perspective that is based on out-of-pocket expenditures.

Table 10: Insurer Demand Estimation

| | (1) | (2) |
|---|---|---|
| Premium ($00) | -0.63* | -0.22*** |
| | (0.37) | (0.05) |
| NCQA rating | 0.72 | -0.33 |
| | (0.55) | (0.36) |
| NCQA accreditation | 0.63 | 0.46 |
| | (0.42) | (0.39) |
| VI-insurer | 0.96** | 0.19 |
| | (0.48) | (0.44) |
| FLF-insurer | 0.22 | 0.76 |
| | (0.33) | (0.78) |
| Age | -0.001 | 0.007* |
| | (0.005) | (0.004) |
| Weiss rating | -0.10*** | -0.11*** |
| | (0.03) | (0.04) |
| Physicians | 0.008*** | 0.001*** |
| | (0.0003) | (0.0002) |
| Prevention | -0.94** | -0.15 |
| | (0.44) | (0.30) |
| Expected utility | 0.46*** | 0.49*** |
| | (0.03) | (0.03) |
| PPO | 1.10* | -0.001 |
| | (0.66) | (0.26) |
| Large Insurer FE | Yes | Yes |
| BCBS FE | Yes | Yes |
| State FE | Yes | Yes |

*Notes:* Results from GMM estimation. N=989 insurers from 50 states and Washington DC. Robust, clustered (at the state level) standard errors in parentheses. First column follows Berry (1994), second column follows BLP (1995). Physicians are per 100 enrollee population. *** statistically significant at 1% level, ** statistically significant at 5% level, * statistically significant at 10% level.

## 6.3 Bargaining

The estimating equation for the bargaining model is:

$$\Delta_j \Pi_h = (\alpha_0 + \beta H_h + \eta M_h)[\Delta W_j(\mathcal{M}) - \Delta C_h(D_h(\mathcal{M})) - \Delta R_j(p_{\mathcal{M}j}) + p_{hj} D_h(\mathcal{M})] = \alpha_h \times \Delta S_{hj}(\mathcal{M}) \quad (35)$$

where $\Delta\Pi_h$ are the additional profits earned by hospital $h$ when it successfully signs a contract with insurer $j$, the first term in parentheses is the decomposition of hospital bargaining power, while the second term in brackets is the total surplus generated by insurer $j$ and hospital $h$ successfully negotiating a contract. In this model, hospital $h$'s bargaining power is determined by market characteristics $M_h$ such as hospital market HHI, as well as its own characteristics $H_h$ such as FLF status, VI status, teaching status, rural status, market share, integration with a physician group, system membership etc. The elements of total surplus are each calculated using the models in the previous steps, and their calculations are detailed below. In an ideal world, the bargaining model would be estimated using data on negotiated prices between each hospital-insurer pair, the insurance plan the individual is enrolled in, and the hospital he/she visited. While I observe the hospital visited by each patient in my data, I do not observe what health plan he/she is enrolled in. Lewis and Pflum (2015) encounter the same problem, and aggregate the bargaining equation to the hospital level, and estimate the following equation at the hospital level:

$$\sum_{j \in \mathcal{H}} \Delta_j \Pi_h = \alpha_h \times \sum_{j \in \mathcal{H}} \Delta S_{hj}(\mathcal{M}) \quad (36)$$

where every term is summed across the set of insurers the hospital contracts with, $\mathcal{H}$. They also use assumed shares from insurers at a hospital when calculating individual elements of surplus. I extend their setup by using data on insurers and estimating the bargaining model at the hospital-insurer level. Since I observe insurer networks as well as market shares, I disaggregate the hospital level components of surplus into hospital-insurer level by using the share of insurer for each hospital. For example, in the model outlined above, $\Delta W_j(\mathcal{M})$ is the total WTP of the enrollees for hospital $h$ to be included in insurer $j$'s network. Following equation (36), I first calculate the total WTP of *all* patients in the market for hospital $h$. Then, I split this WTP measure into WTP for hospital $h$ of each insurer, based on market shares.[47] Therefore, if the share of patients at hospital $h$ from insurer $j$ is $s_{jh}^{ins-hosp}$, then the WTP of enrollees of $j$ for hospital $h$ is $\Delta W_h \times s_{jh}^{ins-hosp}$ where $\Delta W_h$ is the total WTP for hospital $h$.

[47]Calculation of the share of patients at a hospital from an insurer is outlined in the appendix.

In what follows, I discuss the calculation of each element of surplus and then discuss results from the bargaining model. I use data from California OSHPD to estimate the bargaining model and obtain welfare results. One thing to note is that hospital demand model forms the basis of components in bargaining model, therefore I start by estimating hospital demand. Different from the hospital demand model discussed above, I use choice sets based on Hospital Referral Regions (HRRs)[48] and omit the individual-level variables.[49] I estimate this demand model with the same set of SID states, and then use the parameter estimates to predict what hospital shares would be in California. I cannot adopt the alternative approach to create choice sets based on ZIP codes as patient ZIP codes are not reported in California data.[50]

For each component of total surplus, I follow calculations in Lewis and Pflum (2015) to obtain hospital level aggregate measures. The basis of the first element, $\Delta W_h$, is the total WTP for hospital $h$. I calculate the WTP of each individual for each hospital as $ln(1/(1-\hat{s_{ih}}))$ where $\hat{s_{ih}}$ is the predicted probability that patient $i$ will choose hospital $h$. Then, I sum this across individuals for every hospital to obtain $\Delta W_h$. Then, the WTP of enrollees of insurer $j$ for hospital $h$ to be included in $j$'s network is calculated as $\Delta W_j(\mathcal{M}) = \frac{1}{\gamma} \times \Delta W_h \times s_{jh}^{ins-hosp}$ following Equation (11).

The second element of surplus, $\Delta C_h(D_h(\mathcal{M}))$, is the expected change in hospital costs when hospital $h$ joins insurer $j$'s network. This term is calculated using both hospital demand model and hospital cost function. First, I estimate the hospital cost function where hospital costs depend on hospital inputs and outputs such as inpatient days from private payers, inpatient days from Medicare etc. Next, using results from hospital demand, I predict private[51] inpatient days at a hospital ($predDays$) by multiplying predicted probability of choice with length of stay, and summing it across private patients for a particular hospital. Finally, using parameter estimates from hospital cost function, I predict hospital costs using two output measures: $predDays$ and $(1 - s_{jh}^{ins-hosp}) \times predDays$. The difference between the two predicted costs gives the change

[48] I also tried using the estimates from choice sets based on the narrower geographic measure, Hospital Service Areas (HSAs). Since many of the HSAs have only 1 hospital, these choice sets needed to be dropped while calculating change in insurer reimbursements when an agreement is not reached, as this calculation requires patients to re-allocate themselves to other hospitals in their choice sets. Moreover, having a single hospital in a choice set is not credible as many patients choose from multiple hospitals as opposed to one. Therefore, I conducted my analysis using HRR choice sets.

[49] Bargaining model uses data from California OSHPD. While previous papers in the literature used individual-level variables from the same data source, OSHPD stopped reporting these variables to protect patient confidentiality starting with 2012 data. Therefore, I am unable to use terms involving sex and age in hospital demand estimation and prediction, as these are unavailable in the California data.

[50] Starting with 2012 data, OSHPD only reports the first 3 digits of a patient's ZIP code. Based on this, I randomly assign individuals to ZIP codes starting with those 3 digits based on ZIP codes' population weights, and use this to calculate distance to be used in prediction. However, as these ZIP code assignments are not precise, I refrain from using them as choice sets in hospital demand estimation. Estimating hospital demand using exact distances from SID gives me more accurate coefficients for the covariates, and especially for distance.

[51] As I only work with commercial insurers in my data, I only consider changes in the private line of business.

in costs at a hospital if it contracts with insurer $j$.

The third component of surplus, $\Delta R_j(p_{\mathcal{M}j})$, is the change in reimbursements of an insurer if it does not include hospital $h$ in its network and its patients re-allocate themselves to the remaining hospitals.[52] The outline of the calculation is as follows. For every hospital $h$, I remove $h$ from the market, focus on the choice set affected by removal of $h$, re-assign $h$'s patients to other hospitals in this choice set based on the demand model, and then calculate extra revenues at these hospitals. Finally, I get an extra revenue per hospital which is equivalent to extra reimbursement the hospital gets from all the insurers it contracts with. I then decompose this to hospital-insurer level based on shares $s_{jh}^{ins-hosp}$. The detailed calculation of these steps is as follows. First, for every hospital that is removed from the market, I calculate predicted extra days at other hospitals that are in the same choice set. If hospital $h$ is removed from the choice set, then the predicted extra days at hospital $h'$ are calculated as: $extraPredDays_{h'} = \left[\frac{prob_{h'}}{1-prob_h} - prob_{h'}\right] \times LOS_h$.[53] In this expression, *prob* are original choice probabilities based on hospital demand estimation and $LOS_h$ is the patient's length of stay at hospital $h$ as reported by OSHPD. The term in brackets represents the *increased* choice probability of hospital $h'$ while $LOS_h$ represents the extra days available with the removal of hospital $h$. In the next step, I use OSHPD Discharge and Financial Reports to estimate the average revenue per inpatient day at a hospital for each MDC for a non-ER private patient, and then use it to calculate extra revenues at a hospital. For example, average revenues for a patient diagnosed with MDC category 5 that is admitted through non-ER is calculated as:

$$\text{Avg.Rev./Day} = \frac{\text{Net Revenues from Private Payers}}{\text{Gross Charges for Private Payers}} \times \frac{\text{Total IP Charges for Private-non-ER in MDC5}}{\text{Total IP Days for Private-non-ER in MDC5}}$$

Given the average revenue measure, total extra revenues at a hospital are calculated by multiplying predicted extra days with the average revenues that correspond to that MDC, and then aggregating it to the hospital level. This hospital level measure is the total reimbursements from all insurers the hospital contracts with. Therefore, to break it down to hospital-insurer level, I multiply it with $s_{jh}^{ins-hosp}$ and use this final term in estimation.

[52] In an ideal world, I would calculate this component using insurer networks as choice sets. However, I do not observe which individual is enrolled in which health plan, so I am forced to use HRRs as my choice sets. While the re-assignment of patients to other hospitals in patient choice sets would be more precise with insurer networks as choice sets, since the final product is calculated at the hospital level and then decomposed to hospital-insurer level based on $s_{jh}^{ins-hosp}$, I assume the hospital level aggregated measure is close when using HRR choice sets to what it would be if I used insurer networks as choice sets. The choice sets in Lewis and Pflum (2015) are also not based on insurer networks (except top five largest HMOs), and they conduct a similar analysis to obtain the change in reimbursements of insurers using ZIP code choice sets. Finally, I drop choice sets with only 1 hospital as patients are unable to re-allocate themselves if they belong in these choice sets.

[53] This expression is calculated at the individual level and then aggregated to hospital-MDC level at each iteration.

The final component of surplus, $p_{hj}D_h(\mathcal{M})$, is the change in revenues of a hospital when it contracts with an insurer. I use the same average revenue measure in lieu of the reimbursement price vector $p_{hj}$ and calculate the change in expected demand for the hospital $D_h(\mathcal{M})$ based on hospital demand estimation. In particular, I start at the individual level and calculate $predDays$ as above (by multiplying hospital choice probability with length of stay of each private individual), then predict the revenues from these days by MDC using the average revenue measure, and finally aggregate these revenues to the hospital level. This aggregate measure is again broken into hospital-insurer level using $s_{jh}^{ins-hosp}$. Finally, the left hand side variable, $\Delta_j \Pi_h$, is calculated by subtracting the change in costs discussed above from the total extra revenue generated.

I estimate two specifications of the bargaining model. The first specification is estimated at the hospital-insurer level where the negotiating hospital unit is a hospital. This specification assumes hospitals negotiate with insurers individually and not as systems. Belonging to a hospital system can still improve a hospital's bargaining position in this specification through the system membership variable. In the second specification, the negotiating hospital unit is a hospital if the hospital is a member of a non-FLF system or an individual hospital, and a system otherwise. In other words, I only allow FLF-offering systems to negotiate as a system. In this dataset, binary hospital characteristics are calculated as a fraction when the negotiating unit is a system.[54] Hospital and system shares used in HHI are calculated based on hospital visits in that hospital/system and in that HRR. While constructing the components of the bargaining equation, I calculate each variable at the hospital level first, and then aggregate to the system level by summing across member hospitals if the hospital belongs to an FLF-offering system. The disaggregation from system level to system-insurer level variables is again done by using $s_{jh}^{ins-hosp}$, which in these instances represent the share of patients at system $h$ coming from insurer $j$.

The two specifications represent the two extremes of the health care market I observe in the data. In practice, not all hospitals negotiate individually and independently with insurers, as some systems tie their hospitals together and negotiate as a system. However, it is also true that not all systems try to impose full-line of their products to all insurers they negotiate with. The system level bargaining model in this paper improves upon the previous bargaining estimations in the literature that assume *all* hospital systems negotiate as systems[55] by identifying FLF-offering systems as negotiating units. Nonetheless, while the hospitals in these

[54]For example, if 2 out of 4 system hospitals are teaching hospitals, the teaching indicator variable for that system is 0.50.
[55]See Gowrisankaran et al. (2015) and Lewis and Pflum (2015).

Table 11: Determination of Bargaining Power

| *Dep. var.:* $\Delta\Pi_h$ | (1) Individual Level | (2) System Level |
|---|---|---|
| Base bargaining power | 0.98*** | 0.80** |
| | (0.11) | (0.38) |
| VI-hospital | 0.53*** | 0.02** |
| | (0.09) | (0.01) |
| FLF-hospital | -1.19** | -0.14 |
| | (0.54) | (0.34) |
| Hospital share | 1.28*** | 0.11** |
| | (0.47) | (0.05) |
| Hospital market HHI | -2.78*** | 0.57 |
| | (0.81) | (1.13) |
| Predicted days | 0.002 | 0.0003*** |
| | (0.004) | (0.00002) |
| Teaching hospital | -0.34*** | -0.03** |
| | (0.09) | (0.02) |
| Rural hospital | -0.13 | -0.02 |
| | (0.09) | (0.02) |
| For-profit hospital | 0.02 | 0.01 |
| | (0.10) | (0.01) |
| Physician group | 0.04 | -0.48*** |
| | (3.56) | (0.08) |
| System member | 1.33** | 0.14*** |
| | (0.55) | (0.02) |
| $\gamma^{-1}$ (x1000) | 1.24*** | 0.83*** |
| | (0.10) | (0.08) |
| *Bargaining Power:* | | |
| Mean fitted value | 0.69 | 0.63 |
| Standard deviation | 0.19 | 0.20 |
| $N$ | 4936 | 2163 |
| Adjusted $R^2$ | 0.39 | 0.95 |

*Notes:* Results from nonlinear least squares estimation using California data. Individual level sample includes 276 hospitals and 32 insurers, system level sample includes 116 systems/hospitals and 32 insurers. Both specifications include HRR fixed effects. Predicted patient days in thousands. *** statistically significant at 1% level, ** statistically significant at 5% level, * statistically significant at 10% level.

systems are often tied together, they also negotiate individually with some insurers, which is not captured by the system level estimation.

Results from bargaining estimation are presented in Table 11. The coefficient estimates are similar in terms of sign and significance across two specifications. The main finding is that VI hospitals have significantly higher bargaining power in the negotiation process compared to their non-VI counterparts. System membership also increases a hospital's bargaining power, consistent with the previous findings in the literature. At the individual level specification, the positive impact of being a system member on a hospital's bargaining power is greater in terms of magnitude compared to the system level specification. This is as anticipated because system membership is one of the few variables that capture the advantages of being in a hospital system in this specification that does not force systems to tie their products. The bottom part of Table 11 reports the mean fitted values of hospital bargaining power. Like previous papers in the literature, I find that hospitals, on average, have higher bargaining power than insurers in the negotiation process. These results are robust to dropping the VI hospital-insurer pairs from the dataset. These pairs likely do not engage in bargaining since they own one another. I also estimated specifications where I allowed hospital bargaining power to depend on the characteristics of the insurer it contracts with. This approach did not improve the fit of the model and insurer characteristics never proved to be informative in explaining hospital's bargaining power. Given the results in Table 11, I conclude the system level regression better models the market as is, given the better fit and more sensible coefficients such as HHI. Results from this specification are used in counterfactual simulations.

## 7 Welfare Analysis

In this section, I simulate a counterfactual scenario to assess the impact of vertical integration on welfare. In this policy experiment, I remove vertical integration from the market altogether, force previously-VI entities to become independent hospitals and insurers, and measure the change in consumer welfare and producer surplus. It is important to note that these counterfactual simulations are "partial" in the sense that they do not take into account all market responses such as changes in *all* product characteristics, due to the complexity and infeasibility of calculating an equilibrium in both hospital and insurer markets in all dimensions.[56] In particular, I assume that hospital services, staffing, reimbursement rates, governance,

[56]Other papers in the literature use the same approach and keep many attributes constant while estimating the impact on welfare. See, for example, Lee (2013), Gowrisankaran et al. (2015), Ho (2006) among others.

system membership, and insurer quality measures do not change when vertical integration is removed from the market.

After I remove vertical integration from the market, I re-form the hospital networks offered by insurers based on the bargaining model, predict new premiums for insurers under new networks, and calculate the change in consumer welfare and producer surplus. When forming new networks, I use the system level bargaining model to isolate the effect of vertical integration. Upon removal of vertical integration from the market, the previously-VI hospitals act as new players in the bargaining game, but other FLF-offering system hospitals keep negotiating as systems. Using individual level bargaining model would imply allowing both VI and FLF hospitals to negotiate individually, and the counterfactual scenario in this case would capture the combined effect of removal of VI and FLF from the market together. As the purpose here is to understand how VI affects welfare, I only break the VI entities and allow FLF systems to impose full-line of their products. I estimate two counterfactual scenarios as I impose two different assumptions of network formation. In the first one, I allow previously-VI systems to negotiate as systems upon removal of vertical integration. The second counterfactual assumes the previously-VI hospitals will act as individual hospitals and individually negotiate with insurers in the market. A potential policy change that bans vertical integration is not likely to impose constraints on how these systems should negotiate their hospitals (individually or as a system). A plausible scenario is that these VI system hospitals will negotiate as systems with some insurers and individually with others. Hence, the welfare results based on two counterfactuals in this paper should be interpreted as the upper and lower bounds of change in welfare.

Below, I explain welfare calculation in detail. To begin, I re-form insurer networks by following the steps below:

1. Re-predict hospital choice probabilities. First, I remove vertical integration from the market by setting VI dummy to zero for all previously-VI entities. The first counterfactual assumes previously-VI hospitals negotiate as systems, the second counterfactual assumes previously-VI hospitals negotiate individually. In both cases, I re-construct the data accordingly. I also update hospital share and hospital market HHI to reflect the new negotiating units in the data. Finally, I predict new probabilities using hospital demand model. These new shares imply new $predDays$ and $extraPredDays$ at a hospital (or system).

2. Using the new probabilities from the first step, re-calculate every element of the bargaining model. In

particular, I obtain new values for the hospital level variables $\Delta W_h(\mathcal{M})$, $\Delta C_h(D_h(\mathcal{M}))$, $\Delta R_h(p_{h\mathcal{H}})$, and $p_{h\mathcal{H}}D_h(\mathcal{M})$.

3. Break hospital (or system) level variables into hospital-insurer level variables. In the first iteration, I begin by assuming all hospitals/systems contract with all insurers, and do the decomposition using $s_{jh}^{ins-hosp}$ calculated under this assumption.

4. Predict changes in hospital profits, $\Delta_j\Pi_h$, for every hospital-insurer pair by using parameter estimates from the bargaining model, setting VI dummy to zero for all previously-VI entities, and using the re-calculated elements of surplus from step 3. Hospital (or system) $h$ chooses to sign a contract with insurer $j$ if $\Delta_j\Pi_h > 0$, hence new networks are formed.

Given the new networks, I go back to step 3 and re-calculate shares at a hospital from an insurer following the share calculation outlined in the appendix that accounts for insurer networks while imposing capacity constraints. I keep iterating using this routine until equilibrium is reached. At this equilibrium, no hospital has an incentive to deviate and change the set of insurers it contracts with. This routine relies on the assumption that if a hospital is willing to join an insurer's network, the insurer will agree to establish a contract. This assumption is supported by both theory[57] and data.[58]

Next, I calculate new premiums given the new insurer networks. Insurers maximize a standard profit function by choosing premiums:

$$\pi_j = (prem_j - C_j \times p_{ip}) \times s_j \times M \tag{37}$$

where M is market size, $C_j = \sum\limits_{h\in\mathcal{H}} \frac{s_h}{\sum\limits_{k\in\mathcal{H}} s_k} \times p_h$ is the average reimbursement per day by insurer $j$ to the set of hospitals $\mathcal{H}$ in its network, and $p_{ip}$ is the number of inpatient days at a state divided by total population. Maximizing the objective function leads to the following first order condition that determines premiums:

$$s_j + (prem_j - C_j \times p_{ip})\frac{\partial s_j}{\partial prem_j} = 0 \tag{38}$$

Finally, given the new networks and new premiums, I calculate the change in producer surplus and consumer welfare. Producer surplus is the aggregation of all hospital and insurer profits in the market. Hospital profits

[57]Capps et al. (2003) present a model of network formation which shows that the profit maximizing strategy for an insurer is to include all the hospitals in the market in its network.

[58]Lewis and Pflum (2015) state that their communications with a former contract negotiator for a major national insurer revealed PPOs' strategies are to include almost every hospital in the market in their network. In their data, the median HMO covers 84% of hospitals in the market. Ho (2009b) reports, on average, 87% of hospital-HMO pairs successfully sign a contract. In my data, the median HMO and PPO cover 42% and 60% of the hospitals in the market, respectively. Therefore, I assume both HMOs and PPOs aim to cover a substantial portion of the market where possible.

are calculated as: $\pi_h = (p_h \times p_{ip} \times \sum_{j \in \mathcal{M}} Ms_{jh}) - Cost_h$ where $p_h$ is the average revenue per inpatient day at hospital $h$, and $Cost_h$ is the predicted cost of the hospital given new predicted days. Insurer profits are calculated using equation (37).

I use compensating variation (CV) to measure the change in consumer welfare when vertical integration is removed from the market. Compensating variation refers to the amount of money a consumer would need to give up following a change in prices or product quality (hospital networks) in order to reach his pre-change utility level. Following Small and Rosen (1985), the compensating variation for consumer $i$ is given by:

$$CV_i = -\frac{1}{\alpha_i}\left[ln \sum_j exp(V_{ij}^{post}) - ln \sum_j exp(V_{ij}^{pre})\right] \tag{39}$$

where the superscripts *post* and *pre* refer to the removal and presence of vertical integration, respectively. $-\alpha_i = -(\beta_2 + \gamma_2 v_{i2})$ is the negative of the premium coefficient and $j$ still represents an insurer. $V$ is the observed portion of utility defined as:

$$V_{ij} = \xi_j + x_j\hat{\phi} + \hat{\beta}_1 EU_j + \hat{\beta}_2 prem_j + \hat{\gamma}_1\nu_{i1}EU_j(H_j) + \hat{\gamma}_2\nu_{i2}prem_j \tag{40}$$

Compensating variation is then the market size times the integral of compensating variation over the distribution of $\nu$ as given by:

$$CV = M\int CV_i dP_\nu(\nu) \tag{41}$$

where $M$ is market size. Applying this to the random-coefficients model, I calculate the compensating variation by simulation. In particular, I calculate compensating variation for each draw of $\nu$, and then take the average across these $ns$ draws to obtain:

$$CV = M\frac{1}{ns}\sum_{i=1}^{ns} CV_i \tag{42}$$

Counterfactual results are reported in Table 12. I simulate two counterfactual scenarios using California data. Under the first scenario, previously-VI hospitals negotiate as systems upon removal of vertical integration from the market. Results from this counterfactual are reported in the first part of Table 12, labeled "System Level". Under the second counterfactual scenario, I allow previously-VI hospitals to negotiate as individual hospitals. Results from this estimation are reported in the second part of Table 12, labeled "Individual Level". In both counterfactuals, the non-VI FLF hospital systems in the market continue to negotiate

Table 12: Counterfactual Results

| | System Level | | | | Individual Level | | | |
|---|---|---|---|---|---|---|---|---|
| | All | VI | Non-VI | Entrant | All | VI | Non-VI | Entrant |
| **Δ CS:** | **$ -9.5b** | | | | **$ -1.9b** | | | |
| Premiums | 97% | -18% | 124% | 16% | 94% | -34% | 123% | 3% |
| Network Size | 53% | 50% | 54% | 55% | 42% | 39% | 42% | 45% |
| **Δ PS:** | **$ 44b** | | | | **$ 16.8b** | | | |
| Pct hosp. at loss | 34% | 36% | 33% | | 50% | 43% | 54% | |
| Hosp. Π lost | $ -22.8b | $ -4.8b | $ -18b | | $ -35.7b | $ -8.9b | $ -26.8b | |
| Pct hosp. at gain | 66% | 64% | 67% | | 50% | 57% | 46% | |
| Hosp Π gained | $ 65b | $ 26.8b | $ 38.2b | | $ 49b | $ 21b | $ 28b | |
| Pct insurers at loss | 41% | 50% | 38% | | 41% | 67% | 35% | |
| Ins. Π lost | $ -3b | $ -2.25b | $ -750m | | $ -1.3b | $ -706m | $ -543m | |
| Pct insurers at gain | 59% | 50% | 62% | | 59% | 33% | 65% | |
| Ins Π gained | $ 4.9b | $ 16.6m | $ 4.85b | $ 128m | $ 4.7b | $ 2.5m | $ 4.7b | $ 8.9m |

*Notes:* Results from counterfactual simulations where vertical integration is removed from the market. Members of previously-VI hospital systems negotiate as systems under the system level simulation; as individual hospitals under the individual level simulation. Entrant is a non-VI insurer.

as systems. The baseline model in both counterfactuals is the predicted world where VI is present. In this baseline specification, VI hospitals are only included in their own insurer's network, and the remaining network formation in the market is done based on the bargaining model. I present results for previously-VI entities, non-VI entities, and the only entrant to the downstream market which is non-VI. In the first panel where I report sources of change in consumer welfare, percentages represent average changes compared to vertical integration being present in the market. In the second panel, percentages represent the split within their own categories defined by VI status and entity type.

I find that removal of vertical integration from the market decreases consumer welfare by $9.5 billion a year. When exclusionary vertical restraints are abolished, many previously-VI hospitals join other insurers' networks. This leads to formation of larger insurer networks that increase expected utility and benefit consumers. Previously-VI insurers lose exclusive access to their own hospitals, and they compete with the rest of the market by lowering their premiums. Non-VI insurers, on the other hand, increase their premiums on average following the expansion in their networks. The increase in consumer utility induced by larger networks is offset by the increase in premiums, leading to a decline in welfare. Increase is network size is larger in the system level counterfactual as the new players are included as systems in insurer networks. Under the individual level counterfactual, both network size and premiums increase by less on average because not all system members are included in insurer networks. While the result that removal of vertical integration harms consumer welfare remains the same, the magnitude of change drops to \$1.9 billion a year.

My findings further indicate that total producer surplus in the market increases by \$44 billion when vertical integration is banned and previously-VI hospitals negotiate as systems. Majority of this gain comes from the hospital industry. Among the previously-VI hospitals, 36% lose profits while 64% increase profits. The gain in profits is substantially larger compared to the loss (+\$26.8 billion vs. -\$4.8 billion) when previously-VI hospitals are considered. This is mainly due to the fact that hospital visits increase as now they are included in a larger number of insurer networks. While the predicted hospital costs also change under the new networks (as a result of new *predDays*), the main driver of change in hospital profits is change in shares. Overall, the insurer market benefits from removal of vertical integration. The major increase in profits is achieved by previously-non-VI insurers who now have access to previously-VI hospitals which are cheaper.[59] The majority of loss is born by previously-VI insurers, as expected. These insurers lose \$2.25 billion in profits as they lower premiums and no longer enjoy the dedicated market share to them from their own hospitals. When previously-VI hospitals negotiate individually, producer surplus increases by \$16.8 billion a year. Patterns for this individual level counterfactual are very similar to the ones discussed above, although the magnitudes of change are smaller. Since VI hospitals no longer negotiate as systems, the increased market share obtained by tying hospitals together is now absent, resulting in smaller changes in the same direction.

I also find entry barriers are present in the market due to upstream foreclosure.[60] By definition, entry barriers are cost advantages the incumbent enjoys compared to the entrant (Stigler (1968)). In the case of hospital-insurer consolidation, the incumbent VI entities have cost advantages over non-VI entrants as they have access to less expensive hospitals while the entrants do not.[61] When vertical integration is removed

[59] Reimbursement rates in this paper are approximated by average revenue per day per MDC category, as outlined in the previous section. This measure is lower for VI hospitals for 23 out of 25 MDC categories. The average revenue per day measure is 39% lower for VI hospitals, on average, compared to non-VI hospitals.

[60] Downstream foreclosure also leads to entry barriers as insurers prevent hospitals from accessing their markets. However, I only observe entrants in the insurer dataset, hence, I only investigate barriers to entry in the downstream market.

[61] The incumbent VI insurers have cost advantages over non-VI entrants also because they have an already-formed network (which is costly for the entrant insurer to form) and their administrative businesses are settled (for example, entrant insurer needs to find compatible billing systems with the hospitals it contracts with while the VI incumbent can use the same system for its health plan and hospitals). In practice, these cost advantages also act as barriers to entry. Since the incumbent VI entities face lower costs, they are able to charge lower premiums. The non-VI insurers who want to enter the market face higher costs, but they also need to match the low premiums offered by incumbent VI insurers. Therefore, not all the entrants who desire can enter the market. However, my model does not explain these entry barriers as it does not account for network formation costs and administrative costs in counterfactual simulations. As outlined above, the model assumes that the only costs to insurers

and the non-VI entrant is allowed to access the same cost advantages, the entrant is better off. I find that the entrant non-VI insurer increases its premiums by 16% and its network size by 55% by switching to previously-VI hospitals. Since the entrant's market share and premiums increase, its profits go up by $128 million. Figures for individual level counterfactual are again smaller in magnitude, but the same patterns are observed. Therefore, when entry barriers are removed from the market, the current entrant is better off. This result suggests that with reduction in barriers to entry, previously excluded firms may now be able to join the market.

## 8 Conclusion

This paper investigates the impact of vertical integration on market outcomes and welfare in the health care industry. Theory suggests vertical integration can lead to both efficiency gains and market foreclosure, hence the overall impact on welfare is ambiguous. I focus on VI entities that are formed by vertical mergers between hospitals and insurers. In most cases, hospitals and insurers come together to offer their own health plan that restricts enrollees to the parent organization hospitals. To compensate for the restricted choice, these plans promise lower premiums and higher quality.

I conduct reduced-form analysis to examine the relationship between vertical integration and market foreclosure, prices, and quality. I find that VI insurers indeed offer lower-premium, higher-quality plans controlling for market and product characteristics. Results from estimation of a structural model of consumer demand and hospital-insurer bargaining suggest that VI hospitals have higher bargaining power in the negotiation process compared to their non-VI counterparts. Overall, consumers benefit from vertical integration, although majority of hospitals and insurers are better off in its absence. I also find that the presence of vertical integration in the market serves as an entry barrier to the downstream market, mainly due to upstream foreclosure exercised by VI system hospitals.

While this paper quantifies the welfare impacts of vertical integration for consumers and producers, it does not fully model market responses in all product characteristics. Future work in this area should analyze how reimbursement rates would change upon removal of vertical integration, and how the market would respond to these price changes. Improving the model in this way would require estimating a bargaining model using

are reimbursements to hospitals for the patients who get sick, therefore, any cost advantages gained will be through formation of new networks that include less expensive hospitals.

administrative claims data, ideally at the national level. Another direction for future research would be to estimate a dynamic entry model to fully identify the barriers to entry caused by vertical restraints in the health care industry. Such analyses will better inform the policy makers with regard to the potential impacts of vertical integration over time.

# 9 Appendix

## 9.1 Enrollment in VI plans

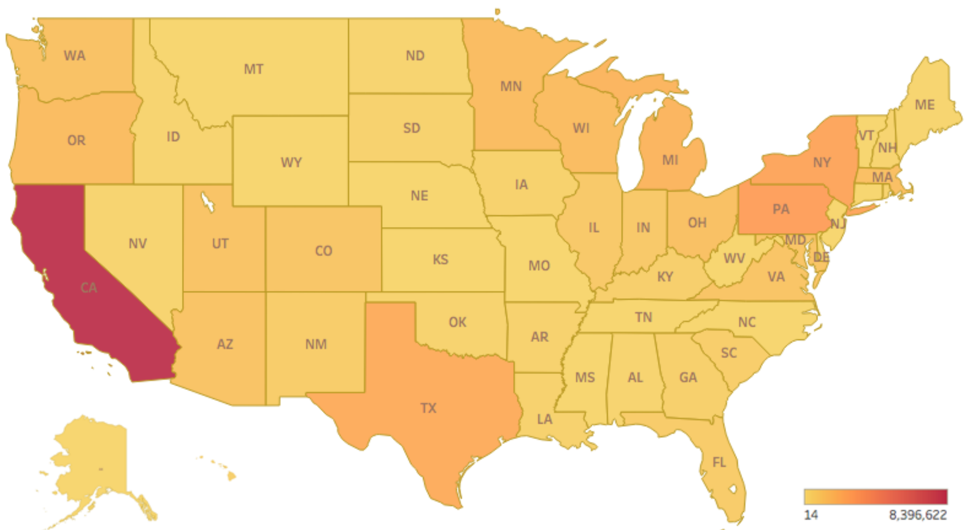


Figure 1: Enrollment Distribution by State

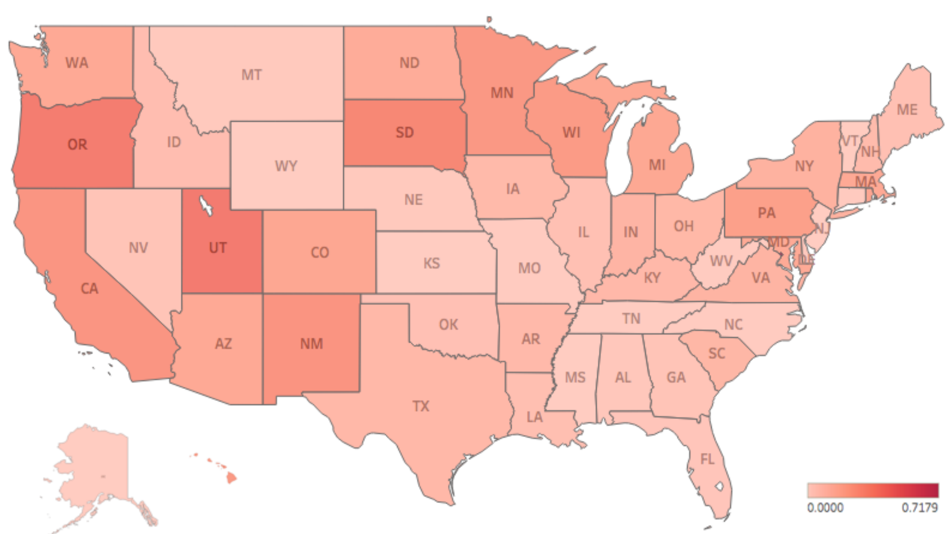


Figure 2: Enrollment as a Percentage of Insured Population

## 9.2 Share of patients at a hospital from an insurer

The share calculation is done based on hospital shares, insurer shares, and capacity constraints. Let $M$ be market size, $cc_h$ be the capacity constraint of hospital $h$, and $s_h$ and $s_j$ represent market shares of hospital $h$ and insurer $j$, respectively. Capacity constraints are defined by the number of beds at a hospital, and hospitals are assumed to operate at full capacity when the constraints are imposed. To begin with, I allocate patients to hospitals based on hospital networks offered by insurers, hospital shares, and insurer shares. Share of patients enrolled in plan $j$ that visit hospital $h$ is calculated as:

$$s_{jh} = s_j \times \frac{s_h}{\sum\limits_{h \in K} s_h} \tag{43}$$

where $K$ is the set of hospitals insurer $j$ contracts with. However, when the allocation is done this way, some hospitals exceed their capacity limits. If this is the case, I remove excess patients from that hospital and re-allocate them to remaining hospitals in the market who have not reached their capacity constraints at this stage, based on shares $s_{jh}$. Following this re-allocation, I calculate shares $s_{jh}^{ins-hosp}$.

Formally, if hospital $h$ exceeds its capacity, then I remove $E_j$ many patients from insurer $j$'s allocated patients to hospital $h$, where $E_j$ is given by:

$$E_j = \left[\left(M \times \sum_{j \in H} s_{jh}\right) - cc_h\right] \times \frac{s_{jh}}{\sum\limits_{j \in R} s_{jh}} \tag{44}$$

In this equation, $H$ is the set of insurers hospital $h$ contracts with and $R$ is the set of insurers contracted with hospital $h$ who will re-allocate a fraction of their patients to other hospitals.[62] First term in brackets is the number of excess patients at hospital $h$, whereas the second term represents the share of insurer $j$ among insurers who will re-allocate their patients from hospital $h$. Re-allocation of these removed patients to other hospitals is done based on insurer shares at the destination hospital. For any hospital $h'$ that is short of its capacity, I first determine the number of patients it can accept, and then re-allocate patients to $h'$ based on their initial shares at $h'$. Formally, I assume $T_j$ many patients will be added to hospital $h'$'s network from insurer $j$ as a result of removing excess capacity from hospital $h$, where $T_j$ is defined as:

$$T_j = \left[cc_{h'} - \left(M \times \sum_{j \in H'} s_{jh'}\right)\right] \times \frac{s_{jh'}}{\sum\limits_{j \in L} s_{jh'}} \tag{45}$$

[62] Insurers in set $R$ are insurers who contracted with hospital $h$ and at least with one other hospital that is short in capacity.

Here, the first term in the brackets represents the number of patients hospital $h'$ can accept, and the second term represents the share of insurer $j$ among other insurers who contracted with hospital $h'$ and with at least one other hospital that exceeds its capacity in the previous step. Given this final allocation, I calculate $s_{jh}^{ins-hosp} = \frac{n_{jh}}{\sum\limits_{j \in H} n_{jh}}$ where $n_{jh}$ is the number of patients allocated to hospital $h$ from insurer $j$.

## 9.3 Results from Full Model Specifications

In what follows, I present the results from the full model specifications of hospital-level foreclosure estimation, estimation of the hospital demand model, and estimation of hospital cost function.

Table A1: Foreclosure Estimation - Hospitals

| | (1) | | (2) | | (3) | |
|---|---|---|---|---|---|---|
| | Number of plans included | Percent of plans included | Number of VI-plans included | Percent of VI-plans included | Number of non-VI-plans included | Percent of non-VI-plans included |
| VI-hospital | -1.45** | -0.04* | 0.85*** | 0.15*** | -2.30*** | -0.08*** |
| | (0.72) | (0.03) | (0.08) | (0.02) | (0.69) | (0.03) |
| Physicians | -0.003* | -0.0001* | 0.0001 | -0.00002 | -0.003** | -0.0001* |
| | (0.002) | (0.00006) | (0.0002) | (0.00005) | (0.002) | (0.00007) |
| Star rating | -0.11 | -0.0003 | 0.002 | 0.01 | -0.11 | -0.001 |
| | (0.28) | (0.009) | (0.03) | (0.009) | (0.27) | (0.01) |
| ln(Expenditure per adm.) | 0.57** | 0.04*** | 0.08*** | 0.05*** | 0.50** | 0.04*** |
| | (0.23) | (0.008) | (0.03) | (0.007) | (0.21) | (0.009) |
| Teaching hospital | -0.004 | 0.03 | -0.07 | 0.006 | 0.07 | 0.03 |
| | (0.97) | (0.04) | (0.13) | (0.03) | (0.92) | (0.04) |
| System member | 0.78* | 0.02 | 0.003 | 0.002 | 0.77* | 0.03 |
| | (0.43) | (0.02) | (0.05) | (0.01) | (0.41) | (0.02) |
| Physician group | 2.48** | 0.08 | -0.09 | -0.01 | 2.57** | 0.09 |
| | (1.11) | (0.05) | (0.09) | (0.04) | (1.07) | (0.05) |
| Birthing room | -0.66 | -0.01 | 0.10 | 0.03 | -0.76 | -0.02 |
| | (0.63) | (0.02) | (0.07) | (0.02) | (0.60) | (0.03) |
| Women health center | 0.69 | 0.02 | -0.05 | -0.005 | 0.74 | 0.03 |
| | (0.62) | (0.02) | (0.06) | (0.02) | (0.60) | (0.03) |
| Burn care | 1.93** | 0.08** | 0.14 | 0.05 | 1.80** | 0.09** |
| | (0.90) | (0.03) | (0.13) | (0.03) | (0.87) | (0.04) |
| Skilled nursing care | -0.36 | -0.02 | -0.05 | -0.01 | -0.31 | -0.02 |
| | (0.59) | (0.02) | (0.06) | (0.02) | (0.57) | (0.02) |
| Blood donor center | 0.57 | 0.02 | 0.08 | -0.02 | 0.50 | 0.02 |
| | (0.78) | (0.03) | (0.10) | (0.02) | (0.74) | (0.03) |
| Cardiology services | 0.27 | 0.01 | 0.14** | 0.04* | 0.13 | 0.01 |
| | (0.60) | (0.02) | (0.07) | (0.02) | (0.58) | (0.03) |
| Chemotherapy | 0.10 | 0.005 | 0.13** | 0.03* | -0.03 | 0.002 |
| | (0.56) | (0.02) | (0.07) | (0.02) | (0.54) | (0.02) |
| Neonatal intensive care | -1.34** | -0.05** | -0.12* | -0.04** | -1.22** | -0.05** |
| | (0.63) | (0.02) | (0.07) | (0.02) | (0.61) | (0.03) |
| Medical/surgical int. care | -0.74 | -0.02 | 0.04 | -0.007 | -0.78 | -0.02 |
| | (0.70) | (0.03) | (0.08) | (0.02) | (0.67) | (0.03) |
| Cardiac intensive care | 0.24 | 0.01 | 0.02 | 0.006 | 0.22 | 0.01 |
| | (0.59) | (0.02) | (0.07) | (0.02) | (0.57) | (0.02) |
| Kidney transplant | 1.21 | 0.05 | 0.02 | 0.02 | 1.19 | 0.05 |
| | (1.10) | (0.04) | (0.14) | (0.04) | (1.05) | (0.04) |
| Pain management center | 0.02 | -0.002 | -0.08 | -0.03* | 0.10 | 0.004 |
| | (0.51) | (0.02) | (0.06) | (0.02) | (0.50) | (0.02) |
| Physical rehabilitation | 0.04 | -0.002 | -0.0007 | -0.004 | 0.04 | -0.0006 |
| | (0.49) | (0.02) | (0.06) | (0.02) | (0.47) | (0.02) |
| Mammography | 0.67 | 0.03 | -0.05 | -0.006 | 0.72 | 0.03 |
| | (0.62) | (0.02) | (0.07) | (0.02) | (0.60) | (0.03) |
| MRI | 0.36 | 0.007 | 0.03 | 0.009 | 0.39 | 0.008 |
| | (0.62) | (0.02) | (0.08) | (0.02) | (0.59) | (0.03) |
| $R^2$ | 0.85 | 0.85 | 0.69 | 0.71 | 0.84 | 0.85 |

*Notes:* Results from ordinary least squares estimation. $N = 1285$ hospitals. Robust standard errors in parentheses. All specifications include state fixed effects, and control for hospital bed size category, primary service category, and control authority.

Table A2: Hospital Demand Estimation

| Variable | Coefficient | Variable | Coefficient |
|---|---|---|---|
| Distance (miles) | -0.03*** (0.0001) | Neonatal intensive care | 0.12*** (0.007) |
| Distance squared | 0.00002*** (0.0000001) | Neonatal intermediate care | 0.23*** (0.006) |
| VI hospital | 0.05*** (0.005) | Pediatric intensive care | 0.02*** (0.004) |
| VI x Distance | -0.003*** (0.00009) | Burn care | -0.18*** (0.005) |
| Age (0-17) x Distance | -0.03*** (0.0002) | Physical rehabilitation care | -0.19*** (0.003) |
| Age (18-34) x Distance | -0.02*** (0.0002) | Alcohol/drug abuse care | 0.22*** (0.004) |
| Age (35-44) x Distance | -0.01*** (0.0002) | Psychiatric care | -0.14*** (0.003) |
| Age (45-54) x Distance | 0.0008*** (0.0002) | Skilled nursing care | -0.20*** (0.005) |
| Female x Distance | -0.01*** (0.0001) | Intermediate nursing care | -0.04*** (0.006) |
| For-profit hospital | -0.40*** (0.005) | Acute long term care | -0.32*** (0.008) |
| Nurses per bed | 0.22*** (0.002) | Alzheimer center | 0.32*** (0.005) |
| Teaching hospital | 0.06*** (0.005) | Arthritis treatment center | 0.04*** (0.005) |
| Bed size (6-24) | -2.81*** (0.03) | Birthing room | -0.02 (0.02) |
| Bed size (25-49) | -2.40*** (0.01) | Breast cancer screening/mammograms | -0.08*** (0.006) |
| Bed size (50-99) | -1.48*** (0.008) | Adult cardiology services | -0.55*** (0.007) |
| Bed size (100-199) | -0.77*** (0.006) | Diagnostic catheterization | -0.07*** (0.006) |
| Bed size (200-299) | -0.41*** (0.005) | Interventional cardiac catheterization | 0.12*** (0.006) |
| Bed size (300-399) | -0.20*** (0.005) | Adult cardiac surgery | 0.10*** (0.004) |
| Bed size (400-499) | 0.01** (0.005) | Cardiac rehabilitation | 0.06*** (0.003) |
| Medical/surgical care | 0.26*** (0.02) | Chemotherapy | 0.17*** (0.007) |
| Obstetrics care | -0.24*** (0.02) | Computer assisted orthopedic surgery | 0.10*** (0.003) |
| Medical/surgical intensive care | -0.16*** (0.01) | Optical colonoscopy | -0.17*** (0.005) |
| Cardiac intensive care | -0.11*** (0.004) | Endoscopic ultrasound | -0.11*** (0.004) |

Table A2: Hospital Demand Estimation - continued

| Variable | Coefficient | Variable | Coefficient |
|---|---|---|---|
| Ablation of Barrett's esophagus | -0.19***<br>(0.003) | Full-field digital mammography | 0.08***<br>(0.006) |
| Endoscopic retrograde cholangiopancreatography (ERCP) | -0.14***<br>(0.006) | Magnetic resonance imaging (MRI) | -0.21***<br>(0.007) |
| Extracorporeal shock waved lithotripter (ERCP) | -0.06***<br>(0.004) | Intraoperative magnetic resonance imaging | -0.11***<br>(0.004) |
| Fertility clinic | 0.02***<br>(0.006) | Magnetoencephalography (MEG) | -0.16***<br>(0.004) |
| Geriatric services | 0.18***<br>(0.004) | Multislice spiral computed tomography <64 slice | -0.09***<br>(0.005) |
| Health screenings | -0.32***<br>(0.007) | Multislice spiral computed tomography 64+ slice | 0.07***<br>(0.005) |
| Hemodialysis | -0.20***<br>(0.004) | Positron emission tomography (PET) | -0.03***<br>(0.004) |
| HIV-AIDS services | 0.27***<br>(0.004) | PET/CT | 0.02***<br>(0.005) |
| Immunization program | 0.05***<br>(0.004) | Single photon emission computerized tomography (SPECT) | -0.05***<br>(0.004) |
| Indigent care clinic | -0.06***<br>(0.004) | Ultrasound | 0.35***<br>(0.02) |
| Linguistic/translation services | 0.04***<br>(0.004) | Image-guided radiation therapy | 0.20***<br>(0.007) |
| Neurological services | 0.03***<br>(0.007) | Intensity-modulated radiation therapy (IMRT) | -0.34***<br>(0.009) |
| Oncology services | 0.14***<br>(0.009) | Proton beam therapy | -0.22***<br>(0.009) |
| Orthopedic services | 0.02**<br>(0.01) | Shaped beam radiation system | 0.22***<br>(0.008) |
| Pain management program | 0.15***<br>(0.005) | Stereotactic radiosurgery | -0.02***<br>(0.004) |
| Palliative care program | -0.04***<br>(0.004) | Robotic surgery | 0.41***<br>(0.004) |
| Inpatient palliative care unit | 0.10***<br>(0.003) | Sleep center | 0.13***<br>(0.003) |
| Electrodiagnostic services | 0.11***<br>(0.003) | Sports medicine | -0.04***<br>(0.003) |
| Physical rehabilitation outpatient services | -0.09***<br>(0.005) | Tobacco treatment services | 0.16***<br>(0.004) |
| Psychiatric geriatric services | -0.11***<br>(0.004) | Bone marrow transplant services | 0.22***<br>(0.006) |
| Computed-tomography (CT) scanner | -0.70***<br>(0.02) | Heart transplant | -0.25***<br>(0.006) |
| Diagnostic radioisotope facility | 0.05***<br>(0.006) | Kidney transplant | 0.006<br>(0.006) |
| Electron Bean Computed Tomography | 0.02***<br>(0.004) | Liver transplant | -0.11***<br>(0.007) |

Table A2: Hospital Demand Estimation - continued

| Variable | Coefficient | Variable | Coefficient |
|---|---|---|---|
| Lung transplant | 0.11*** | Neonatal intermediate care | -0.03*** |
| | (0.007) | x Female | (0.006) |
| Tissue transplant | 0.01*** | Burn care x Burns | 3.75*** |
| | (0.004) | | (0.13) |
| Virtual colonoscopy | 0.20*** | Alcohol/drug abuse care | 2.94*** |
| | (0.003) | x Alcohol/drug induced mental disorders | (0.03) |
| Women's health center | -0.03*** | Psychiatric care | 1.87*** |
| | (0.007) | x Mental diseases and disorders | (0.02) |
| Adult cardiology services | 0.61*** | Birthing room | 0.43*** |
| x Age (35-54) | (0.007) | x Pregnancy, childbirth and puerperium | (0.02) |
| Acute long term care | 0.26*** | Birthing room | 0.33*** |
| x Age (45-54) | (0.01) | x Newborn and other neonates | (0.02) |
| Arthritis treatment center | 0.12*** | Birthing room | 0.15*** |
| x Age (45-54) | (0.008) | x Female | (0.02) |
| Medical/surgical care | -0.69*** | Blood donor center | 0.19*** |
| x Circulatory system | (0.05) | x Circulatory system | (0.02) |
| Medical/surgical care | 0.76*** | Blood donor center | 0.71*** |
| x Hepatobiliary and pancreas | (0.06) | x Blood and blood forming organ disorders | (0.05) |
| Medical/surgical care | 0.73*** | Breast cancer screening/mammograms | 0.40*** |
| x Skin, subcutaneous tissue and breast | (0.05) | x Skin, subcutaneous tissue and breast | (0.05) |
| Medical/surgical care | 0.43*** | Adult cardiology services | 0.17*** |
| x Male reproductive system | (0.08) | x Circulatory system | (0.05) |
| Medical/surgical care | 0.85*** | Diagnostic catheterization | 0.38*** |
| x Female reproductive system | (0.03) | x Kidney and urinary tract | (0.03) |
| Medical/surgical care | -2.58*** | Interventional cardiac catheterization | 0.58*** |
| x Pregnancy, childbirth and puerperium | (0.05) | x Circulatory system | (0.04) |
| Medical/surgical care | 1.16*** | Adult cardiac surgery | 1.30*** |
| x Multiple significant trauma | (0.05) | x Circulatory system | (0.02) |
| Obstetrics care | 4.58*** | Cardiac rehabilitation | 0.01 |
| x Pregnancy, childbirth and puerperium | (0.05) | x Circulatory system | (0.02) |
| Obstetrics care | 1.89*** | Chemotherapy | 1.36*** |
| x Newborn and other neonates | (0.02) | x Ear, nose, mouth and throat | (0.10) |
| Obstetrics care | -0.05** | Chemotherapy | 0.11** |
| x Female | (0.02) | x Respiratory system | (0.05) |
| Cardiac intensive care | 0.03 | Chemotherapy | 0.15*** |
| x Circulatory system | (0.02) | x Digestive system | (0.03) |
| Neonatal intensive care | 0.12*** | Chemotherapy | 0.31*** |
| x Pregnancy, childbirth and puerperium | (0.008) | x Hepatobiliary and pancreas | (0.06) |
| Neonatal intensive care | 0.15*** | Chemotherapy | 0.39*** |
| x Newborn and other neonates | (0.007) | x Skin, subcutaneous tissue and breast | (0.06) |
| Neonatal intensive care | 0.06*** | Chemotherapy | 0.59*** |
| x Female | (0.007) | x Male reproductive system | (0.09) |
| Neonatal intermediate care | -0.20*** | Chemotherapy | 0.14*** |
| x Pregnancy, childbirth and puerperium | (0.007) | x Female reproductive system | (0.03) |
| Neonatal intermediate care | -0.17*** | Chemotherapy | 0.79*** |
| x Newborn and other neonates | (0.007) | x Blood disorders | (0.11) |

Table A2: Hospital Demand Estimation - continued

| Variable | Coefficient | Variable | Coefficient |
|---|---|---|---|
| Optical colonoscopy x Digestive system | 0.39*** (0.02) | Diagnostic radioisotope facility x Circulatory system | 0.24*** (0.03) |
| Endoscopic ultrasound x Digestive system | 0.15*** (0.02) | Full-field digital mammography x Skin, subcutaneous tissue and breast | -0.06 (0.04) |
| Ablation of Barrett's esophagus x Digestive system | -0.04** (0.01) | MRI x Nervous system | 0.50*** (0.03) |
| ERCP x Digestive system | 0.55*** (0.02) | MRI x Respiratory system | 0.34*** (0.04) |
| ERCP x Hepatobiliary and pancreas | 0.73*** (0.04) | MRI x Circulatory system | 0.04 (0.03) |
| ESWL x Hepatobiliary and pancreas | 0.08*** (0.02) | MRI x Digestive system | 0.46*** (0.02) |
| ESWL x Kidney and urinary tract | 0.30*** (0.02) | MRI x Male reproductive system | 0.63*** (0.07) |
| Fertility clinic x Female reproductive system | 0.14*** (0.02) | Multislice spiral computed tomography <64 slice x Nervous system | 0.14*** (0.02) |
| Fertility clinic x Female | -0.05*** (0.006) | Multislice spiral computed tomography <64 slice x Respiratory system | 0.21*** (0.03) |
| Hemodialysis x Kidney and urinary tract | 0.34*** (0.03) | Multislice spiral computed tomography <64 slice x Circulatory system | -0.30*** (0.02) |
| HIV-AIDS services x Infectious and parasitic DD | 0.60*** (0.03) | Multislice spiral computed tomography 64+ slice x Nervous system | 0.18*** (0.03) |
| Neurological services x Nervous system | 0.92*** (0.03) | Multislice spiral computed tomography 64+ slice x Respiratory system | 0.21*** (0.03) |
| Oncology services x Ear, nose mouth and throat | 0.09 (0.10) | Multislice spiral computed tomography 64+ slice x Circulatory system | -0.11*** (0.03) |
| Oncology services x Respiratory system | -0.16*** (0.05) | PET/CT x Nervous system | 0.35*** (0.01) |
| Oncology services x Digestive system | -0.32*** (0.03) | PET/CT x Respiratory system | 0.24*** (0.02) |
| Oncology services x Hepatobiliary and pancreas | -0.41*** (0.07) | PET/CT x Circulatory system | 0.07*** (0.02) |
| Oncology services x Skin, subcutaneous tissue and breast | -0.41** (0.06) | PET/CT x Skin, subcutaneous tissue and breast | 0.12** (0.02) |
| Oncology services x Male reproductive system | -0.18* (0.10) | Ultrasound x Pregnancy, childbirth and puerperium | -0.57*** (0.02) |
| Oncology services x Female reproductive system | -0.07 (0.04) | Ultrasound x Female | -0.04*** (0.01) |
| Oncology services x Blood disorders | 0.76*** (0.11) | Heart transplant x Circulatory system | 0.45*** (0.02) |
| Psychiatric geriatric services x Mental diseases and disorders | 0.10*** (0.02) | Kidney transplant x Kidney and urinary tract | 0.88*** (0.02) |
| Diagnostic radioisotope facility x Ear, nose, mouth and throat | 0.34*** (0.06) | Liver transplant x Digestive system | 0.46*** (0.02) |
| Diagnostic radioisotope facility x Respiratory system | -0.06*** (0.03) | Lung transplant x Respiratory system | 0.51*** (0.03) |

Table A2: Hospital Demand Estimation - continued

| Variable | Coefficient | Variable | Coefficient |
|---|---|---|---|
| Tissue transplant<br>x Skin, subcutaneous tissue and breast | 0.20***<br>(0.02) | Women's health center<br>x Female reproductive system | 0.36***<br>(0.02) |
| Virtual colonoscopy<br>x Digestive system | -0.07***<br>(0.01) | Women's health center<br>x Female | 0.06***<br>(0.008) |
| FLF hospital | 0.04***<br>(0.03) | | |

Notes: Results from maximum likelihood estimation. Standard errors in parentheses. Omitted bed size category is 500+ beds, omitted age category is 45-54. Fixed effects included for hospitals missing AHA data. *** statistically significant at 1% level, ** statistically significant at 5% level, * statistically significant at 10% level.

Table A3: Cost Function Estimates

| | Coef. | × For-Profit Coef. | × Govt. Coef. | × Rural Coef. | × Teaching Coef. | × VI Coef. | × FLF Coef. |
|---|---|---|---|---|---|---|---|
| Percent ER | 0.40** | -0.13 | 0.10 | 0.48** | -0.29 | -0.52*** | -0.07 |
| | (0.17) | (0.15) | (0.23) | (0.22) | (0.27) | (0.20) | (0.16) |
| Percent Medicare | -0.35 | 0.18 | -0.60 | -0.99** | 1.60 | 1.43 | -0.09 |
| | (0.52) | (0.46) | (0.62) | (0.44) | (1.13) | (0.97) | (0.51) |
| Other OP | 0.54** | -0.75*** | -0.55* | -0.02 | 0.53 | -0.27 | -0.27 |
| | (0.24) | (0.26) | (0.33) | (0.24) | (0.63) | (0.39) | (0.21) |
| Other OP$^2$ | -0.004 | 0.003 | 0.005 | 0.01 | -0.04** | -0.007 | 0.006 |
| | (0.006) | (0.006) | (0.01) | (0.006) | (0.02) | (0.02) | (0.004) |
| Medicare OP | -0.45* | 0.30 | 0.48* | -0.06 | 0.24 | 0.62 | 0.23 |
| | (0.24) | (0.25) | (0.29) | (0.25) | (0.84) | (0.38) | (0.24) |
| Medicare OP$^2$ | 0.007 | -0.004 | -0.01 | 0.02*** | -0.003 | -0.04** | -0.005 |
| | (0.005) | (0.006) | (0.009) | (0.008) | (0.03) | (0.02) | (0.006) |
| Medicare IP × Private IP | 0.005 | 0.02 | 0.006 | -0.003 | 0.02 | -0.03 | -0.0002 |
| | (0.01) | (0.02) | (0.01) | (0.01) | (0.03) | (0.02) | (0.01) |
| All IP × Other OP | -0.04* | 0.07*** | 0.04 | -0.02 | 0.02 | 0.04 | 0.02 |
| | (0.02) | (0.02) | (0.03) | (0.02) | (0.06) | (0.04) | (0.02) |
| Other IP × Medicare IP | -0.02 | -0.06** | -0.09* | -0.07* | -0.15* | -0.05 | -0.02 |
| | (0.02) | (0.02) | (0.05) | (0.03) | (0.09) | (0.09) | (0.03) |
| Other IP × Private IP | -0.01 | 0.02** | 0.02** | 0.01 | 0.0003 | 0.04** | -0.01 |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.02) | (0.02) | (0.01) |
| All IP × Medicare OP | 0.04 | -0.02 | -0.03 | -0.03 | -0.02 | 0.005 | -0.02 |
| | (0.02) | (0.03) | (0.03) | (0.02) | (0.07) | (0.04) | (0.03) |
| All IP × Private OP | 0.005 | -0.002 | 0.002 | 0.006 | -0.007** | -0.01*** | 0.001 |
| | (0.003) | (0.003) | (0.004) | (0.004) | (0.004) | (0.004) | (0.003) |
| Private FFS OP × Other IP | -0.01 | 0.008 | -0.01 | 0.01 | 0.03 | -0.02 | 0.008 |
| | (0.01) | (0.008) | (0.02) | (0.01) | (0.02) | (0.02) | (0.009) |
| Other IP | 0.19 | 0.23 | 0.55* | -0.38* | 0.44 | 0.08 | 0.33 |
| | (0.27) | (0.26) | (0.30) | (0.22) | (0.59) | (0.44) | (0.24) |
| Other IP$^2$ | 0.01 | 0.004 | -0.002 | 0.04* | 0.05 | 0.03 | -0.01 |
| | (0.02) | (0.02) | (0.03) | (0.02) | (0.04) | (0.04) | (0.02) |
| Medicare IP × Prvt FFS OP | -0.01 | 0.0009 | -0.008 | 0.02 | -0.07 | 0.01 | -0.004 |
| | (0.01) | (0.02) | (0.01) | (0.01) | (0.04) | (0.02) | (0.01) |
| Private FFS OP | 0.18* | -0.08 | 0.11 | -0.32*** | 0.44 | 0.07 | 0.02 |
| | (0.11) | (0.14) | (0.14) | (0.11) | (0.36) | (0.14) | (0.10) |
| Private FFS OP$^2$ | 0.002 | -0.001 | 0.003 | 0.003 | -0.00003 | 0.002 | -0.006 |
| | (0.003) | (0.004) | (0.004) | (0.003) | (0.007) | (0.008) | (0.004) |
| Medicare IP | -0.45* | 0.42 | 0.77 | -0.41 | -0.11 | -0.45 | -0.55* |
| | (0.27) | (0.34) | (0.33) | (0.37) | (0.61) | (0.43) | (0.28) |
| Medicare IP$^2$ | 0.02* | -0.003 | 0.01 | 0.04*** | 0.10* | 0.07 | -0.004 |
| | (0.01) | (0.004) | (0.02) | (0.01) | (0.06) | (0.05) | (0.01) |
| Fixed Assets | 0.02 | 0.23 | -0.35 | -0.17 | 2.16** | 1.86** | -0.17 |
| | (0.22) | (0.28) | (0.27) | (0.22) | (0.84) | (0.76) | (0.23) |
| Fixed Assets$^2$ | 0.007 | -0.009 | -0.01 | 0.002 | 0.02 | -0.05* | 0.0007 |
| | (0.008) | (0.009) | (0.01) | (0.008) | (0.02) | (0.02) | (0.008) |
| Number of Beds | -1.95*** | 1.90** | 1.56* | 1.27* | 1.35 | 1.13 | -0.76 |
| | (0.61) | (0.94) | (0.82) | (0.68) | (2.60) | (1.25) | (0.60) |

Table A3: Cost Function Estimates - continued

| | Coef. | × For-Profit Coef. | × Govt. Coef. | × Rural Coef. | × Teaching Coef. | × VI Coef. | × FLF Coef. |
|---|---|---|---|---|---|---|---|
| Number of Beds$^2$ | -0.13*** | 0.11** | 0.10** | 0.08* | 0.09 | 0.01 | 0.03 |
| | (0.03) | (0.05) | (0.05) | (0.04) | (0.13) | (0.07) | (0.03) |
| Fixed Assets × Beds | -0.03 | -0.04** | -0.01 | 0.007 | 0.08 | -0.04 | 0.07*** |
| | (0.02) | (0.02) | (0.02) | (0.02) | (0.10) | (0.06) | (0.02) |
| RN Hours | 1.12** | 0.32 | -1.18** | 0.61 | -2.87** | -0.90 | -0.54 |
| | (0.44) | (0.80) | (0.46) | (0.41) | (1.11) | (0.86) | (0.39) |
| RN Hours$^2$ | -0.04** | -0.006 | 0.05** | -0.02 | 0.11** | 0.03 | 0.02 |
| | (0.02) | (0.03) | (0.02) | (0.02) | (0.04) | (0.03) | (0.02) |
| RN Hours × Staff Hours | -0.05*** | 0.01 | 0.05** | -0.009 | 0.004 | 0.03 | 0.03** |
| | (0.02) | (0.02) | (0.02) | (0.01) | (0.02) | (0.03) | (0.02) |
| Empl Hrs × Beds | 0.27*** | -0.16* | -0.17** | -0.15** | -0.28 | -0.04 | -0.07 |
| | (0.06) | (0.09) | (0.08) | (0.07) | (0.19) | (0.12) | (0.06) |
| Empl Hrs × Private IP | 0.005 | -0.005 | -0.001 | -0.002 | 0.003 | 0.005 | -0.003 |
| | (0.004) | (0.003) | (0.004) | (0.003) | (0.005) | (0.006) | (0.003) |
| Empl Hrs × Medicare IP | 0.03*** | -0.005 | -0.002 | 0.04** | -0.01 | -0.05* | -0.02 |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.02) | (0.03) | (0.01) |
| Empl Hrs × Fixed Assets | -0.007 | 0.02 | -0.02 | 0.02 | 0.08 | -0.01 | -0.01 |
| | (0.02) | (0.02) | (0.03) | (0.02) | (0.05) | (0.03) | (0.02) |
| Private IP | -0.04 | -0.18 | -0.11 | 0.05 | -0.06 | 0.03 | 0.09 |
| | (0.13) | (0.13) | (0.13) | (0.09) | (0.22) | (0.11) | (0.12) |
| Private IP$^2$ | 0.007*** | -0.01** | -0.01*** | -0.009** | -0.01 | -0.01* | 0.002 |
| | (0.003) | (0.006) | (0.003) | (0.004) | (0.008) | (0.006) | (0.004) |

Notes: Results from ordinary least squares estimation. Clustered standard errors (at the hospital level) in parentheses. All outputs and inputs are in logs. ER is emergency room visits, OP is outpatient visits, IP is inpatient days, FFS is fee-for-service, Empl Hrs is employee hours, RN is registered nurses. Specification includes hospital fixed effects and a time trend. $R^2$ is above 0.99. *** statistically significant at 1% level, ** statistically significant at 5% level, * statistically significant at 10% level.

## 10 References


Asker, J. (2016). "Diagnosing foreclosure due to exclusive dealing." *Journal of Industrial Economics*, 64(3): 375-410.

Baker, L.C., Bundorf, M. K., and Kessler, D.P. (2014). "Vertical integration: Hospital ownership of physician practices is associated with higher prices and spending." *Health Affairs*, 33(5): 756-763.

Berry, S. (1994). "Estimating discrete choice models of product differentiation." *RAND Journal of Economics*, 25: 242-262.

Berry, S., Levinsohn, J., and Pakes, A. (1995). "Automobile prices in market equilibrium." *Econometrica*, 63(4): 841–890.

Bolton, P. and Whinston, M. D. (1991). "The "foreclosure" effects of vertical mergers." *Journal of Institutional and Theoretical Economics*, (147): 207–226.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak." *Journal of the American Statistical Association*, 90(430): 443-450.

Brooks, J. M., Dor, A., and Wong, H. S. (1997). "Hospital-insurer bargaining: An empirical investigation of appendectomy pricing." *Journal of Health Economics*, 16: 417-434.

Capps, C. (2005). "The quality effects of hospital mergers." Unpublished manuscript. Bates White LLC.

Capps, C., Dranove, D., and Lindrooth, R. C. (2010). "Hospital closures and economic efficiency." *Journal of Health Economics*, 29(1): 87–109.

Capps C., Dranove, D., and Satterthwaite, M. (2003). "Competition and market power in option demand markets." *RAND Journal of Economics*, 33(4): 737-763.

Chipty, T. (2001). "Vertical integration, market foreclosure, and consumer welfare in the cable television industry." *The American Economic Review*, 91(3): 428-453.

Crawford, G., Lee, R. S., Whinston, M. D., and Yurukoglu, A. (2017). "The welfare effects of vertical integration in multichannel television markets." NBER Working Paper 21832.

Crémer, J. and Riordan, M. (1987). "On governing multilateral transactions with bilateral contracts." *RAND Journal of Economics*, 18(3): 436–451.

Dafny, L. S. (2015). "Issue brief: Evaluating the impact of health insurance industry consolidation: Learning from experience." The Commonwealth Fund, November 2015.

Dafny, L. S., Duggan, M., and Ramanarayanan, S. (2012). "Paying a premium on your premium? Consolidation in the US health insurance industry." *American Economic Review*, 102(2): 1161-1185.

Ericson, K. M., and Starc, A. (2015). "Measuring consumer valuation of limited provider networks." *American Economic Review*, 105(5): 115-119.

Fournier, G. M. and Mitchell, J. M. (1997). "New evidence on the performance advantages of multihospital services." *Review of Industrial Organization*, 12: 703–718.

Frakt, A.B., Pizer, S. D., and Feldman, R. (2013). "Plan-provider integration, premiums, and quality in the Medicare Advantage market." *Health Services Research*, 48(6): 1996-2013.

Ginsburg, P. B. (2016). "Health care market consolidations: Impacts on costs, quality and access." *California Legislature, Senate Committee on Health Informational Hearing*, March 16, 2016.

Gaynor, M., and Town, R. J. (2012). "Competition in health care markets." In McGuire, T. G., Pauly, M. V., and Pita Barros, P., editors, *Handbook of Health Economics*, volume 2, chapter 9. Elsevier North-Holland, Amsterdam and London.

Gowrisankaran, G., Nevo, A., and Town, R. (2015). "Mergers when prices are negotiated: Evidence from the hospital industry." *The American Economic Review*, 175(1): 172-203.

Grossman, S. and Hart, O. (1986). "The costs and benefits of ownership: A theory of vertical and lateral integration." *Journal of Political Economy*, 94: 691-796.

Hart, O. and Tirole, J. (1990). "Vertical integration and market foreclosure." *Brookings Papers on Economic Activity, Microeconomics*, 205-285.

Ho, K. (2006). "The welfare effects of restricted hospital choice in the US medical care market." *Journal of Applied Econometrics*, 21(7): 1039–1079.

Ho, K. (2009a). "Barriers to entry of a vertically integrated health insurer: An analysis of welfare and entry costs." *Journal of Economics and Management Strategy*, 18(2): 487-545.

Ho, K. (2009b). "Insurer-provider networks in the medical care market." *The American Economic Review*, 99(1): 393–430.

Ho, K., and Lee, R. S. (2017a). "Insurer competition in health care markets." *Econometrica*, 85(2): 379-417.

Ho, K., and Lee, R. S. (2017b). "Equilibrium insurer-provider networks: Bargaining and exclusion in health care markets." NBER working paper 23742.

Ho, V., and Hamilton, B. (2000). "Hospital mergers and acquisitions: Does market consolidation harm patients?" *Journal of Health Economics*, 19(5): 767-791.

Joskow, P. (2008). "Vertical integration." In *Issues in competition law and policy.* Chicago: ABA Section of Antitrust Law.

Khanna, G., Narula, D., and Rao, N. (2016). "The market evolution of provider-led health plans." *McKinsey & Company, Healthcare Systems and Services Practice.*

Lee, R. (2013). "Vertical integration and exclusivity in platform and two-sided markets." *The American Economic Review*, 103(7): 2960-3000.

Lewis, M. and Pflum, K. (2015). "Diagnosing hospital system bargaining power in managed care networks." *American Economic Journal: Policy*, 7(1): 243-274.

McFadden, D. (1974). "Conditional logit analysis of qualitative choice behavior." In *Frontiers in Econometrics*, Ed. P. Zarembka. *Academic Press: New York*, 105-142.

Ordover, J., Saloner, G., and Salop, S. (1990). "Equilibrium Vertical Foreclosure." *American Economic Review*, 80: 127-142.

Perry, M. K. (1989). "Vertical integration: Determinants and effects." In Schmalensee, R. and Willig, R., editors, *Handbook of Industrial Organization*, volume 1. Elsevier North-Holland, Amsterdam.

Salop, S. C. and Scheffman, D. T. (1983). "Raising rivals' costs." *American Economic Review*, 73(2): 267-271.

Small, K., and H. Rosen (1985). "Applied welfare economics with discrete choice models." *Econometrica*, 49: 105-130.

Song, Z., Landrum, M. B., and Chernew, M. E. (2012). "Competitive bidding in Medicare: Who benefits from competition?" *The American Journal of Managed Care*, 18(9): 546–552.

Stigler, G. J. (1968). "The organization of industry." *Chicago: University of Chicago Press.*

Tirole, J. (1988). "The theory of industrial organization." *MIT Press.*

Town, R., Wholey, D., Feldman, R., and Burns, L. (2006). "The welfare consequences of hospital mergers." NBER Working Paper 12244.

Trish, E. E. and Herring, B. J. (2015). "How do health insurer market concentration and bargaining power with hospitals affect health insurance premiums?" *Journal of Health Economics*, 42: 104-114.

Vita, M. G. (2000). "Regulatory restrictions on vertical integration and control: The competitive impact of gasoline divorcement policies." *Journal of Regulatory Economics*, 18(3): 217–233.

Williamson, O. (1971). "The vertical integration of production: Market failure considerations." *American Economic Review*, 61(2): 112-123.