2021

Public policy and vertical relationships in the healthcare market

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

PUBLIC POLICY AND VERTICAL RELATIONSHIPS IN THE HEALTHCARE MARKET

by

XIAOXI ZHAO

B.S., Beihang University, 2011
M.Eng., Beihang University, 2014
M.S., Ecole Centrale de Lyon, 2014
M.A., Boston University, 2015

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2021

Approved by

First Reader

Randall P. Ellis, PhD
Professor of Economics

Second Reader

Marc Rysman, PhD
Professor of Economics

Third Reader

Keith M. Ericson, PhD
Associate Professor of Markets, Public Policy and Law

Fourth Reader

Jordi Jaumandreu, PhD
Senior Academic Researcher of Economics

To my great parents and my beloved spouse.

Acknowledgments

First and foremost, I would like to express my most profound appreciation to my main advisors, Randall P. Ellis and Marc Rysman, for their invaluable advice and unwavering guidance throughout my while Ph.D. training program. Their immense knowledge and ample experience have encouraged me in all the time of my research.

I would also extend my sincere thanks to Jordi Jaumandreu, for all the in-depth discussions we had since the beginning of my dissertation work and his constructive advice for Chapter 1.

I am extremely grateful to Leila Agha, Keith M. Ericson, and Christopher Whaley for the mentorship and coauthorship. In particular, Chapter 2 is a joint work with Leila and Keith, and Chapter 3 is a joint work with Chris. In addition to their contribution to those two chapters, I also very much appreciate their insightful suggestions for Chapter 1.

I also greatly acknowledge financial support from the National Institutes of Health (P30AG012810) and the Agency for Healthcare Research and Quality (1U19HS024067-01). The contents of this dissertation do not represent the official views of the NIH or AHRQ. All errors and omissions belong to the authors.

Finally, I would like to thank my family, friends, and peers, without whose support and encouragement, this dissertation would not have been possible.

PUBLIC POLICY AND VERTICAL RELATIONSHIPS IN THE HEALTHCARE MARKET

XIAOXI ZHAO

Boston University, Graduate School of Arts And Sciences, 2021

Major Professor: Randall P. Ellis, PhD
Professor of Economics

ABSTRACT

In complex, rapidly evolving healthcare markets, vertical relationships play an increasingly important role in pricing and health care delivery, creating or eliminating barriers to market efficiency. Despite their importance, few studies have addressed how vertical relationships affect care coordination and utilization and how public policies affect the market through those relationships. This dissertation focuses on two types of vertical relationships—the price negotiation between insurers and health care providers and the relationship between health care organizations and physicians.

Chapter 1 examines the effect of medical loss ratio (MLR) regulation on insurer pricing behavior in light of insurer-provider price negotiation. Such regulation could disincentivize insurers from efficient cost-cutting, resulting in higher health service prices and consumer welfare loss. Using a Nash-in-Nash bargaining model with a regulation constraint, I find that MLR regulation rules out price bargaining equilibria with low health service price. The empirical examination of the Affordable Care Act MLR regulation shows that the regulation significantly increases medical cost for non-compliant insurers. Either restricting health service prices or replacing the regulation with a low-margin public option could keep prices low and improve consumer welfare.

Chapter 2 utilizes patient changes of their primary care physicians (PCP) to examine the effect of health care organization boundaries on patient healthcare utilization. We define a visit-based organizational care concentration metric to measure how patient health care utilization spreads across organizations. Applying that metric to Medicare data, we capture the substantial variation of organizational care concentration across physicians and regions. Leveraging on PCP changes due to the exit of original PCPs, we find that patients who switched to PCPs with higher organizational care concentration experience a significant decrease in their total healthcare utilization without significant changes in quality of care.

Chapter 3 examines the consequences of vertical integration between hospitals and physician practice groups, which could incentivize physicians to refer more patients to hospitals and induce higher medical spending. Employing an event-study framework and a discrete choice model, we find consistent and robust evidence that vertical integration increases the probability of choosing hospital-based facilities for outpatient surgical procedures, increasing both medical spending and travel distances.

Contents

List of Tables

List of Figures

List of Abbreviations

ACA	.....	Affordable Care Act
AMA	.....	American Medical Association
ASC	.....	Ambulatory Surgical Center
BETOS	.....	Berenson-Eggers Type of Service
CCIIO	.....	Center for Consumer Information and Insurance Oversight
CMS	.....	Centers for Medicare & Medicaid Services
FFM	.....	Federally-Facilitated Marketplaces
FFS	.....	Fee-For-Service
HCPCS	.....	Healthcare Common Procedure Coding System
HHI	.....	Herfindahl-Hirschman Index
HHS	.....	Health and Human Services
HOPD	.....	Hospital Outpatient Departments
HRR	.....	Hospital Referral Region
IV	.....	Instrument Variable
MLR	.....	Medical Loss Ratio
NPI	.....	National Provider Identifier
NPPES	.....	National Plan and Provider Enumeration System
OOP	.....	Out-Of-Pocket
PCP	.....	Primary Care Provider
SD	.....	Standard Deviation
TIN	.....	Tax ID Number
VI	.....	Vertical Integration

Chapter 1

Medical Loss Ratio Regulation and Insurer Pricing

1.1 Introduction

Regulating a firm's profitability is a common way for governments to try to obtain efficient outcomes in critical markets. It is particularly important for the healthcare market where private firms exploit their pricing leverage when complementing public provisions (Einav, Finkelstein and Polyakova, 2018; Curto et al., 2019; Duggan, Starc and Vabson, 2016). Unlike traditional rate-of-return regulation, which regulates the ratio of profit to capital costs, the Affordable Care Act (ACA) limits the ratio of a variable cost measurement (approximately) to revenue, which the ACA terms the Medical Loss Ratio (MLR). The ACA MLR regulation requires each insurer in the private insurance market to spend at least 80% or 85% of the premium revenue on medical claims and quality improvement. The regulator's goal is presumably to limit insurer markups, hoping that this will improve the quality of insurance plans while keeping plan premiums low. Between 2014 to 2018, 16 million individuals purchased insurance plans in the individual private insurance market each year, accounting for $74 billion of premium payment and $65 billion of medical claims spending (Cox, Fehr and Levitt, 2019; Fehr, Cox and Levitt, 2019). Understanding the mechanism by which regulation affects pricing and medical spending is critical for evaluating this regulation and for designing future policies.

It is far from obvious that higher MLRs translate into lower absolute markups for insurance firms, better insurance plans, or more affordable health care services. Along with the fact that the average MLR in the individual market jumped from 84.1% in 2011 to 92.9% in 2016 (CCIIO, 2017), existing studies find that the ACA MLR regulation is associated with improvements of insurers' financial performance (Cox, Fehr and Levitt, 2019; Fehr, McDermott and Cox, 2020) and increases in medical costs (McCue, Hall and Liu, 2013; McCue and Hall, 2015; Cicala, Lieber and Marone, 2019; Callaghan, Plummer and Wempe, 2020). Those papers link the regulation to insurers' financial measures but fail to identify viable mechanisms for these effects.

One possible unintended effect of MLR regulation is that it disincentivizes insurance firms from containing costs when negotiating health service prices. Recent papers model insurers as bargaining with health care providers on health service prices to reduce the medical bills they pay to providers (Gowrisankaran, Nevo and Town, 2015; Craig, Ericson and Starc, 2018; Gaynor, Ho and Town, 2015; Barrette, Gowrisankaran and Town, 2020). This bargaining process opens a channel for insurers to change their cost containment behavior due to MLR regulation. When insurers have relatively large bargaining power, they may realize that pursuing optimal solutions with low service prices will make them non-compliant with the regulation. Consequently, insurers may concede part of their bargaining power and profit, allowing higher health service prices. Moreover, higher health service prices will lead to higher out-of-pocket payment for consumers through a non-zero consumer cost-sharing system and higher premiums if insurers price it in.

In this paper, I develop a model of how insurers negotiate health service prices and determine premiums under MLR regulation. This model reveals the mechanism through which MLR regulation affects the price bargaining equilibrium and insurers' premium choices. By combining this model with demand estimation, I empirically show how the ACA MLR regulation affects negotiated prices, insurers' cost, and consumer welfare in the ACA Health Insurance Exchange Marketplace, without observing pre-regulation data.

To estimate the effects of MLR regulation on price and consumer welfare, I start by modeling the demand for insurance plans. I adopt a discrete choice model with random coefficients from previous studies (Berry, Levinsohn and Pakes, 1995; Nevo, 2001) to describe how consumers choose insurance plans based on premiums and other plan features. Having estimated the demand for insurance plans, I build a two-stage model with constraint on insurers' pricing under MLR regulation. At the first stage, the price negotiation stage, insurers and providers bargain on health service prices, both aware of MLR regulation and know the demand function. The bargaining equilibrium depends on the MLR regulation, the relative bargaining power of the two sides, and the demand function. At the second stage, where insurers maximize their profits by choosing premiums, I introduce MLR regulation as a constraint imposed on insurers. This constraint will be binding if insurers' firm-level MLRs are below the threshold required by the regulation.

Following that, I show graphically how binding MLR regulation rules out bargaining equilibria with low negotiated service prices. By imposing MLR regulation onto the optimal prices and premium choices, I show how that premium-price combinations with low negotiated price are no longer optimal because of the regulation.

This two-stage model provides a straightforward empirical approach for estimating the effect of MLR regulation. I estimate this model using data from the individual Health Insurance Exchange Marketplace, which has covered 10 million people every year since it launched in 2014. After estimating the demand for health insurance plans, I structurally estimate the effect of the ACA MLR regulation on health service prices, insurers' bargaining power, and fixed cost. Consistent with my theoretical model, my estimates suggest that the MLR regulation leads to higher negotiated prices and higher costs to insurers.

The estimates inform welfare analysis of counterfactual market settings that change insurers' pricing and the market outcomes. Based on those structural estimates, I introduce price negotiation and the ACA MLR regulation sequentially to decompose their effects on the negotiated prices, premiums, and welfare. Results from the counterfactual analysis imply that if service prices are fixed at the no-regulation negotiated level, insurance plans would be more affordable and cover more people. However, when service prices are negotiated by insurers and health care providers, the MLR regulation results in higher service prices and higher premiums. Therefore, consumers will need to pay more out of pocket for both health care services and insurance plans.

One of this paper's primary contributions is using a structural model to reveal the mechanism by which MLR regulation can potentially affect insurers' costs and pricing. Building on the reduced-form evidence from previous studies (Cicala, Lieber and Marone, 2019; Callaghan, Plummer and Wempe, 2020), this paper explicitly shows how insurers respond to the ACA MLR regulation via price negotiation and structurally estimates the effect of MLR regulation. This model has three attractive features. First, by structurally estimating this model, I am able to draw conclusions about the regulation's effects on welfare. Second, this model does not require data on negotiated prices, which are rarely available in the private insurance market. Third, this structural estimation is more attractive when pre-regulation data is not available—when the market started after the regulation was implemented.

The model structure and estimates inform a novel counterfactual analysis, which investigates the interaction between MLR regulation and price negotiation. First, I consider two counterfactual scenarios: one with price negotiation and no regulation; and a second with MLR regulation but fixed service prices. By comparing these two scenarios with the real ACA exchange marketplace, I find that it is the combination of the regulation and negotiation that leads to higher service prices, higher premiums, and consequently a similar profit level but higher OOP payment for consumers. I further test an alternative option—introducing a public insurance plan with 80% MLR and removing MLR regulation for the other private plans. The results suggest that such a public option could lower prices and improve consumer welfare.

This paper also contributes to the empirical studies of the ACA Exchange Marketplace. First, it estimates the effect of MLR regulation on the ACA Exchange Marketplace, which is absent in the existing literature. Second, the demand estimation in this paper provides baseline estimates of demand in all the Federally-Facilitated Marketplaces (FFM), which complement previous studies that focus on the Massachusetts pre-ACA marketplace (Ericson and Starc, 2015; Jaffe and Shepard, 2017) and that focus on one or two state-based ACA exchange marketplaces (Abraham et al., 2017; Tebaldi, 2017; Saltzman, 2019; Drake, 2019).

The paper proceeds as follows. Section 2 provides an overview of existing related literature. Section 3 describes the ACA MLR regulation and the Health Insurance Exchange Marketplace. Section 4 presents the theoretical framework along with graphs illustrating the effect of MLR regulation on insurers' pricing. Section 5 discusses sample construction and provides descriptive statistics of the final sample. In Section 6 describes my empirical strategy and identification assumptions. Section 7 presents the main results. Section 8 simulates the markets in alternative settings. Section 9 concludes.

1.2 Related Literature

This paper mainly relates to three strands of literature: rate-of-return regulation, insurer-provider negotiation, and the ACA marketplace.

My work relates to a broad literature on rate-of-return regulation. Averch and Johnson (1962) explain how rate-of-return regulation distorts resource allocation. In the health care market, Acemoglu and Finkelstein (2008) shows the effect of the Medicare payment reform on the capital-labor ratios. Unlike those policies that focus on the capital ratio, MLR regulation imposes a requirement on the ratio of variable cost to revenue, approximately. Cicala, Lieber and Marone (2019) discuss the similarity of the MLR regulation and the rate-of-return regulation with a single input.

With regard to MLR regulation, the literature is relatively small. Karaca-Mandic, Abraham and Simon (2015) use pre-ACA data to show that the medical loss ratio is a good measure of insurers' price-cost margins. They also find that MLRs are significantly lower in monopoly markets than in markets with two or more insurers. Abraham, Karaca-Mandic and Simon (2014) find that MLRs increased in the individual market in the first year after the ACA MLR regulation. McCue, Hall and Liu (2013) assess the changes between 2010 and 2011 in insurers' financial performance and find that in 2011, the administrative cost ratios and the operating margins both decreased, especially for individual health insurers. Later, McCue and Hall (2015) study the data from the second year and find a continuous increase in MLRs and a continuous decrease in the administrative cost ratios. Recently, Cicala, Lieber and Marone (2019) use the difference-in-differences method to study the effect of the ACA MLR regulation on medical cost and premiums. They find that the regulation is associated with higher claims cost but has statistically no effect on the premiums. Callaghan, Plummer and Wempe (2020) find similar consequences following the implementation of the ACA MLR regulation. They show that the MLRs increased remarkably since 2011, and that increase in MLRs comes primarily from the increase in claim cost, not the reduction in premium. Building on these reduced-form results, I introduce the price negotiation between insurers and health care providers to the model. Price negotiation is an essential channel for insurers to contain medical costs and strategically maintain their compliance status under MLR regulation. Another difference between my paper and existing studies is that those studies exploit the pre-post variation, which does not exist for the Marketplace. To fill this vacuum, I employ a structural model and use the cross-sectional variation to estimate the effect of MLR regulation in the relatively new market.

My work also relates to papers that study the insurer-provider negotiation in different institutional contexts. Reduced-form evidence shows that an increase in bargaining power (due to mergers or measured by market HHI) relates to a reduction in physician earnings and an increase in the substitution of nurses for physicians (Dafny, Duggan and Ramanarayanan, 2012). Roberts, Chernew and McWilliams (2017) show that low service prices are associated with large market shares. Cooper et al. (2019) find that hospital prices are lower in higher concentrated insurer markets and lower concentrated hospital markets. Trish and Herring (2015) find that high concentration of provider markets is associated with the higher premiums after control the concentration of insurance market. Using the bilateral Nash bargaining model developed by Horn and Wolinsky (1988), researchers structurally examine the role of price negotiation in the health care market. Grennan (2013, 2014) studies the bargaining between hospitals and medical devices suppliers. Gowrisankaran, Nevo and Town (2015) examine the bargaining between hospitals and insurers in Virginia. Ho (2009), Ho and Lee (2017) and Ho and Lee (2019) examine the bargaining between hospitals and insurers in California, primarily focusing on the network formation. Those papers highlight the importance of the insurer-provider negotiation. Because of the time period studies, MLR regulation was absent in those papers. I incorporate the price negotiation and MLR regulation in this paper, using a similar but richer Nash bargaining model.

From the theoretical perspective, Aghadadashli, Dertwinkel-Kalt and Wey (2016) shows how bargaining power and demand elasticity affect the profit share in vertical relations if upstream and downstream firms bargain over linear input prices. This setting is very similar to the scenario where no regulation is imposed on the MLR. Despite the deep and broad discussion of the rate-of-return regulation and the bargaining model, very few studies discuss the consequences when such regulation is implemented in a bargaining setting. Hendricks (1975) introduced the wage negotiation when the rate-of-return of capital is regulated. My model suggests that the division of profit between insurers and health care providers is distorted by MLR regulation.

The third strand of literature that this paper relates to is the ACA health insurance exchange marketplace literature. Being a new market with new policies, the Marketplace attracts many research works. A large part of them focus on the demand estimation in this market. Abraham et al. (2017) found that demand was highly elastic in 2014-2015. Tebaldi (2017) studies how the subsidy policy affects demand in California. Saltzman (2019) estimates the demand for insurance plans in California and Washington. Drake (2019) estimates the demand in 2017 California's Marketplace. In this paper, I employ a discrete choice model with random coefficients to estimate the demand in all the Federally-Facilitated Marketplaces, from 2014 to 2017. My estimates fall within the range of previous estimates, including those from the Massachusetts pre-ACA marketplace (Ericson and Starc, 2015; Jaffe and Shepard, 2017). The competition among insurers is another popular topic, spanning from the relation between premium increase and monopoly (Parys, 2018) to insurers' massive exits (Griffith, Jones and Sommers, 2018). This study raises another concern related to insurers' large bargaining power—increasing medical spending.

1.3 Background

1.3.1 The ACA Medical Loss Ratio Regulation

The Medical Loss Ratio regulation under the ACA requires each insurer to spend 80%—for individual and small group market—or 85%—for large group market—of the total premium revenue on medical claims and health care quality improvement. Insurers that do not reach the threshold must publicly disclose their non-compliance and refund to their enrollees the portion of premium that exceeds the limit. If insurers are not compliant for three years, they will not be allowed to enter the market.

The measurement is implemented at the insurer-market segment level. That is, in each state, one insurer will have three MLRs if the firm participates in all three segments—individual, small group, and large group. Along with MLRs, the compliance status and rebates are also at the insurer-segment level.

More specifically, the MLR calculation includes a preliminary calculation of the ratio and a credibility-based adjustment. The preliminary MLR of an insurer in a market segment is

$$Preliminary\ MLR\ = \frac{Total\,Medical\,Care\,Claims + Quality\,Improvement\,Expenses}{Total\,Premium\,Revenue - Taxes,\ Licensing\,and\,Regulatory\,Fees}.$$

The core part and the most flexible part of MLR is the medical claims over premium revenue. On average, the quality improvement expenses are only one percent of the medical care claims, and the taxes and fees together count for 17% of premium revenue.[1] Although insurers have some degree of flexibility in what items to be

[1] Calculated by author, based on the 2011-2018 MLR reports, individual segment.

counted as quality improvement expenses, it takes time and effort for them to get approvals from the state commissioner.

Because of the uncertainty in medical claims, the regulation allows a credibility-based adjustment of MLRs for small insurers,

$$Adjusted\,MLR = Preliminary\,MLR + Base\,Credibility\,Factor * Deductible\,Factor.$$

The $base\,credibility\,factor$, as shown in Panel A of table 1.1, decreases in the number of enrollees. Insurers having less than 1000 enrollees in a market are exempted from the regulation. For insurers with 1,000 to 75,000 enrollees, adjusted MLRs are higher than their preliminary ones.

The other adjustment factor is the $deductible\,factor$. The deductible line is a threshold for enrollees below which enrollees need to pay 100% of the medical bills by themselves. The deductible factor is calculated based on the average deductible of all plans provided by the insurer, as shown in Panel B of table 1.1.

By the regulation, insurers need to provide rebates to their enrollees if their adjusted MLRs are above the 80% threshold in the individual market. Nationwide, average annual rebates in the individual market were $260 million from 2011 to 2018.[2]

1.3.2 The ACA Health Insurance Exchange Marketplaces

The Health Insurance Exchange Marketplace is an online platform for individuals and small groups to shop for health insurance plans. It opened in October 2013 when people could enroll in health insurance plans for 2014. The exchange marketplace is an important market for the individual insurance market. Every year, around 10 million individuals purchased their insurance at the Marketplace. That is 55% to 72% of the total private individual market, depending on the year (Cox, Fehr and Levitt, 2019). Unlike in the large group markets, consumers and insurers in the Marketplace

[2] Calculated based on the 2011 to 2018 data from the CMS.

do not bargain on the premiums or plan design. When people go to the platform, they choose insurance plans by comparing predetermined plan characteristics listed on the website, such as premiums, coverage rates, etc.

To ease the cross-plan comparison, the ACA requires insurers to classify all the plans into five categories (four metal categories plus one "catastrophic" category) based on plans' actuarial values. The table 1.2 shows the actuarial values for each category. The more precious the metal, the higher the fraction of medical bills the insurance plan will cover. Among all five categories, silver is the most commonly offered plan, no matter measured by the number of plans or the number of enrollees. Two factors contribute here. On the insurer side, by regulation, insurers must provide at least one silver plan to gain the entrance ticket. The second factor is that low-income silver plan enrollees are eligible for a premium subsidy (Tebaldi, 2017).

Plans in the same metal category have the same actuarial values but could be differentiated by other features such as premiums, deductibles, out-of-pocket maximums, and cost-sharing, as well as provider networks. Unfortunately, limited by the data, I can not observe the provider network. However, premiums, metal categories, and cost-sharing rates are much more salient than the network from consumers' view. Therefore, I leave the network feature for future research.

The marketplaces are state-level markets. Every state could choose between the Federally-Facilitated Marketplace (FFM) and the state-based Marketplace. The FFM is operated by the U.S. Department of Health and Human Services (HHS). States choosing FFM share similar marketplace design and regulatory environments. Therefore, this paper focuses on the FFMs. From 2014 to 2017, 40 states participated in the FFM for at least one year.

1.4 Theoretical Framework

This section presents the model of insurance plan demand and insurers' pricing with MLR regulation. The model consists of three parts. In the first part, insurers and health care providers bargain on health service prices. Following that, insurers determine premiums of the insurance plans and sell them in the market. Finally, consumers shop for insurance plans based on premiums and plan characteristics. If a consumer visits a health care provider, the insurer will pay part of the medical bill according to the plan, and the consumer will pay the other part. The health care provider treats the patient and gets payments from the insurer and the patient. In the following subsections, I solve the model by backward induction.

1.4.1 Demand for insurance plans

To model the demand for insurance plans, I adopt a methodology set out by Berry, Levinsohn and Pakes (1995) that has been used in both health (Ho, 2009) and non-health sectors (Nevo, 2001). I assume that the utility of individual i from choosing plan j in state s and year t is

$$u_{ijst} = X_{jst}\beta_i + \alpha_i \phi_{jst} + \xi_s^s + \xi_t^t + \varepsilon_{ijst}, \tag{1.1}$$

where X_{jst} is a vector of plan characteristics such as deductible and maximum Out-Of-Pocket (OOP) allowed by the plan. ϕ_{jst} is the plan premium. ξ_s^s and ξ_t^t are state fixed effects and year fixed effects, respectively. ε_{ijst} is the idiosyncratic error term, which follows EV Type I. α_i and β_i are random coefficients both of which consist of a constant part and a random part:

$$\begin{pmatrix}\alpha_i \\ \beta_i\end{pmatrix} = \begin{pmatrix}\alpha \\ \beta\end{pmatrix} + \Pi D_{ist} + \Sigma \nu_i,$$

where D_{ist} is a random draw of demographic characteristics in state s year t, and ν_i follows a i.i.d. standard normal distribution. Π and Σ are the matrices of random coefficients to be estimated.

Then the utility individual i receives from choosing plan j can be written as

$$u_{ijst} = X_{jst}\beta+\alpha\phi_{jst}+\xi_s^s+\xi_t^t+\sum_k x_{jkst}\pi_k^\beta d_{ikst}^\beta+\pi^\alpha d_{ist}^\alpha\phi_{jst}+\sum_k x_{jkst}\sigma_k^\beta\nu_{ik}^\beta+\sigma^\alpha\nu_i^\alpha\phi_{jst}+\varepsilon_{ijst}$$

where the sum of first four terms, $X_{jst}\beta + \alpha\phi_{jst} + \xi_s^s + \xi_t^t$, represents the mean utility in market st. The following four terms, including the two summation terms, have all the random coefficients. The random part could be considered as an individual deviation from the market-level average.

The market share of plan j in market t is the probability that plan j provides the highest utility to individuals. More specifically, based on the utility function, the market share of plan j in state s, year t is

$$s_{jst} = \int \frac{e^{u_{ijst}(X_{jst},\phi_{jst},\xi_{jst},D_{ist},\nu_i;\alpha,\beta,\Pi,\Sigma)}}{1+\sum_k e^{u_{ikst}(X_{kst},\phi_{kst},\xi_{kst},D_{ist},\nu_i;\alpha,\beta,\Pi,\Sigma)}}\varphi(\nu_i)d\nu_i. \tag{1.2}$$

By multiplying the market share with the market size M_{st}, I obtain the demand function of plan j in market st

$$D_{jst}(\phi_{jst}, X_{jst}) = M_{st}s_{jst}.$$

1.4.2 Insurers' premium choices

Insurers choose premiums to maximize their total profits given service prices and MLR regulation. Consider an insurer f that provides a set of insurance plans J_f.[3] The objective in this part is to choose a set of premiums $\{\phi_j\}_{j\in J}$ that maximizes the total profit Π_f^I with a MLR no lower than the threshold $\bar{R}$. Following the regulation, I introduce MLR regulation as a constraint imposed at insurer-level. Therefore, the

[3] To simplify the notation, I omit the subscript s and t for the rest of theoretical model.

constrained profit-maximization problem that insurer f needs to solve is

$$\max_{\{\phi_j\}_{j\in J_f}} \Pi_f^I = \sum_{j\in J_f}(\phi_j - \tilde{p}_j\kappa_j\theta)D_j(\boldsymbol{\phi}) - C^F$$

$$s.t. \quad \sum_{j\in J_f}\tilde{p}_j\kappa_j\theta D_j \geq \bar{R}\sum_{j\in J_f}\phi_j D_j,$$

where ϕ_j is the premium of plan j, $\tilde{p}_j$ is the health service price, κ is plan j's coverage rate, θ is the fraction of enrollees who seek care, and C^F is the fixed cost to the insurer.

The insurers' optimal solution to the problem can be rewritten as

$$\max_{\{\phi_j\}_{j\in J_f}} \mathfrak{L}^I = \sum_{j\in J_f}(\phi_j - mc_j)D_j(\boldsymbol{\phi}) - C^F + \lambda(\sum_{j\in J_f} mc_j D_j - \bar{R}\sum_{j\in J_f}\phi_j D_j), \tag{1.3}$$

where λ is the Lagrangian multiplier, a measure of the shadow cost brought by MLR regulation. By the Kuhn-Tucker theorem, $\lambda = 0$ when issuers are compliant with MLR regulation, $\lambda > 0$ otherwise. In the appendix section A, I prove that $\lambda \in (0, 1)$ if the regulation is binding.

The first order condition yields the relation between the optimal premiums and the health service prices in the presence of MLR regulation. That is,

$$\boldsymbol{\phi} = \frac{1-\lambda}{1-\lambda\bar{R}}\tilde{p} \odot \boldsymbol{\kappa}\theta - \mathbf{J}^{-1}\mathbf{D} \tag{1.4}$$

where $\odot$ indicates element-wise multiplication and J is the Jacobian matrix of $\boldsymbol{D}$ with respect to $\boldsymbol{\phi}$. This best response function shows that for insurers with MLRs above the threshold required by the regulation, their optimal premiums equal the medical cost plus a markup. Insurers that are non-compliant with the regulation will put less weight on the medical cost for the premium choices.

In the appendix section A, I prove that $\lambda \in (0, 1)$ is the sufficient and necessary condition of $0 < \frac{1-\lambda}{1-\lambda\bar{R}} < 1$. Putting the above discussion in an Averch-Johnson model, λ ranging between 0 and 1 is the crucial condition for the existence of the Averch-Johnson effect (Averch and Johnson, 1962; Takayama, 1969; Stein and Borts, 1972).

1.4.3 Price negotiation

In the first stage, insurers and health care providers negotiate on health service price $\tilde{\mathbf{p}}$. In practice, insurers usually bargain for a set of insurance plans at once. Thus, I assume that plans provided by the same insurer in the same market have the same service price, that is $\tilde{p}_j = \tilde{p}, \forall j \in J_f$. Following previous studies (Crawford and Yurukoglu, 2012; Grennan, 2013; Gowrisankaran, Nevo and Town, 2015; Ho and Lee, 2017), I adopt the bilateral Nash bargaining model proposed in Horn and Wolinsky (1988). I assume that each insurer-provider pair maximizes a bilateral Nash product, taking the outcomes of the others as given. One important difference between my model and previous ones is the constraint introduced by MLR regulation. This regulation is well known to all the insurers and providers. Therefore, I assume that the insurer's profit in the Nash product is constrained.

Due to data limitations, I do not observe provider networks or the demand for services for each provider. Thus, I use one provider to represent one market. This assumption is strong if the paper focuses on provider network negotiation. It is less concerning for my question, which focuses on a regulation imposed on insurer profits. Including more providers will change alternative profits of both sides and give insurers more bargaining power. It will not affect the conclusion that MLR regulation rules out bargaining equilibrium with low prices.

Consider one insurer-provider pair. They seek to find the service price $\tilde{p}$ that maximizes the Nash product of their profits below

$$\Pi^{Nash} = \left(\Pi^I - \Pi^I_0\right)^\tau \left(\Pi^H - \Pi^H_0\right)^{(1-\tau)},$$

where τ is the bargaining parameter which measure insurer's bargaining power. Π^I and Π^H are insurer's constrained profit and health care provider's profit, respectively, when they agree on price $\tilde{p}$. In the case of no agreement is reached, the insurer will not enter this market and receive Π_0^I; health care providers will see patients covered by other insurance plans and receive Π_0^H. Therefore, the four relevant profit functions in the bargaining problem are

$$\begin{aligned}
&\Pi^I = \textstyle\sum_{j\in J_f} \left(\phi_j - \tilde{p}_j\kappa_j\theta\right) D_j(\boldsymbol{\phi}) - C^F + \lambda \left(\sum_{j\in J_f} \tilde{p}_j\kappa_j\theta D_j - \bar{R}\sum_{j\in J_f} \phi_j D_j\right), \\
&\Pi^H = \textstyle\sum_{j\in J_f} (\tilde{p} - \tilde{c_h})\theta D_j, \\
&\Pi_0^I = 0, \\
&\Pi_0^H = \textstyle\sum_{j\in J_f} (\tilde{p}_0 - \tilde{c_h})\theta D_j.
\end{aligned} \tag{1.5}$$

$\tilde{c_h}$ is the service cost per patient to the health care provider, which is assumed to be the same for patients covered by different insurance plans. $\tilde{p}_0$ is the alternative service price.

In the following subsections, I solve the Nash product maximization problem defined above to obtain the negotiated prices at equilibrium. To better illustrate the intuition, I start with a single-product scenario and provide the multi-product solution after that. It is important to note that one insurer could and usually does offer multiple plans in a market. When the regulation is imposed at the insurer-level, insurers could make a large profit from one plan at the expense of profit of another plan.

Single-product scenario

When the insurer provides only one insurance plan in the market, the objective function of the bargaining model could be simplified as

$$\Pi^{Nash} = \left[(\phi - \tilde{p}\kappa\theta)D - C^F + \lambda(\tilde{p}\kappa\theta D - \bar{R}\phi D)\right]^{\tau} \left[(\tilde{p} - \tilde{p}_0)\theta D\right]^{(1-\tau)}.$$

By the envelope theorem and the relation in equation (1.4), the F.O.C. of the objective function above with respect to p yields,

$$\tilde{p} = \tilde{p}_0 - \frac{(1-\lambda\bar{R})\frac{D^2}{D'} + C^F}{\frac{\tau}{1-\tau}(1-\lambda)\kappa\theta D + \frac{D'}{D}\phi'[(1-\lambda\bar{R})\frac{D^2}{D'} + C^F]}. \tag{1.6}$$

As shown by the equation above, the equilibrium price consists of two parts—the opportunity cost to health care provider $\tilde{p}_0$ and a unit markup for the provider. The latter part depends on the bargaining power τ, the effect of MLR regulation measured by λ, and the demand function $D(\cdot)$. The negotiated price $\tilde{p}$ is decreasing in insurers' bargaining power τ. Intuitively, insurers with larger bargaining power could push the negotiated price lower and leave smaller profit to health care providers. The extreme case is when an insurer has all the bargaining power, then the negotiated price will equal to $\tilde{p}_0$, and the provider will gain zero profit from having an agreement with this insurer. The relation between the negotiated price and MLR regulation depends on insurers' best response function $\phi(\cdot)$ and demand function $D(\cdot)$. From previous section, I obtain the best response function $\phi(\tilde{p})$. It allows me to solve out the ϕ' in the denominator, details are in the appendix section B. Applying the implicit function theorem to the equation (1.4), I obtain

$$\phi' = \frac{(1-\lambda)\kappa}{(1-\lambda\bar{R})A}, \text{ with } A = 2 - \frac{DD''}{(D')^2}. \tag{1.7}$$

The above equation shows that both the curvature of demand function and MLR regulation affect how optimal premium responds to changes in serve price. Essentially, the MLR regulation reduces the sensitivity of premium to the service price. As proved in the Appendix section A, $0 < \frac{1-\lambda}{1-\lambda\bar{R}} < 1$. Then premiums will be less responsive to changes in price, no matter an increase or a reduction, if insurers are affected by MLR regulation.

Multi-product scenario

In the real world, insurers provide more than one plan and strategically select premiums to maximize total profit. Thus, it is necessary to generalize previous analysis to a multi-product scenario. Similar to the single-product case, I apply the envelope theorem and use the relation in equation (1.4) to obtain the equation below, which is analog to equation (1.6),

$$\tilde{p} = \tilde{p}_0 - \frac{\mathbf{1}^T \boldsymbol{D}}{\frac{\tau}{(1-\tau)} \frac{(1-\lambda)\boldsymbol{\kappa}^T \boldsymbol{D}}{(1-\lambda\bar{R})[\boldsymbol{J}^{-1}\boldsymbol{D}]^T \boldsymbol{D} + C^F} + \mathbf{1}^T \boldsymbol{J} \frac{\partial \boldsymbol{\phi}}{\partial p}}. \tag{1.8}$$

The equation above explicitly shows the effect of regulation (λ) and bargaining power (τ) on the service price. The difference between single-product and multi-product solutions is the Jacobian matrix $\boldsymbol{J}$. When this matrix degenerates to a diagonal matrix, solutions are the same regardless of the number of plans offered by an insurer. When the off-diagonal entries in the Jacobian matrix are not zero, plans subsidize each other and achieve the MLR requirement as a whole.

For the $\frac{\partial \boldsymbol{\phi}}{\partial p}$, I again use the implicit function theorem and derive the following equation from equation (1.4),

$$\left(\begin{bmatrix} [\boldsymbol{J}^{-1}\boldsymbol{D}]^T H^1 \\ \vdots \\ [\boldsymbol{J}^{-1}\boldsymbol{D}]^T H^J \end{bmatrix} - 2\boldsymbol{J} \right) \frac{\partial \boldsymbol{\phi}}{\partial p} = -\frac{1-\lambda}{1-\lambda\bar{R}} \boldsymbol{J}\boldsymbol{\kappa}. \tag{1.9}$$

H^j is the Hessian matrix of plan j's demand function with respect to premiums of plans offered by the same insurer. The term in the large parentheses corresponds to the A in equation (1.7). The technical details are in the appendix section B.

1.4.4 Graphical illustration

I use the framework sketched above to provide a graphical representation of the effect of MLR regulation. The graph shows a simplified situation where the insurer provides only one plan. This simplification affects the estimation but does not change the general conclusion that MLR regulation rules out bargaining equilibria with low negotiated service prices. In addition, I assume that the profit will be zero for both the insurer and the provider if they do not reach any agreement on the service price.

Setup

Consider an insurer I that offers one insurance plan, and a health care provider H that offers one medical service. In the absence of MLR regulation, both the profit maximization problem and the Nash product maximization problem are unconstrained.

More specifically, at the first stage, the insurer and the provider maximize the Nash product of their profits with respect to the service price p. The Nash product is defined as

$$\Pi^{Nash} = \left[\phi D(\phi) - p\kappa\theta D(\phi) - C^F\right]^{\tau} \left[(p - c_h)\theta D(\phi)\right]^{(1-\tau)}.$$

At the second stage, the insurer maximizes the profit by setting the premium

$$\phi = arg \max_{\phi} \left[\phi D(\phi) - p\kappa\theta D(\phi) - C^F\right].$$

Price negotiation and premium choice

As shown in the previous sections, I solve this simplified model by backward induction. The solution from the second stage is insurer's best response function,

$$\phi(p) = p\kappa\theta - \frac{D}{D'}. \tag{1.10}$$

This is a special version of equation (1.4)—when MLR regulation is not effective. The second term on the right-hand-side is positive as I assumed the demand decreases in the premium. The optimal premium is increasing in service price if and only if $2(D')^2 - DD'' > 0$. This condition is satisfied when the demand function is concave or mildly convex, such as a linear demand function and a logit-form demand function. Appendix section C discusses the conditions for an increasing best response function.

Having the best response in mind, the insurer negotiates the service price with the health care provider. The negotiated price solution in this simplified scenario is

$$p = c_h - \frac{\frac{D^2}{D'} + C^F}{\frac{\tau}{1-\tau}\kappa D + \frac{D'}{D}\phi'(\frac{D^2}{D'} + C^F)}. \tag{1.11}$$

When the insurer has all of the bargaining power, $\tau \to 1$, the service price is equal to the cost of the provider. In this case, the provider does not enjoy any profit while the insurer takes all the profit. When the health care provider has all the bargaining power, $\tau = 0$, the equilibrium price is $\bar{p} = c_h - \frac{D(\phi(\bar{p}))}{D'(\phi(\bar{p}))\phi'(\bar{p})}$. This is the highest price that the insurer would accept. Any price higher than $\bar{p}$ means negative profit for the insurer.

Figure 1·1 shows the insurer's best response and the corresponding profit. On the left panel, the x-axis represents the negotiated price; the y-axis represents the premium. When the insurer and the provider bargain on the price, the equilibrium price moves between c_h and $\bar{p}$ along the x-axis. The location of the negotiated price depends on the bargaining power τ. When $\tau = 1$, $p = c_h$. When $\tau = 0$, $p = \bar{p}$. For each possible negotiated price, the blue curve shows the insurer's best premium choice as a response to the price. Along the best response curve, points A and C correspond to the equilibria when the insurer has all the bargaining power and none bargaining power, respectively. The dashed line $p\kappa\theta$ represents the marginal cost to the insurer. On the right panel, the black curve is the demand function for insurance plans. Then the profit could be represented by a rectangle below the demand curve. The real profit equals the premium revenue minus the medical cost minus the fixed cost. To simplify the figure, I show the raw profit—real profit plus fixed costs—rather than the real profit. In other words, the raw profit is the product of demand and the difference between the premium and marginal cost. This simplification does not affect the comparison between profits corresponding to different (p, ϕ) pairs within a firm. For example, as shown in the figure, the yellow shaded area represents π_A, the raw profit corresponding to the point A where the insurer has all the bargaining power.

Effect of MLR regulation

When MLR regulation takes effect, the insurance firm maximizes its profit subject to a minimal ratio $\bar{R}$ for medical loss. That is, for the insurer,

$$p\kappa\theta \geq \bar{R}\phi(D).$$

Figure 1·2 shows how the regulation affects the bargaining equilibrium and insurer's optimal premium choice. The orange line represents the MLR threshold, and the area underneath the line, the orange shared area, indicates all the eligible (p, ϕ) pairs under the regulation. In other words, if the insurer has little bargaining power so that the optimal solution without MLR regulation locates between points B and C, then the insurer will not be affected by the regulation. However, when the no-regulation optimal premium choice is between points A and B, that is when the insurer has large bargaining power and low negotiated price, the insurer has to either lower the premium and/or concede part of her profit to the provider to reach the MLR requirement. Continue using the example in the Figure 1·1. The MLR at point A is below the regulatory threshold. If the service price is fixed, the insurer will have no choice other than moving to the point A' and gain the profit $\Pi_{A'}$ indicated by the green shaded area. When the service price is negotiated, the insurer could move to other locations between points A' and B. For example, the insurer could move to point A'' and gain the profit $\Pi_{A''}$ indicated by the blue shaded area. Where locates the new optimal choice is an empirical question. It depends on the profit corresponding to each point, which is related to the curvature of the demand function. As for the premium, it could become lower or higher than ϕ_A. Again, it depends on the relation between the regulation threshold and the curvature of the demand function.

This simple model shows that if an insurer has little bargaining power and a high negotiated price, MLR regulation will not affect the price. For an insurer with large bargaining power and a low negotiated price, MLR regulation will rule out the (p, ϕ) combination that the insurer initially chooses. Consequently, the service price increases under the above condition. Whether the premium increases or decreases depends on the relative concavity of the demand function to the regulation threshold.

1.5 Data

To empirically estimate the effect of MLR regulation, I apply the model developed in the previous section to the ACA Health Insurance Exchange Marketplaces. I use data from multiple sources to construct my sample. The two primary data sources this paper relies on are healthcare.gov and the Center for Consumer Information & Insurance Oversight (CCIIO). The former is the website where consumers shop for their insurance plans. From the website, I gather data on premiums and other plan characteristics. In this dataset, each plan has a state-specific plan ID, which is uniquely defined by state, insurer, and plan.

From the CCIIO, I obtain the enrollment data and insurers' MLR reports. The unit of observation in the enrollment data is state-plan. I merge the enrollment data with the plan characteristics by plan IDs and generate a dataset of demand, premium, and plan characteristics. A caveat with the enrollment data is that the number of enrollees is aggregated to the state level, not at county or rating area level. Therefore, I use the average premium as the state-level premium. Insurers' MLR reports are filed annually and contain financial information such as MLR, total revenue, total cost, and other insurer information such as the Employer Identification Number (EIN), whether it is for-profit, etc. These data are merged to the main dataset by state-specific issuer IDs. In the following analysis, I use this state-specific issuer ID as the insurer ID. The EIN is used for generating instruments, described in section 1.6.1.

To complete the dataset, I obtain the market and population demographics data from the Area Health Resources Files (AHRF). More specifically, I estimate the income distribution in each state using the statistics of uninsured people's income groups and use the number of uninsured people as the market size.

The final sample has 7,570 unique plan IDs offered by 306 insurers (238 EINs) in 149 state-year markets. That are 12,384 plan-year observations. Table 1.3 shows the descriptive statistics of key variables in the sample.

As mentioned earlier, plan features could vary within the same metal category. In this paper, I include the following plan characteristics. The *deductible* is the amount an enrollee pays for health care services before the insurance plan starts to pay. After reaching the deductible line, the enrollee will pay either a fixed amount—*copayment*—for the health care service or a fixed fraction—*coinsurance*—of the total spending. When the total out-of-pocket payment—the sum of deductibles, copayments, and coinsurance—reaches the Out-Of-Pocket maximum (*OOP max*), the insurance plan will cover all spendings after that.

Figure 1·3 shows the average of premiums and other plan characteristics by metal category and year. Panel A shows that, on average, premium increases from catastrophic plans to platinum plans. It also increases significantly over the years. For the deductible and OOP max, the difference across metal categories is significant, but the over-year changes are much smaller than those in premiums. Panel D and E present the copayment and coinsurance, respectively. Because only 52.35% of plans have copayment and 38.53% have coinsurance, the variation of these two plan characteristics is much larger. As Parys (2018) shows, premiums increased significantly from year to year. The time trend of other plan characteristics is relatively flat. This variation in premiums and plan characteristics is one of the key identification sources in my analysis.

Another important source of variation in my analysis is from MLRs. Figure 1·4 shows the MLRs reported by insurers from 2014-2017. Grey bars and red contour bars present the preliminary MLR and credibility-adjusted MLR, respectively. Most of the insurers have MLRs between 0.8 and 1. Those after the credibility-adjustment are higher than the preliminary ones overall. Callaghan, Plummer and Wempe (2020) find that, all markets pooled together from 2011-2015, MLRs of both non-compliant and compliant plans move towards to the threshold over years. In the Marketplace, from 2014-2017, the width of MLR distribution does not shrink and it is relatively continuous at threshold 0.8, as shown in the appendix figure A1.

This study grapples with the same two major data limitations as in other studies. The first one is that I do not observe negotiated prices. To address this problem, I structurally introduce insurer premium choices into the bargaining solution. Second, because the marketplace started in 2014, three years later than the regulation, pre-regulation data do not exist. Therefore, reduced-form approaches are not applicable in this case.

1.6 Empirical Strategy

In this section, I present the empirical estimation strategy which is based on the theoretical model developed in section 1.4.

1.6.1 Demand for insurance plans

In the discrete choice model for the insurance plan demand, a couple of variables and distributions need to be further specified. First, the plan characteristics X_j includes plan j's deductible, OOP max, copayment, coinsurance, metal category, and plan type. These are the most salient plan characteristics that consumers could observe in the marketplace.[4] Second, for the demographic distribution, I use income group relative to the federal poverty level. Another important variable for the demand estimation is the market size M_t. Following previous studies (Tebaldi, 2017; Drake, 2019; Saltzman, 2019), I assume that the outside option is uninsured. The uninsured population under age 65 represents the "demand" of outside option.

The premium is endogenous in the model. Therefore, I use three sets of instrument variables together to mitigate the endogeneity problem. The first set contains the average plan characteristics of plans offered by the rivals in the same market. Those instruments are missing in single-insurer markets. Therefore, I use an indicator of single-insurer markets and replace missing values by zero in the average plan characteristics. The second set of instruments includes the average premiums and the average plan characteristics of plans offered by the same insurer in other markets. The same insurer is identified by the EIN. Because there are local insurers that only participated in one market (i.e., one year in one state), I use a similar method to deal with missing values—add an indicator of single-market insurer and assign zero to those missing values in the average instruments. The last set of instruments includes cost shifters. It includes the average Medicare wage index across all core-based statistical areas where the plan was offered.

[4] Due to the potential collinearity issue, the main specification leaves OOP max out. As OOP max is a threshold of the accumulative amount of patient OOP payments, only patients who use lots of health care services will reach this threshold. Therefore, I assume consumers care less about the OOP max than other features.

1.6.2 Effect of MLR regulation on marginal cost

Before estimating the full model, I examine the effect of MLR regulation on insurers' marginal cost, without including the price negotiation.

Insurers' best response function, the equation (1.4), establishes the relation between marginal cost and marginal revenue under MLR regulation. That is

$$mc = \left(\bar{R} + \frac{1-\bar{R}}{1-\lambda}\right) mr,$$

with $mc = \tilde{p}\kappa\theta$ and $mr = \phi + \boldsymbol{J}^{-1}\boldsymbol{D}$. From the demand estimation, the marginal revenue mr_j is calculable. However, the marginal cost is unobservable. Therefore, I adopt a widely used assumption about marginal cost—that the log of marginal cost is a linear function of product characteristics. With this assumption, I estimate the effect of MLR regulation by the following equation

$$ln(mr_j) = Const - ln\left(\bar{R} + \frac{1-\bar{R}}{1-\lambda\mathbb{1}(Rebate_{f(j)})}\right) + w_j\gamma + \omega_j, \tag{1.12}$$

where $\mathbb{1}(Rebate_{f(j)})$ is an indicator of whether the insurer f needs to rebate, w_j is a vector of plan characteristics including same variables as in X_j, λ measures the effect of MLR regulation on non-compliant insurers. From the discussion in the section 1.4, I expect that $\lambda > 0$. A positive λ implies that MLR regulation breaks the efficient equality–$marginal\, cost = marginal\, revenue$. Affected by the MLR, marginal cost is higher than the marginal revenue.

1.6.3 Effect of MLR regulation and price negotiation

Using the model developed in section 1.4, I derive a set of moments for the estimation. I start by adapting my model to the available data. The heterogeneity in the service prices is an interesting topic but requires a richer dataset. As this paper focuses on the essence of price negotiation, I use one representative price for all the health services. More specifically, I use the average spending as a proxy of the service price. Moreover, due to the data limitation, I use the average spending per enrollee instead of the average spending per patient. In this way, θ is not identifiable, thus, I use p and p_0 instead of $\tilde{p}\theta$ and $\tilde{p}_0\theta$, respectively. For the alternative price p_0, I use the average Medicaid spending per enrollee as a proxy and denote it as p^{MCD}. This payment can be considered as the lowest payment a provider receives.

One big challenge for the estimation is that negotiated prices are not observed. In this part of estimation, instead of making an assumption about marginal cost, I combine equation (1.4) and equation (1.8) to eliminate service price in the final estimation moments.

After those changes, I am able to jointly estimate λ, τ, and C^F using two-step GMM under the assumption that $E[Z'\xi] = 0$. The error term ξ is defined as

$$\begin{aligned}\xi = \quad & \frac{\tau(1-\lambda\bar{R})(\boldsymbol{\phi} + \boldsymbol{J}^{-1}\boldsymbol{D})^T\boldsymbol{D}}{(1-\lambda\bar{R})[\boldsymbol{J}^{-1}\boldsymbol{D}]^T\boldsymbol{D} + C^F} - \frac{\tau(1-\lambda)p^{MCD}\boldsymbol{\kappa}^T\boldsymbol{D}}{(1-\lambda\bar{R})[\boldsymbol{J}^{-1}\boldsymbol{D}]^T\boldsymbol{D} + C^F} \\ & - \frac{(1-\tau)(1-\lambda)}{1-\lambda\bar{R}} p^{MCD} \frac{\boldsymbol{1}^T\boldsymbol{J}\boldsymbol{A}^{-1}\boldsymbol{J}\boldsymbol{\kappa}}{\boldsymbol{1}^T\boldsymbol{D}} - (1-\tau)(\frac{\boldsymbol{1}^T\boldsymbol{J}\boldsymbol{A}^{-1}\boldsymbol{J}(\boldsymbol{\phi} + \boldsymbol{J}^{-1}\boldsymbol{D})}{\boldsymbol{1}^T\boldsymbol{D}} + 1)\end{aligned} \quad (1.13)$$

In the equation, $\boldsymbol{\phi}$ and $\boldsymbol{D}$ are observed, $\boldsymbol{J}$ and $\boldsymbol{H}$ are calculated based on the demand estimates. Appendix section B shows how $\boldsymbol{J}$ and $\boldsymbol{H}$ are calculated. For the instruments in Z, I use the number of firms in the market, year dummies, and state dummies.

1.7 Empirical Results

1.7.1 Demand for insurance plans

Before estimating the demand with random coefficients, I explore the relation between demand and plan characteristics by nested-logit regressions.

$$ln(s_{jst}) - ln(s_{0st}) = \alpha ln\phi_{jst} + X'_{jst}\beta + \sigma ln s_{j/g} + \gamma^s_s + \gamma^t_t + \gamma^f_{f(j)} + \xi_j \quad (1.14)$$

In the equation, s_{jst} and s_{0st} are market shares of plan j and outside option in state s and year t, respectively. ϕ_{jst} is plan j's average premium. X_{jst} is a vector of plan characteristics. $s_{j/g}$ is the nested market share where g represents the metal category that plan j belongs to. γ^s and γ^t are state fixed effects and year fixed effects, respectively. $\gamma^f_{f(j)}$ is insurer fixed effect, $f(j)$ indicates the insurer who provides plan j. Since the conditional market share $s_{j/g}$ is also endogenous, I use the number of plans in the nest as an additional instrument.

Table 1.4 presents the estimates of demand for insurance plans in the exchange marketplace. The first two columns show the results of nested-logit regressions while the following columns show the results with random coefficients. In columns (1) and (3), plan characteristics include deductible and OOP max. In the other columns, I replace OOP max with copayment and coinsurance. Across all the specifications, consumers significantly dislike high premium. After controlling for the copayment and coinsurance rates, consumers significantly prefer low deductible. It is interesting to note that consumers significantly prefer the copayment feature and do not care much about the amount of copayment. As for the coinsurance, consumers do not like this feature and are very sensitive to the coinsurance level. This finding echoes what Loewenstein et al. (2013) find—consumers understand copayment the best among all the four plan features. In other words, consumers prefer insurance plans that they can easily understand.

In the following study, I use the model and estimates in column (5) as the demand for insurance plans. More specifically, I use those estimates to calculate Jacobian and Hessian matrices of demand as well as the derivatives of the best response function with respect to the health service price.

Based on the selected specification, the average own-price elasticity is -2.23 (SD 0.51), which locates well in the range of estimates in previous studies (-10.6, -1.7)

(Abraham et al., 2017; Saltzman, 2019; Drake, 2019; Tebaldi, 2017). Figure 1·5 shows the kernel density of own-price elasticity by metal category. From the least to the most generous category, mean elasticity does not move much while the width of distribution shrinks. One outlier is the "catastrophic" category. Enrollees who choose plans in this category are on the borderline between purchasing and not. The coverage of "catastrophic" plans is very limited. Thus the price elasticity is relatively high.

1.7.2 Effect of MLR regulation

From the demand estimation, I obtain plan-level marginal revenue $mr_j = \phi_j + [\boldsymbol{J}^{-1}\boldsymbol{D}]_{jj}$. Figure 1·6 shows the distribution of marginal revenue for compliant and non-compliant insurers . In the figure, the marginal revenue of plans provided by non-compliant insurers are higher than that of plans provided by compliant insurers. Using these calculated marginal revenues, I estimate equation (1.12) to tease out the effect of MLR regulation on marginal cost by non-linear least squares.

As shown in the table 1.5, λ is significantly different from zero, regardless of model specifications. From column (1) to (4), I add year fixed effects, state fixed effects, and firm characteristics one by one. Inspired by previous studies (Dafny, 2019; McCue, Hall and Liu, 2013), I include a not-for-profit indicator and the number of counties the insurer enters as firm characteristics. With more controls, the magnitude of the effect reduces. Column (5) includes the same control variables as column (4) except using OOP max instead of copayment and coinsurance. The estimate of λ is robust to this change of plan characteristics. In the setting of column (4), $\lambda = 0.0568$ means that, compared to compliant insurers, 9.7% of non-compliant insurers' marginal cost is induced by the regulation given plan design. In other words, the ratio of marginal cost to marginal revenue equals one when there is no regulation. When the MLR regulation is effective, this ratio increases to 1.097, which suggests an inefficiency.

That increase could be driven by an increase of marginal cost and/or a decrease of the marginal revenue at equilibrium.

Among all the plan characteristics, the most significant contributor is copayment. As expected, higher copayment relates to lower marginal cost and, therefore, lower marginal revenue. The constant term estimates, shown in the last row in the table, implying that the marginal cost of the base plan—a plan without any cost-sharing in the catastrophic category—is about $1700, on average.

1.7.3 Effect of MLR regulation and price negotiation

Combining the demand estimates with the full model, I jointly estimate the bargaining parameter, the effect of MLR regulation, and insurers' fixed cost. Table 1.6 presents the results with different specifications. Because the estimation in this part relies on the demand estimation discussed in section 1.7.1. I use a parametric bootstrap method to calculate the standard errors (Efron and Tibshirani, 1986). More specifically, I assume that the demand parameters follow the distributions estimated by the demand model, and bootstrap based on those estimated distributions to evaluate the accuracy of the point estimates in this section.

Column (1) in Table 1.6 shows the baseline estimates, where I assume that insurers have the same bargaining parameter and same fixed cost. Following that, I allow non-compliant insurers to have a different bargaining parameter where τ_1 measures the difference. Results displayed in column (2) suggest that non-compliant insurers have an insignificantly lower bargaining power although the magnitude is large—28% less powerful than compliant insurers on average. In column (3), all insurers have the same bargaining parameter, and not-for-profit insurers could have a different fixed cost. The estimates suggest that, compared to for-profit insurers, not-for-profit insurers have higher fixed cost. The last column combines the specification of columns (2) and (3).

The first row presents the estimates of λ—a measure of the effect of MLR regulation. The results are relatively robust to changes in the specifications. To interpret λ, I use estimates in column (4) to calculate the negotiated service prices with and without the MLR regulation, assuming all other settings are unchanged. Table 1.7 shows a summary of the calculated negotiated prices. The average negotiated price of compliant insurers does not change when there is the MLR regulation. For non-compliant insurers, the MLR regulation raises the negotiated price by \$86.62. Those changes in negotiated price translate into \$21.3 million and \$62.8 million more OOP payment for patients enrolled in compliant plans and non-compliant plans every year, respectively. This extra OOP payment is for health care services, excluding any changes in the premiums. To answer what would premiums be and what is the total welfare effect, I conduct the following counterfactual analysis.

1.8 Counterfactual Analysis

To study the welfare effect and to decompose the total effect into a regulation effect and a bargaining effect, I compare price and premiums at equilibrium in the following three scenarios:

1. No regulation on profit and health service prices are negotiated;
2. MLR regulation is effective, but prices are fixed at no-regulation negotiated level;
3. MLR regulation is effective, and prices are negotiated;

Public health insurance option is a widely debated option for improving health care affordability. In the counterfactual analysis, following the previous three scenarios, I constructed the fourth one to test if a public option could be a solution here.

4. No regulation on profit, health service prices are negotiated, one public option.

In all the scenarios, I constructed a synthetic market of 100,000 individuals and three insurers. Each insurer provides one silver plan with different plan characteristics shown in the table below.

	Insurer1 (plan A)	Insurer2 (plan B)	Insurer3 (plan C)
Deductible ($)	2000	3000	5000
Copayment ($)	10	25	30

In the first scenario, there is no regulation on insurer profit. Insurers and health care providers negotiate service prices, and insurers set the premium to maximize the profit. This setting would mimic the market if the Marketplace existed before the MLR regulation. When there is no regulation on the profit, insurers will exploit the demand and pricing freely. The top panel in Table 1.8 shows the solution at equilibrium in this scenario. The negotiated service prices are around $2700, and the optimal premiums are around $2600. Of the 100,000 individuals in the market, about 52,000 purchased insurance plans. The uninsurance rate is higher in this constructed market than in the real market because I eliminate cheap plans for simplification. Together, it means around $45 million of premium revenue and $33 million of medical loss for one insurer.

The second scenario is when the MLR regulation is effective, and prices are fixed at the no-regulation level. The no-regulation level is the negotiated result when there is no MLR regulation, as in the first scenario. This setting allows me to measure how much the regulation would bring down premiums if the price negotiation channel were shut down. With the same level of service prices, because of the MLR regulation, insurers have to lower premiums to achieve the requirement. Therefore, premiums in this scenario are lower than those in the first scenario. Consequently, demands and medical loss are higher than in the previous case. Despite the increase in demands, the reduction in premiums drives to lower premium revenues for all three insurers. The second panel in Table 1.8 presents the results at equilibrium in this setting.

Comparing to scenario 1, if service prices were fixed, the MLR regulation would limit insurer profit and make the insurance plans more affordable for more people. The average enrollee would spend $224 less on insurance plans. With lower premiums, 1.6% of individuals would become insured.

The third scenario uses the real world settings that described in the previous sections—the MLR regulation is effective, and service prices are negotiated by insurers and health care providers. In the constructed market, as shown by the third panel in Table 1.8, the negotiated prices increase to $4200, and the premiums increase to $3700. Although the demands reduce by more than one thousand for each plan, the premium revenues are higher than the other scenarios. The pre-post regulation comparison shows that, changing from scenario 1 to 3, insurers will allow higher service prices, concede part of the profit to health care providers, and increase premiums. Although the demand drops, the total profit remains at the same level as before regulation.

In the last scenario, plan C is a public option with a fixed MLR at 0.8 level and the other two plans are private plans that are not regulated by the MLR regulation [5]. As shown in the bottom panel in Table 1.8, insurers bargain hard to lower health service price. For the public option, because it has a fixed MLR, the premium is set at a low level, like in the second scenario. For the private options, premiums are lower than the no-regulation level because of the competition induced by the public option in the insurance market. With the public option, both health insurance plans and health care services become more affordable compare to the no-regulation scenario.

Figure 1·7 shows the changes in consumer welfare if the market setting changes from that of the pre-regulation scenario to the other three scenarios. When health service prices were fixed at the pre-regulation level, MLR regulation could save consumers $7.95 million for insurance plans while slightly more spending, $1.31 million on health care services due to higher demand. Together, the consumer welfare in-

[5] Results are robust regardless which plan is turned into the public option.

creases by 4%. When health service prices are negotiated under the MLR regulation, consumers will need to pay $67 million more out of pocket——$47 million for the insurance plans and the other $20 million for health care services. The last pair of bars shows the savings if we have a "regulated" public option instead of regulating private insurers on their MLR. The results suggest that the introduction of the public option could save consumers $1.75 million for medical services and $7.18 million for health insurance plans, which in total is a 5% increase in consumer welfare.

1.9 Conclusion

This paper builds a structural model of the effects of Medical Loss Ratio regulation on health service prices and insurance premiums. The model incorporates the regulation and insurer-provider price negotiation. This constrained bargaining model reveals how MLR regulation leads to higher health service prices via price negotiation. I also provide a graphical representation of the effect, which illustrates the intuition that MLR regulation rules out bargaining equilibria with low service prices. Vertical relation—e.g., price negotiations—could open a channel through which the effect of regulations passes from the regulated side to the other side. Particularly, in markets where the private provision of public service is essential, a successful regulation needs to align private firms' incentives with the social optimum.

Applying this model to the ACA Health Insurance Exchange Marketplace, I estimate the effect of MLR regulation on prices, premiums, and patient welfare. My estimates imply that, along with the increase in MLRs, the MLR regulation results in higher health service prices and higher patient OOP payments. If MLR regulation were implemented in a market where health care service prices were fixed, consumers would spend less out of pocket. However, because of the insurer-provider price negotiation, consumers need to spend more when MLR regulation is effective.

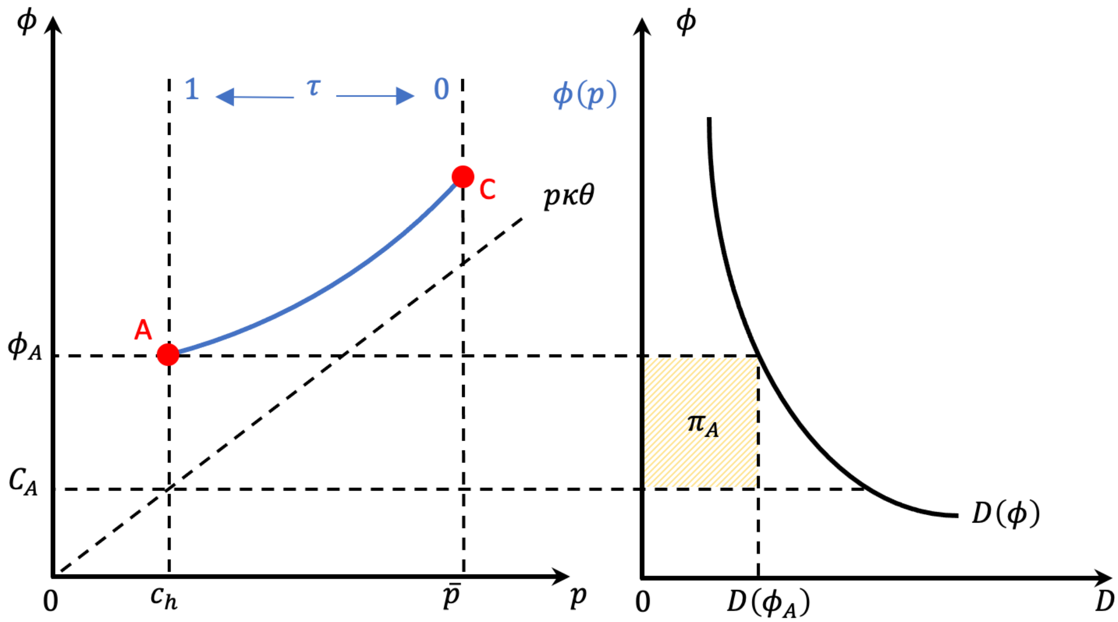
This paper examines pricing behavior in a static and partial equilibrium framework, assuming that demand for health care services only varies proportionally to the demand for insurance plans, without adjusting for any preemptive pricing behavior among insurers. Therefore, the conclusion from this paper should be interpreted as short-run effects. As data availability increases over time, future work could extend my model by adding dynamic components.

I also do not consider heterogeneity among health care providers within a market. Estimates in this paper provide an average estimation of the effect of MLR regulation on health service prices. Both the heterogeneity in the negotiation—e.g., hospitals and independent practice groups might have different bargaining parameters—and the heterogeneity in the menu of service prices—e.g., effects on low-value services and high-value services might be different—could help understand the impact of this regulation on the provider market. However, I argue that this heterogeneity is a second-order concern relative to the direct effects of MLR regulation on insurer pricing, as it does not significantly alter insurer behavior. Future work could build this heterogeneity into the model; alternatively, future work might integrate provider networks into estimation. Ho and Lee (2017) and Ho and Lee (2019) model the bargaining on network formation before the ACA. More data of the provider market is needed to introduce this into the model empirically.

From the perspective of patients, a better health outcome could be more important than the OOP payment. It is still unclear whether the higher medical spending translates into better health status. To answer this question, researchers will need good measurements of service utilization and health outcomes.

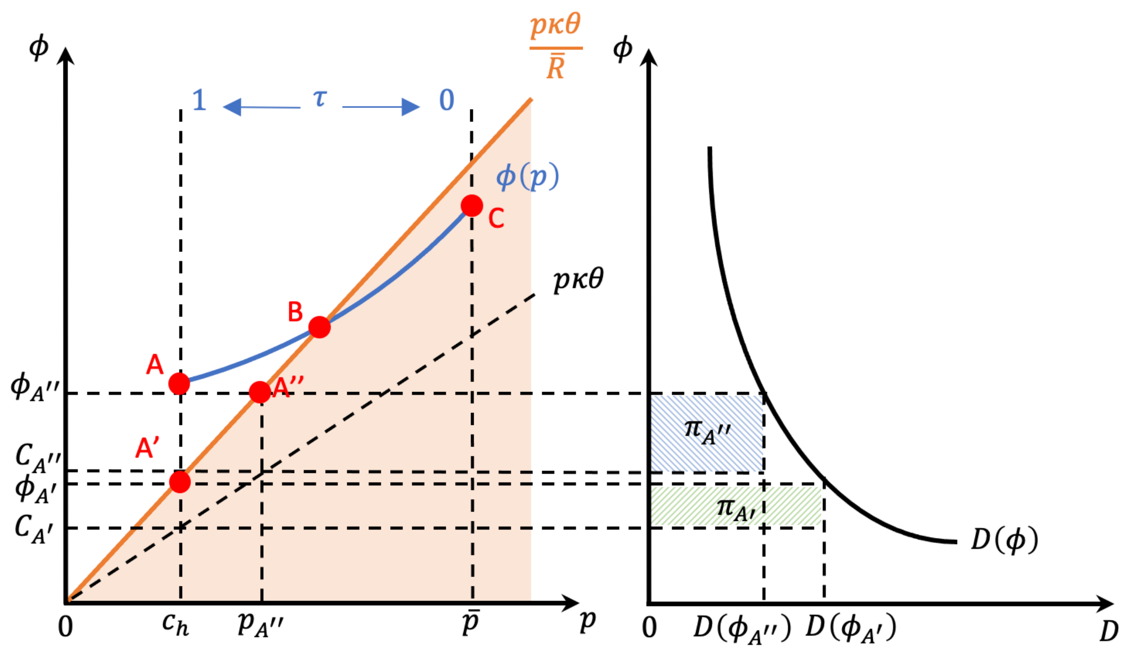
Besides its MLR regulation, there are many other provisions under the ACA. Some of them—such as risk-corridors, reinsurance, and the individual mandate—could affect this paper's estimates. For example, the individual mandate could affect insurer pricing because the penalty level affects the outside option for consumers and, therefore, the curvature of demand function. It is an exciting agenda to incorporate those related provisions and tease out the effect of each provision and how they interact with each other.

Figure 1·1: Insurer's best response and the negotiated health service price



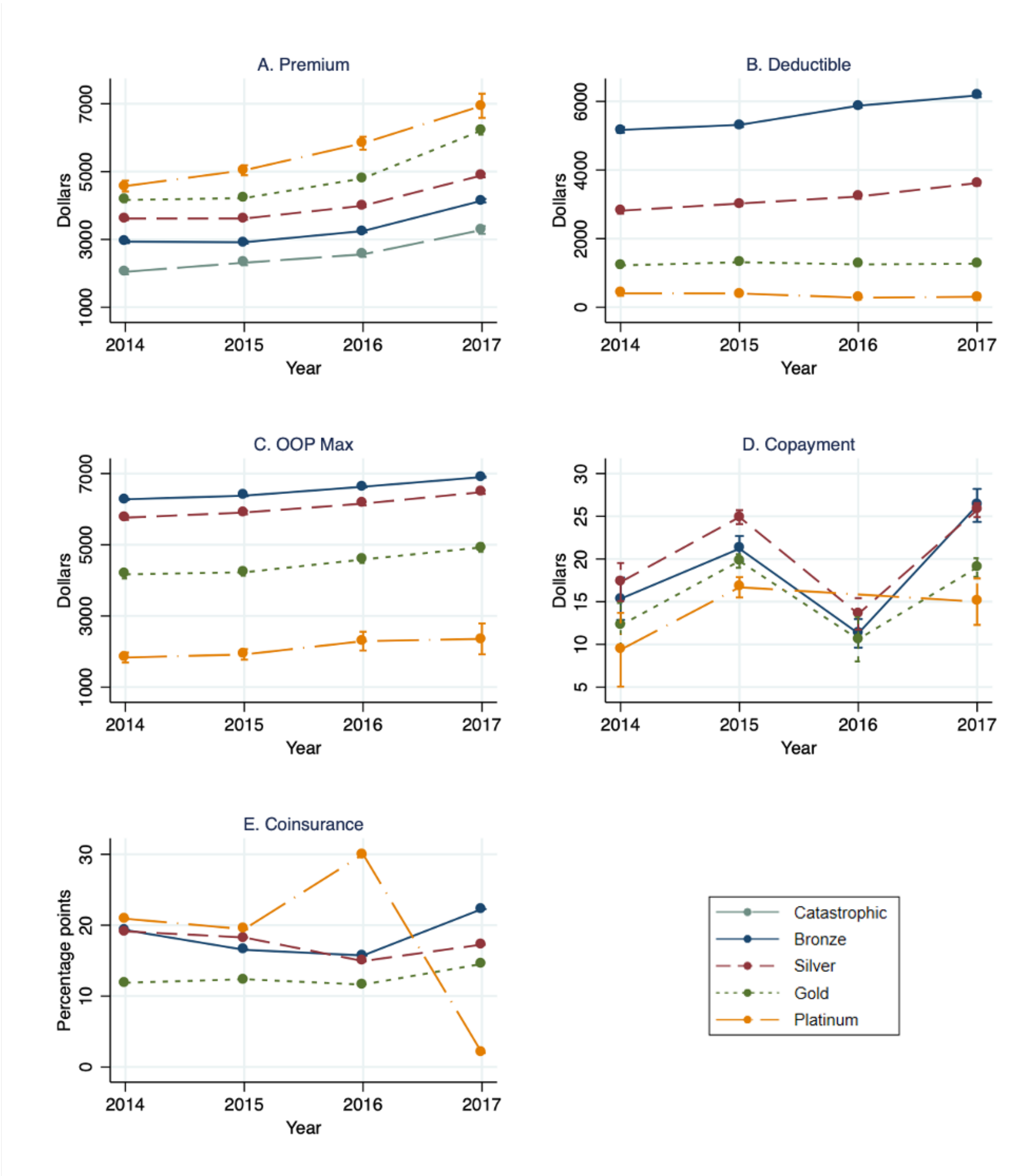
Notes: This figure represents the negotiated prices and insurer's premium choice when the insurer only provides one insurance plan without MLR regulation. The negotiated price varies along the x-axis in the left panel. Insurer's bargaining power τ affects where the negotiated price locates between c_h and $\bar{p}$. The solid curve in the left panel depicts insurer's best response of premium to price. Points A and C indicates what are the negotiated price and premium when insurer has all bargaining power or none bargaining power, respectively. In the right panel, the solid curve is the demand for insurance plan. The shaded area shows the profit (including fixed cost) if the insurer has the max bargaining power.

Figure 1·2: Impact of MLR regulation



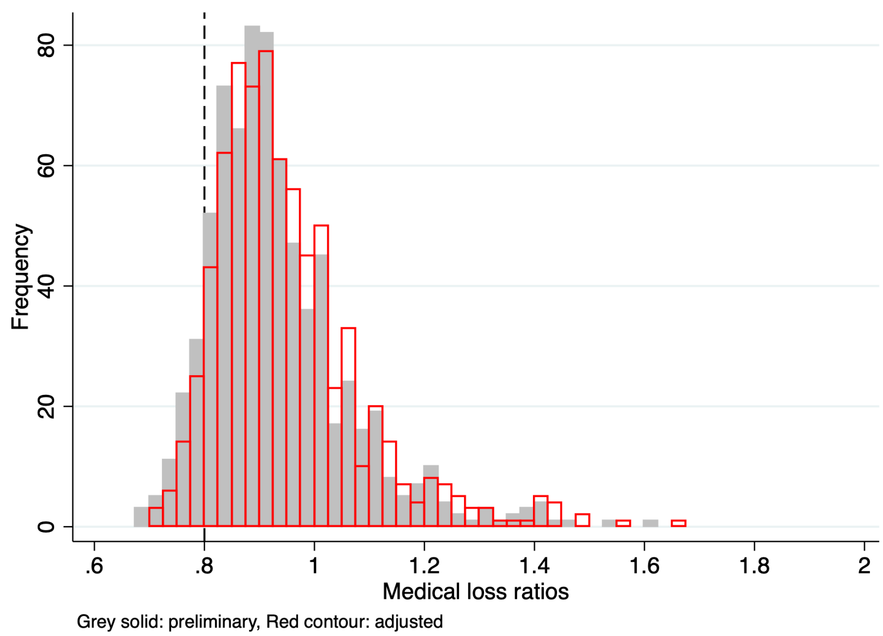
Notes: This figure presents how the MLR regulation affects the choice of negotiated price and premium. The solid line in the left panel represents MLR regulation threshold and the shaded triangle below that line indicates all the eligible choices of price and premiums under MLR regulation.

Figure 1·3: Variation in premium and plan characteristics, by metal category and year



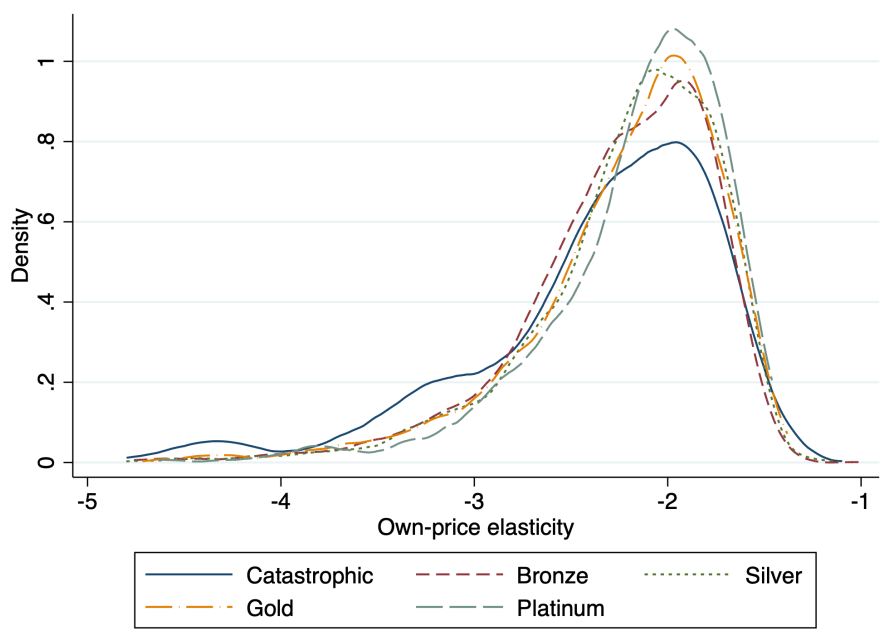
Notes: This figure shows the variation of premium, deductible, Out-Of-Pocket maximum, copayment, and coinsurance, by metal category and year. The bars show the 95% confidence interval of the means.

Figure 1·4: Distribution of MLRs



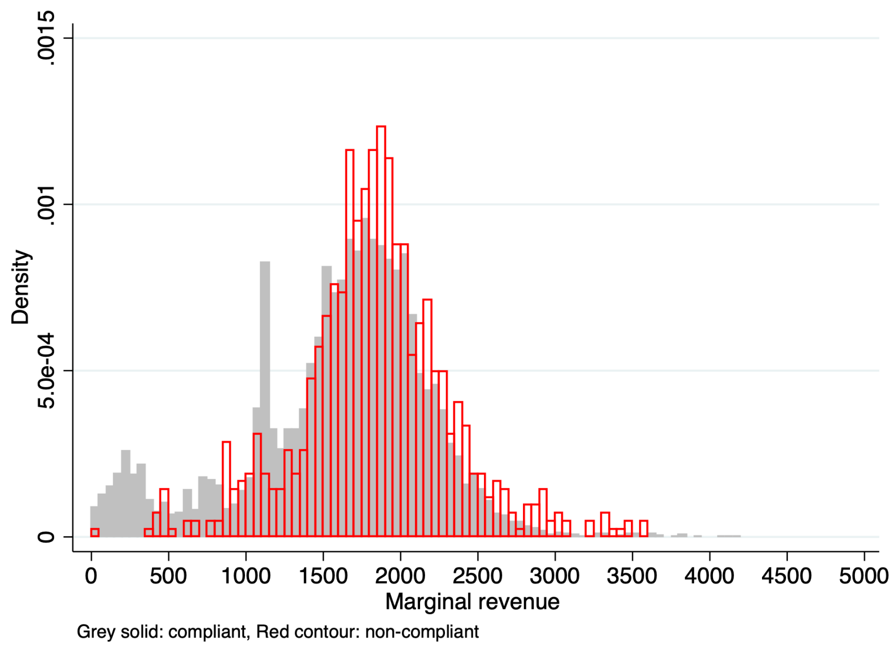
Notes: This figure presents the distribution of medical loss ratios. Data is from the MLR reports. The sample includes insurers in the Federally-Facilitated Marketplace from 2014-2017. Grey solid bars present the preliminary MLRs. The red contoured bars present the credibility-adjusted MLRs.

Figure 1·5: Distribution of own-price elasticity of demand



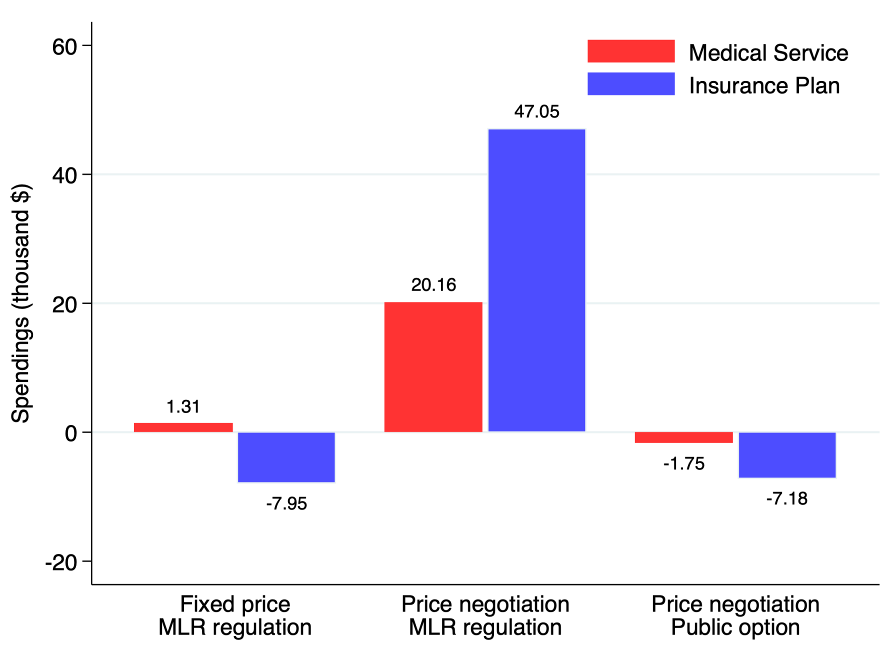
Notes: This figure presents the empirical density distribution of own-price elasticity of demand for insurance plan. The elasticities are calculated based on the demand estimation. Unit observation is plan.

Figure 1·6: Empirical density of estimated marginal revenue



Notes: This figure presents the density distribution of marginal revenues at plan-year level. The marginal revenues are calculated based on the demand estimation. The compliance status are at insurer-year level.

Figure 1·7: Consumer welfare changes compared to the pre-regulation scenario



Notes: This figure shows the changes in consumer welfare by type of spending. Comparing to the pre-regulation scenario simulated in the counterfactual analysis, the two bars, red and blue, for each of the other three scenarios present the changes in patients' medical service spending and insurance spending, respectively.

Table 1.1: Adjustment factors of medical loss ratio

Panel A. Base credibility factor

Life Years	Base Credibility Factor
<1,000	Not Credible
1,000	8.3%
2,500	5.2%
5,000	3.7%
10,000	2.6%
25,000	1.6%
50,000	1.2%
75,000	0.0%

Panel B. Deductible factor

Deductible	Deductible Factor
$0	1.000
$2,500	1.164
$5,000	1.402
$10,000	1.736

Notes: This table shows the two adjustment factors used in the ACA MLR regulation.

Table 1.2: Actuarial value of metal category

Category	Actuarial value
Catastrophic	NA
Bronze	60%
Silver	70%
Gold	80%
Platinum	90%

Notes: This table presents names and corresponding actuarial values of the five categories of plan in the Marketplace.

Table 1.3: Summary statistics of key variables

Panel A. Continuous variables

	Mean	Std. Dev.	Min	Max
N enrollees	2663	9953	11	274,497
Premium ($)	3939.08	1189.48	1008	13,087
Deductible ($)	3645.06	2153.96	0	7150
Copayment ($)	20.54	17.09	0	150
Coinsurance (percent)	16.6	16.65	0	80
OOP max ($)	5773.81	1443.86	500	7150
Plan market share (percent)	0.308	0.906	0.00019	28.445

Panel B. Category variables

	Freq.	Percent
Metal category		
Bronze	3751	30.29
Silver	4803	38.78
Gold	2596	20.96
Platinum	494	3.99
Catastrophic	740	5.98
Plan type		
EPO	986	7.96
HMO	6320	51.03
POS	1369	11.05
PPO	3709	29.95
Year		
2014	2296	18.54
2015	3511	28.35
2016	3698	29.86
2017	2879	23.25

Notes: This table shows the summary statistics of key variables. Full name of plan type: Exclusive provider organization (EPO), Health maintenance organization (HMO), Point-of-service (POS), Preferred provider organization (PPO)

Table 1.4: Results of demand estimation

	Nested-logit		Random Coefficients		
	(1)	(2)	(3)	(4)	(5)
log(premium)	-0.565	-0.713	-2.32	-2.19	-0.615
	(0.101)	(0.086)	(0.416)	(0.34)	(0.803)
log(deductible)	0.002	-0.008	0.01	-0.031	-0.028
	(0.005)	(0.004)	(0.014)	(0.013)	(0.013)
log(OOP max)	0.099		0.475		
	(0.027)		(0.073)		
Copayment		0.0003		0.0004	0.0004
		(0.0005)		(0.0017)	(0.0018)
Coinsurance		-0.225		-0.956	-0.99
		(0.058)		(0.193)	(0.216)
Having copayment		0.00071		0.199	0.196
		(0.017)		(0.056)	(0.067)
Having coinsurance		-0.052		-0.336	-0.359
		(0.017)		(0.054)	(0.059)
ρ	0.768	0.808			
	(0.013)	(0.011)			
Metal Category FE	X	X	X	X	X
Plan Type FE	X	X	X	X	X
State FE	X	X	X	X	X
Year FE	X	X	X	X	X
RC premium			X	X	X
RC premium (interact with demo)					X

Notes: This table presents the results of demand estimation. The first two columns are from nested-logit regressions with metal category being nest. The following three columns report estimates of discrete choice model with random coefficients. Columns (3) and (4) allow for random coefficients for the premium. Column (5) additionally includes individual income level and allows it to interact with the premium.

Table 1.5: Effect of MLR regulation on marginal cost

	(1)	(2)	(3)	(4)	(5)
λ	0.120	0.119	0.0832	0.0568	0.0554
	(0.00725)	(0.00742)	(0.00830)	(0.00821)	(0.00821)
log(deductible)	0.00909	0.00749	0.00472	0.00889	0.00972
	(0.00490)	(0.00485)	(0.00436)	(0.00432)	(0.00425)
Copayment	-0.000724	-0.00132	-0.00254	-0.00191	
	(0.000618)	(0.000621)	(0.000553)	(0.000549)	
Coinsurance rate	-0.0488	-0.0861	0.140	0.131	
	(0.0630)	(0.0622)	(0.0593)	(0.0597)	
Having copayment	-0.0166	-0.0109	0.0430	0.0215	
	(0.0177)	(0.0182)	(0.0171)	(0.0171)	
Having coinsurance	0.0765	0.0652	-0.0248	0.0116	
	(0.0196)	(0.0191)	(0.0168)	(0.0171)	
log(OOP max)					0.0431
					(0.0220)
Constant	7.257	7.341	7.471	7.441	7.051
	(0.0537)	(0.0532)	(0.0803)	(0.0731)	(0.201)
Year FE		Yes	Yes	Yes	Yes
State FE			Yes	Yes	Yes
Firm characteristics				Yes	Yes
N	10,859	10,859	10,859	10,859	10,859
R^2	0.022	0.035	0.278	0.302	0.300

Notes: This table show the estimates of the effect of MLR regulation on marginal cost, without structurally including the price negotiation. Column (1) shows the baseline estimates, which include metal level fixed effects and plan type fixed effects. From column (2) to column (4), I add year fixed effects, state fixed effects, and firm characteristics one-by-one. Column (5) has same control variables as column (4) except that column (5) use OOP max instead of copayment and coinsurance as plan characteristics. λ is the measure of the effect of MLR regulation.

Table 1.6: Effect of MLR regulation and bargaining

		(1)	(2)	(3)	(4)
Effect of MLR Regulation	λ	0.133	0.143	0.145	0.152
		(0.01)	(0.007)	(0.01)	(0.009)
Nash Bargaining Parameters	τ_0	0.498	0.428	0.31	0.403
		(0.028)	(0.591)	(0.03)	(0.041)
	τ_1		-0.1		-0.035
			(0.089)		(0.216)
Insurer Fixed Cost	C^F	1.285	1.227	0.255	1.226
		(0.041)	(1.366)	(0.065)	(0.035)
	C^F_{NFP}			0.047	0.063
				(0.02)	(0.015)
N observations		796	796	796	796

Notes: This table shows the estimates of the full model. In columns (2) and (4), I allow non-compliant insurers having a different bargaining parameter. τ_1 is the estimate of the difference. In column (3) and (4), not-for-profit insurers could have different level of fixed costs. Standard errors are calculated by the parametric bootstrapping.

Table 1.7: Changes in price due to MLR regulation

	(1) Compliant	(2) Non-compliant	(3) Pooled
Price w/o MLR regulation			
Mean	2659.5	2952.56	2680.42
Std. Dev.	5152.09	6312.45	5243.65
Price w/ MLR regulation			
Mean	2659.31	3039.17	2686.42
Std. Dev.	5138.87	6828.58	5277.97
Change in price			
Mean	-0.20	86.62	6.00
Std. Dev.	39.53	561.54	156.31
N Observations	10,629	817	11,446

Notes: This table shows the changes in negotiated price due to the MLR regulation, assuming premium and demand unchanged.

Table 1.8: Statistics of the counterfactual scenarios

	Insurer 1	Insurer 2	Insurer 3
Scenario 1. No regulation, with bargaining			
Service price (in thousands)	2.73	2.72	2.71
Premium (in thousands)	2.61	2.61	2.6
Demand (in thousands)	17.61	17.54	17.39
Premium revenue (in millions)	45.95	45.72	45.23
Medical loss (in millions)	33.68	33.45	32.96
Scenario 2. With regulation, no bargaining			
Service price (in thousands)	2.73	2.72	2.71
Premium (in thousands)	2.39	2.38	2.37
Demand (in thousands)	18.06	18.05	18.03
Premium revenue (in millions)	43.19	43.03	42.72
Medical loss (in millions)	34.55	34.43	34.18
Scenario 3. With regulation, with bargaining			
Service price (in thousands)	4.29	4.27	4.23
Premium (in thousands)	3.76	3.74	3.7
Demand (in thousands)	16.33	16.4	16.56
Premium revenue (in millions)	61.33	61.3	61.3
Medical loss (in millions)	49.07	49.04	49.04
Scenario 4. No regulation, with bargaining & public option			
Service price (in thousands)	2.57	2.56	2.59
Premium (in thousands)	2.54	2.54	2.27
Demand (in thousands)	16.55	16.49	20.21
Premium revenue (in millions)	42.05	41.85	45.82
Medical loss (in millions)	29.78	29.58	36.66

Notes: This table characterizes the market in four scenarios. In the first scenario, there is no regulation and insurers and providers negotiate the service price. In the second scenario, service prices are fixed at the pre-regulation level and the regulation is effective. The settings of the third scenario is the same as in the ACA health exchange marketplaces. In the last scenario, the third insurance plan is a public option and regulated by the medical loss ratio regulation, while the other two private plans are not regulated.

Chapter 1 Appendices

A Conditions for $0 < \frac{1-\lambda}{1-\lambda\bar{R}} < 1$

In this section, I prove that $0 < \lambda < 1$ and $0 < \frac{1-\lambda}{1-\lambda\bar{R}} < 1$ are equivalent.

By definition, $0 < \bar{R} < 1$ and the Lagrangian multiplier $\lambda > 0$ when the constraint is binding. Therefore,

$$0 < \lambda < 1 \quad \Leftrightarrow \quad 0 < 1 - \lambda < 1 - \lambda\bar{R}.$$

It is not hard to prove that

$$0 < 1 - \lambda < 1 - \lambda\bar{R} \quad \Rightarrow \quad 0 < \frac{1-\lambda}{1-\lambda\bar{R}} < 1.$$

For the other direction,

$$0 < \frac{1-\lambda}{1-\lambda\bar{R}} < 1 \quad \Rightarrow \quad 0 < 1 - \lambda < 1 - \lambda\bar{R} \text{ or } 0 > 1 - \lambda > 1 - \lambda\bar{R}$$

Because $0 > 1 - \lambda > 1 - \lambda\bar{R} \Rightarrow \lambda < \lambda\bar{R}$, which contradict to $0 < \bar{R} < 1$ and $\lambda > 0$. Only the first case is possible, that is

$$0 < \frac{1-\lambda}{1-\lambda\bar{R}} < 1 \quad \Rightarrow \quad 0 < 1 - \lambda < 1 - \lambda\bar{R}.$$

Therefore, $0 < \lambda < 1$ is a sufficient and necessary condition for $0 < \frac{1-\lambda}{1-\lambda\bar{R}} < 1$.

B Price negotiation

- Multi-product case

When an insurer providers more than one products, she could strategically select premiums to maximize the overall profit. In such case, the Nash product function for a insurer providing J plans could be written as

$$f^{Nash} = \tau ln \left[\sum_j (\phi_j - p\kappa_j)D_j - C^F + \lambda(\sum_j p\kappa_j D_j - \bar{R} \sum_j \phi_j D_j) \right]$$

$$+(1-\tau)ln \left[(p - p^{MCD}) \sum_j Dj \right].$$

The F.O.C. w.r.t. p is

$$\frac{\partial f^{Nash}}{\partial p} = \frac{\tau}{\mathfrak{L}^I}(\frac{\partial \mathfrak{L}^I}{\partial p} + \sum_j \frac{\partial \mathfrak{L}^I}{\phi_j}\frac{\partial \phi}{\partial p}) + \frac{(1-\tau)}{(p-p^{MCD})\sum_j D_j}\left[\sum_j D_j + (p - p^{MCD}) \sum_k \frac{\partial D_j}{\partial \phi_k}\frac{\partial \phi_k}{\partial p} \right] = 0$$

Similar to the single-product case, I apply the envelope theorem and use the relation in the equation (1.4) to obtain the equation below

$$\frac{\partial f^{Nash}}{\partial p} = \frac{\tau}{\mathfrak{L}^I}(\frac{\partial \mathfrak{L}^I}{\partial p}) + \frac{(1-\tau)}{(p-p^{MCD})\sum_j D_j} \sum_j \left[D_j + (p - p^{MCD}) \sum_k \frac{\partial D_j}{\partial \phi_k}\frac{\partial \phi_k}{\partial p} \right] = 0$$

Then, I obtain the equation below,

$$\frac{-\tau(1-\lambda)\sum_j \kappa_j D_j}{\sum_j \left[(1-\lambda\bar{R})\phi_j - (1-\lambda)p\kappa_j \right] D_j - C^F} + \frac{(1-\tau)}{p - p^{MCD}} + (1-\tau)\frac{\sum_j \sum_k \frac{\partial D_j}{\partial \phi_k}\frac{\partial \phi_k}{\partial p}}{\sum_j D_j} = 0 \quad (15)$$

Rearrange the denominator of the first term to

$$[(1-\lambda\bar{R})\boldsymbol{\phi} - (1-\lambda)p\boldsymbol{\kappa}]^T\mathbf{D}$$

Rearrange the equation (1.4),

$$(1-\lambda)p\boldsymbol{\kappa} - (1-\lambda\bar{R})\boldsymbol{\phi} = (1-\lambda\bar{R})\mathbf{J}^{-1}\mathbf{D}$$

By plugging the above equations to equation (15),

$$\Rightarrow \quad \frac{\tau(1-\lambda)\sum_j \kappa_j D_j}{(1-\lambda\bar{R})[J^{-1}D]^T D + C^F} + \frac{(1-\tau)}{p - p^{MCD}} + (1-\tau)\frac{\sum_j \sum_k \frac{\partial D_j}{\partial \phi_k}\frac{\partial \phi_k}{\partial p}}{\sum_j D_j} = 0$$

$$\Rightarrow \quad (p - p^{MCD}) \left[\frac{\tau(1-\lambda)\sum_j \kappa_j D_j}{(1-\lambda\bar{R})[J^{-1}D]^T D + C^F} + (1-\tau)\frac{\sum_j \sum_k \frac{\partial D_j}{\partial \phi_k}\frac{\partial \phi_k}{\partial p}}{\sum_j D_j} \right] = -(1-\tau)$$

$$\Rightarrow \quad (p - p^{MCD}) \left[\frac{\tau(1-\lambda)\boldsymbol{\kappa}^T \mathbf{D}}{(1-\lambda\bar{R})[\mathbf{J}^{-1}\mathbf{D}]^T \mathbf{D} + C^F} + (1-\tau)\frac{\mathbf{1}^T \mathbf{J}\frac{\partial \boldsymbol{\phi}}{\partial p}}{\mathbf{1}^T \mathbf{D}} \right] = -(1-\tau) \quad (16)$$

where $\mathbf{1}$ is a $J \times 1$ vector of ones.

Calculation of ϕ'

One important variable in the expression of optimal price, equation(1.6) and equation (1.8), is $\frac{\partial \phi}{\partial p}$. This

- Single product case

Starting from the equation (1.4), in the single-product case,

$$(1-\lambda)p\kappa = (1-\lambda\bar{R})(\phi + \frac{D}{D'})$$

Then, the total derivative w.r.t. p when $\phi = \phi(p)$ yields

$$(1-\lambda)p\kappa = (1-\lambda\bar{R})\phi' \left[2 - \frac{DD''}{(D')^2} \right]$$

Rearrange to get the equation(1.7)

$$\phi' = \frac{(1-\lambda)\kappa}{(1-\lambda\bar{R})A}$$

where $A = 2 - \frac{DD''}{(D')^2}$.

- Multi-product case

When each insurers provide more than one plans, ϕ' could be derived as following.

For each plan $j \in \{1, ..., J\}$,

$$\frac{1-\lambda}{1-\lambda\bar{R}}\; p\sum_{k=1}^{J}\frac{\partial D_j}{\partial \phi_k}\kappa_k = \sum_{k=1}^{J}\frac{\partial D_j}{\partial \phi_k}\phi_k + D_j$$

As $\phi_k, \forall k \in \{1, ..., J\}$ is a function of p, by taking derivative w.r.t. p, I obtain,

$$\frac{1-\lambda}{1-\lambda\bar{R}}\sum_{k=1}^{J}\frac{\partial D_j}{\partial \phi_k}\kappa_k + \frac{1-\lambda}{1-\lambda\bar{R}}\; p\sum_{k=1}^{J}\kappa_k\sum_{l=1}^{J}\frac{\partial^2 D_j}{\partial \phi_k \partial \phi_l}\frac{d\phi_l}{dp}$$

$$= \sum_{k=1}^{J}\phi_k\sum_{l=1}^{J}\frac{\partial^2 D_j}{\partial \phi_k \partial \phi_l}\frac{d\phi_l}{dp} + \sum_{k=1}^{J}\frac{\partial D_j}{\partial \phi_k}\frac{d\phi_k}{dp} + \sum_{k=1}^{J}\frac{\partial D_j}{\partial \phi_k}\frac{d\phi_k}{dp}$$

in matrix form

$$\Rightarrow \quad \frac{1-\lambda}{1-\lambda\bar{R}}\mathbf{J}_{j:}\boldsymbol{\kappa} + \frac{1-\lambda}{1-\lambda\bar{R}}\; p\boldsymbol{\kappa}^T\mathbf{H^j}\frac{\partial \boldsymbol{\phi}}{\partial p} = \boldsymbol{\phi}^T\mathbf{H}^j\frac{\partial \boldsymbol{\phi}}{\partial p} + 2\mathbf{J}_{j:}\frac{\partial \boldsymbol{\phi}}{\partial p}$$

where $\mathbf{J}_{j:}$ is the j^{th} row of the Jacobian matrix of demand function, $\mathbf{H^j}$ is the Hessian matrix of plan j's demand function.

$$\Rightarrow \quad \frac{1-\lambda}{1-\lambda\bar{R}}\mathbf{J}_{j:}\boldsymbol{\kappa} + (\frac{1-\lambda}{1-\lambda\bar{R}}\; p\boldsymbol{\kappa}^T - \boldsymbol{\phi}^T)\mathbf{H^j}\frac{\partial \boldsymbol{\phi}}{\partial p} = 2\mathbf{J}_{j:}\frac{\partial \boldsymbol{\phi}}{\partial p}$$

Plug in the matrix form of equation (1.4),

$$\frac{1-\lambda}{1-\lambda\bar{R}}p\boldsymbol{\kappa}^T - \boldsymbol{\phi}^T = [\mathbf{J}^{-1}\mathbf{D}]^T,$$

the equation becomes

$$\frac{1-\lambda}{1-\lambda\bar{R}}\mathbf{J}_{j:}\boldsymbol{\kappa} + [\mathbf{J}^{-1}\mathbf{D}]^T\mathbf{H^j}\frac{\partial \boldsymbol{\phi}}{\partial p} = 2\mathbf{J}_{j:}\frac{\partial \boldsymbol{\phi}}{\partial p}$$

After a rearrangement, the equation used to calculate $\frac{\partial \phi_j}{\partial p}, \forall j$ is

$$[\mathbf{J}^{-1}\mathbf{D}]^T\mathbf{H^j}\frac{\partial \boldsymbol{\phi}}{\partial p} = \mathbf{J}_{j:}(2\frac{\partial \boldsymbol{\phi}}{\partial p} - \frac{1-\lambda}{1-\lambda\bar{R}}\boldsymbol{\kappa}),\ \forall j \qquad (17)$$

$$([\mathbf{J}^{-1}\mathbf{D}]^T\mathbf{H}^{\mathbf{j}} - 2\mathbf{J}_{j:})\frac{\partial\boldsymbol{\phi}}{\partial p} = -\frac{1-\lambda}{1-\lambda\bar{R}}\mathbf{J}_{j:}\boldsymbol{\kappa},\ \forall j$$

Finally, the equation (1.9)

$$\left(\begin{bmatrix}[\mathbf{J}^{-1}\mathbf{D}]^T H^1\\ \vdots \\ [\mathbf{J}^{-1}\mathbf{D}]^T H^J\end{bmatrix} - 2\mathbf{J}\right)\frac{\partial\boldsymbol{\phi}}{\partial p} = -\frac{1-\lambda}{1-\lambda\bar{R}}\mathbf{J}\boldsymbol{\kappa}$$

Calculation of derivatives of demand

The next step is to get D'' from demand estimation. Based on the market share function specified by equation(1.2), the demand of insurance plan j is

$$D_j \equiv M_t s_{jt} = M_t \int \frac{e^{\delta_j+\mu_{ij}(X_j,\phi_j,D_i,\nu_i;\Pi,\Sigma)}}{1+\sum_k e^{\delta_k+\mu_{ik}(X_k,p_k,D_i,\nu_i;\Pi,\Sigma)}}\varphi(\nu_i)d\nu_i$$

where M_t is the market size of market t. Then, by the Leibniz integral rule,

$$\frac{\partial D_j}{\partial\phi_j} = M\int(\alpha+\pi^\alpha d_i^\alpha+\sigma^\alpha\nu_i^\alpha)s_{ij}(1-s_{ij})\varphi(\nu_i)d\nu_i$$

$$\frac{\partial D_j}{\partial\phi_k} = -M\int(\alpha+\pi^\alpha d_i^\alpha+\sigma^\alpha\nu_i^\alpha)s_{ij}s_{ik}\varphi(\nu_i)d\nu_i$$

and

$$\frac{\partial^2 D_j}{\partial\phi_j^2} = M\int(\alpha+\pi^\alpha d_i^\alpha+\sigma^\alpha\nu_i^\alpha)^2(1-2s_{ij})s_{ij}(1-s_{ij})\varphi(\nu_i)d\nu_i$$

$$\frac{\partial^2 D_j}{\partial\phi_j\partial\phi_k} = -M\int(\alpha+\pi^\alpha d_i^\alpha+\sigma^\alpha\nu_i^\alpha)^2(1-2s_{ij})s_{ij}s_{ik}\varphi(\nu_i)d\nu_i$$

$$\frac{\partial^2 D_j}{\partial\phi_k\partial\phi_l} = M\int 2(\alpha+\pi^\alpha d_i^\alpha+\sigma^\alpha\nu_i^\alpha)^2 s_{ij}s_{ik}s_{il}\varphi(\nu_i)d\nu_i$$

The above terms could be approximated numerically by simulation, as in the demand estimation. That is changing integral to summation–sum up all the simulated individuals in the market. Note that the size of Jacobian and Hessian matrices varies across markets.

$$\mathbf{J}_{jj} = \frac{\partial D_j}{\partial \phi_j} \approx M \frac{1}{N} \sum_{i=1}^{N} (\alpha + \pi^{\alpha} d_i^{\alpha} + \sigma^{\alpha} \nu_i^{\alpha}) s_{ij} (1 - s_{ij}) \tag{18}$$

$$\mathbf{J}_{jk} = \frac{\partial D_j}{\partial \phi_k} \approx -M \frac{1}{N} \sum_{i=1}^{N} (\alpha + \pi^{\alpha} d_i^{\alpha} + \sigma^{\alpha} \nu_i^{\alpha}) s_{ij} s_{ik} \tag{19}$$

and

$$\mathbf{H^j}_{jj} = \frac{\partial^2 D_j}{\partial \phi_j^2} \approx M \frac{1}{N} \sum_{i=1}^{N} (\alpha + \pi^{\alpha} d_i^{\alpha} + \sigma^{\alpha} \nu_i^{\alpha})^2 (1 - 2s_{ij}) s_{ij} (1 - s_{ij}) \tag{20}$$

$$\mathbf{H^j}_{jk} = \frac{\partial^2 D_j}{\partial \phi_j \partial \phi_k} \approx -M \frac{1}{N} \sum_{i=1}^{N} (\alpha + \pi^{\alpha} d_i^{\alpha} + \sigma^{\alpha} \nu_i^{\alpha})^2 (1 - 2s_{ij}) s_{ij} s_{ik} \tag{21}$$

$$\mathbf{H^j}_{kl} = \frac{\partial^2 D_j}{\partial \phi_k \partial \phi_l} \approx M \frac{1}{N} \sum_{i=1}^{N} (\alpha + \pi^{\alpha} d_i^{\alpha} + \sigma^{\alpha} \nu_i^{\alpha})^2 s_{ij} s_{ik} s_{il} \tag{22}$$

By plugging equation(18) to (22) into the equation (17), I obtain $\frac{\partial \phi}{\partial p}$ as functions of λ–a parameter to estimate.

In the demand estimation, because I use log of premium instead of premium, I need to divide the derivatives by the premium when I implement this approximation.

Elimination of p

From previous section, we know that insurers choose premiums to maximize their profit under the MLR regulation, according the equation (1.4).

$$p\boldsymbol{\kappa} = \frac{1 - \lambda \bar{R}}{1 - \lambda} (\boldsymbol{\phi} + \mathbf{J}^{-1} \mathbf{D}) \tag{23}$$

Following the Kuhn-Tucker condition, $\lambda > 0$ for insurers triggered the rebate and $\lambda = 0$ otherwise.

Then from the above equation, we could obtain the expression of $\frac{\partial \phi}{\partial p}$, i.e. the equation (1.9)

$$\left(\begin{bmatrix} [\mathbf{J}^{-1}\mathbf{D}]^T H^1 \\ \vdots \\ [\mathbf{J}^{-1}\mathbf{D}]^T H^J \end{bmatrix} - 2\mathbf{J} \right) \frac{\partial \boldsymbol{\phi}}{\partial p} = -\frac{1-\lambda}{1-\lambda \bar{R}} \mathbf{J} \boldsymbol{\kappa}$$

To simplify notation, denote

$$\mathbf{A} = \left(\begin{bmatrix} [\mathbf{J}^{-1}\mathbf{D}]^T H^1 \\ \vdots \\ [\mathbf{J}^{-1}\mathbf{D}]^T H^J \end{bmatrix} - 2\mathbf{J} \right)$$

Then the expression of $\frac{\partial \boldsymbol{\phi}}{\partial p}$ is more explicit,

$$\frac{\partial \boldsymbol{\phi}}{\partial p} = -\frac{1-\lambda}{1-\lambda \bar{R}} \mathbf{A}^{-1} \mathbf{J} \boldsymbol{\kappa} \tag{24}$$

From the bargaining model, the service price p is selected according to the equation(16), one per each market-insurer. After plugged the equation(23) into the equation(16),

$$(p - p^{MCD}) \left[\frac{\tau(1-\lambda)\boldsymbol{\kappa}^T \mathbf{D}}{(1-\lambda \bar{R})[\mathbf{J}^{-1}\mathbf{D}]^T \mathbf{D} + C^F} + (1-\tau) \frac{1-\lambda}{1-\lambda \bar{R}} \frac{\mathbf{1}^T \mathbf{J} \mathbf{A}^{-1} \mathbf{J} \boldsymbol{\kappa}}{\mathbf{1}^T \mathbf{D}} \right] = -(1-\tau)$$

Rearrange

$$\frac{\tau(1-\lambda)(p\boldsymbol{\kappa})^T \mathbf{D}}{(1-\lambda \bar{R})[\mathbf{J}^{-1}\mathbf{D}]^T \mathbf{D} + C^F} + (1-\tau) \frac{1-\lambda}{1-\lambda \bar{R}} \frac{\mathbf{1}^T \mathbf{J} \mathbf{A}^{-1} \mathbf{J} (p\boldsymbol{\kappa})}{\mathbf{1}^T \mathbf{D}}$$

$$-p^{MCD} \left[\frac{\tau(1-\lambda)\boldsymbol{\kappa}^T \mathbf{D}}{(1-\lambda \bar{R})[\mathbf{J}^{-1}\mathbf{D}]^T \mathbf{D} + C^F} + (1-\tau) \frac{1-\lambda}{1-\lambda \bar{R}} \frac{\mathbf{1}^T \mathbf{J} \mathbf{A}^{-1} \mathbf{J} \boldsymbol{\kappa}}{\mathbf{1}^T \mathbf{D}} \right] = -(1-\tau)$$

Then plug the equation(23) into the equation

$$\frac{\tau(1-\lambda \bar{R})(\boldsymbol{\phi} + \mathbf{J}^{-1}\mathbf{D})^T \mathbf{D}}{(1-\lambda \bar{R})[\mathbf{J}^{-1}\mathbf{D}]^T \mathbf{D} + C^F} + (1-\tau) \frac{\mathbf{1}^T \mathbf{J} \mathbf{A}^{-1} \mathbf{J} (\boldsymbol{\phi} + \mathbf{J}^{-1}\mathbf{D})}{\mathbf{1}^T \mathbf{D}}$$

$$-p^{MCD} \left[\frac{\tau(1-\lambda)\boldsymbol{\kappa}^T \mathbf{D}}{(1-\lambda \bar{R})[\mathbf{J}^{-1}\mathbf{D}]^T \mathbf{D} + C^F} + (1-\tau) \frac{1-\lambda}{1-\lambda \bar{R}} \frac{\mathbf{1}^T \mathbf{J} \mathbf{A}^{-1} \mathbf{J} \boldsymbol{\kappa}}{\mathbf{1}^T \mathbf{D}} \right] = -(1-\tau)$$

Finally,

$$\frac{\tau(1-\lambda\bar{R})(\boldsymbol{\phi}+\mathbf{J}^{-1}\mathbf{D})^T\mathbf{D}}{(1-\lambda\bar{R})[\mathbf{J}^{-1}\mathbf{D}]^T\mathbf{D}+C^F}-\frac{\tau(1-\lambda)p^{MCD}\boldsymbol{\kappa}^T\mathbf{D}}{(1-\lambda\bar{R})[\mathbf{J}^{-1}\mathbf{D}]^T\mathbf{D}+C^F}-\frac{(1-\tau)(1-\lambda)}{1-\lambda\bar{R}}p^{MCD}\frac{\mathbf{1}^T\mathbf{J}\mathbf{A}^{-1}\mathbf{J}\boldsymbol{\kappa}}{\mathbf{1}^T\mathbf{D}}$$

$$=(1-\tau)(\frac{\mathbf{1}^T\mathbf{J}\mathbf{A}^{-1}\mathbf{J}(\boldsymbol{\phi}+\mathbf{J}^{-1}\mathbf{D})}{\mathbf{1}^T\mathbf{D}}-1)$$

C Conditions for increasing best response function $\phi(p)$

In this section, I am going to discuss the conditions for $\phi(p)$ being increasing in p.

D Optimal price in different scenarios

Price-taker insurers

In this scenario, health care providers set the price and insurers are simply price-takers. Therefore, the price p is set by

$$\max_p \Pi^H = (p-c_h)\theta D$$

The F.O.C. w.r.t. p is

$$\frac{\partial \Pi_H}{\partial p} = \theta D + (p-c_h)\theta D'\phi' = 0$$

Then the optimal price is

$$p^* = c_h - \frac{D}{D'\phi'} \tag{25}$$

Bargaining without regulation

This time, health care providers and insurers bargaining over price in a context without any constraint. The optimization problem in this scenario is (log of the Nash product)

$$\max_p f^{Nash} = \tau ln[(\phi - p\kappa\theta)D - C^F] + (1-\tau)ln[(p-c_h)\theta D].$$

The F.O.C. w.r.t. p is

$$\frac{\partial f^{Nash}}{\partial p} = \tau\frac{(\phi' - \kappa\theta)D + (\phi - p\kappa\theta)D'\phi'}{(\phi - p\kappa\theta)D - C^F} + (1-\tau)\frac{\theta D + (p - c_h)\theta D'\phi'}{(p - c_h)\theta D}$$

By envelope theorem, $D + (\phi - p\kappa\theta)D' = 0$,

$$\frac{\partial F}{\partial p} = \tau\frac{-\kappa\theta D}{(\phi - p\kappa\theta)D - C^F} + (1-\tau)\frac{\theta D + (p - c_h)\theta D'\phi'}{(p - c_h)\theta D} = 0$$

$$\Rightarrow \tau\frac{\kappa\theta D}{\frac{D^2}{D'} + C^F} + \frac{1-\tau}{p - c_h} + \frac{(1-\tau)D'\phi'}{D} = 0$$

$$\Rightarrow \frac{\tau}{1-\tau}\frac{\kappa\theta D}{\frac{D^2}{D'} + C^F} + \frac{1}{p - c_h} + \frac{D'\phi'}{D} = 0$$

$$\Rightarrow (p - c_h)\left[\frac{\tau}{1-\tau}\kappa\theta D^2 + D'\phi'(\frac{D^2}{D'} + C^F)\right] = -D(\frac{D^2}{D'} + C^F)$$

Therefore, the optimal price in this scenario is

$$p^{**} = c_h - \frac{D(\frac{D}{D'} + \frac{C^F}{D})}{\frac{\tau}{1-\tau}\kappa\theta D + \phi'(D + \frac{D'C^F}{D})}$$

$$= c_h - \frac{D}{D'\phi' + \frac{\frac{\tau}{1-\tau}\kappa\theta D}{\frac{D}{D'} + \frac{C^F}{D}}} \tag{26}$$

Compare to p^*, the second term in the denominator is the extra term and its impact on p^{**} depends on the bargaining power. When $\tau = 0$–insurers have no power at all, it degenerates to price-taker scenario. This term could be rearranged to

$$\frac{\frac{\tau}{1-\tau}\kappa\theta D}{\frac{D}{D'} + \frac{C^F}{D}} = \frac{\tau}{1-\tau}\frac{\kappa\theta D}{\phi(-1/\eta_\phi + C^F/(\phi D))}$$

where $\eta_\phi = -\frac{D'\phi}{D}$. Then the sign of this red term depends only on the tradeoff between premium elasticity and the fraction of fixed cost over premium revenue.

Bargaining with MLR regulation

This is the scenario in the main text, to make things more comparable, I rearrange the equation (1.6) and get

$$p^{***} = c_h - \frac{(1-\lambda\bar{R})\frac{D^2}{D'} + C^F}{\frac{\tau}{1-\tau}(1-\lambda)\kappa\theta D + (1-\lambda\bar{R})D\phi' + \frac{D'\phi'}{D}C^F}$$

$$= c_h - \frac{D\left[(1-\lambda\bar{R})\frac{D}{D'} + \frac{C^F}{D}\right]}{\frac{\tau}{1-\tau}(1-\lambda)\kappa\theta D + D'\phi'\left[(1-\lambda\bar{R})\frac{D}{D'} + \frac{C^F}{D}\right]}$$

$$= c_h - \frac{D}{D'\phi' + \frac{\frac{\tau}{1-\tau}(1-\lambda)\kappa\theta D}{(1-\lambda\bar{R})\frac{D}{D'} + \frac{C^F}{D}}} \tag{27}$$

Again the second term in the denominator will equal to zero if insurers do not have any bargaining power. Moreover, when the constraint is not binding, i.e. $\lambda = 0$, then the term will degenerate to the one in the previous subsection.

The table below summarizes the optimal prices in all six possible scenarios.

Optimal price

	$\nexists$ regulation	$\exists$ regulation
Price-maker ($\tau = 1$)	$p_1 = c_h$	$p_2 = c_h$
Price-taker ($\tau = 0$)	$p_3 = c_h - \frac{D}{D'\phi'}$	$p_4 = c_h - \frac{D}{D'\phi'}$
Bargain	$p_5 = c_h - \frac{D}{D'\phi' + \frac{\frac{\tau}{1-\tau}\kappa\theta D}{\frac{D}{D'} + \frac{C^F}{D}}}$	$p_6 = c_h - \frac{D}{D'\phi' + \frac{\frac{\tau}{1-\tau}(1-\lambda)\kappa\theta D}{(1-\lambda\bar{R})\frac{D}{D'} + \frac{C^F}{D}}}$
Function of ϕ	$\phi = p\kappa\theta - \frac{D}{D'}$	$\phi = \frac{1-\lambda}{1-\lambda\bar{R}}p\kappa\theta - \frac{D}{D'}$

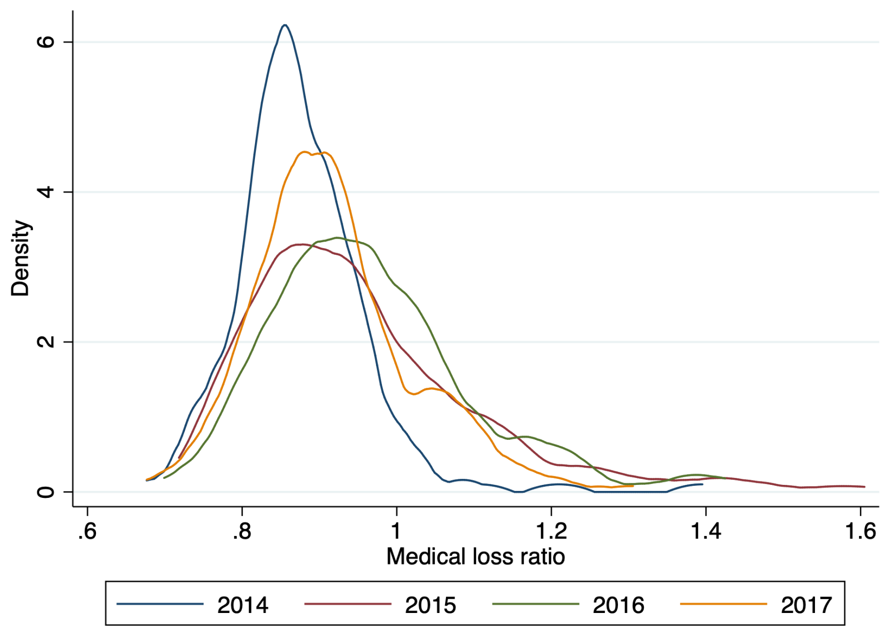
E Additional Tables and Figures

Table A1: Effect of MLR regulation on marginal cost

	(1)	(2)	(3)	(4)	(5)
λ	0.132***	0.125***	0.0413***	0.0298***	0.0288**
	(0.00878)	(0.00900)	(0.00895)	(0.00898)	(0.00886)
ln(deductible)	0.00851	0.00715	0.00454	0.00884*	0.00968*
	(0.00473)	(0.00477)	(0.00449)	(0.00441)	(0.00442)
Copayment	-0.000594	-0.00117	-0.00237***	-0.00179**	
	(0.000601)	(0.000603)	(0.000552)	(0.000549)	
Coinsurance rate	-0.0482	-0.0908	0.136*	0.128*	
	(0.0644)	(0.0624)	(0.0627)	(0.0616)	
Having copayment	-0.0174	-0.0151	0.0389*	0.0184	
	(0.0200)	(0.0198)	(0.0191)	(0.0190)	
Having coinsurance	0.0770***	0.0676***	-0.0223	0.0142	
	(0.0197)	(0.0194)	(0.0176)	(0.0178)	
ln(OOP max)					0.0453*
					(0.0219)
Constant	7.257***	7.344***	7.479***	7.446***	7.036***
	(0.0516)	(0.0538)	(0.0845)	(0.0755)	(0.206)
Year FE		Yes	Yes	Yes	Yes
State FE			Yes	Yes	Yes
Firm characteristics				Yes	Yes
N	10859	10859	10859	10859	10859
R^2	0.022	0.035	0.278	0.302	0.300

Notes: The compliance status is identified by the credibility-adjusted MLR.

Figure A1: Distribution of Medical Loss Ratio in the Marketplace, by year



Notes: This figure shows the Kernel density of MLR of insurers in the Marketplace from 2014-2017. The unit of observation is insurer-year.

Chapter 2

The Impact of Organizational Boundaries on Healthcare Coordination and Utilization

with Leila Agha and Keith Marzilli Ericson

2.1 Introduction

Transaction costs and imperfect information can make it difficult to coordinate production across firm boundaries (Coase, 1937; Williamson, 1985). The determinants of firm boundaries have been the subject of substantial theoretical and empirical investigation, particularly in the literature on vertical integration (Lafontaine and Slade, 2007). Yet, we know less about how firm boundaries affect the firm performance (Mullainathan and Scharfstein, 2001). and empirical studies from different industries find mixed results.[1]

In healthcare, the challenges of cross-firm coordination are particularly salient; patient care is often produced with the input of many healthcare providers working in separate organizations. Geographically and over time, there is substantial variation in the organizational structures those providers operate in. An increasing fraction of US physicians is employed by large practices or hospitals (Welch et al., 2013), which may mitigate these coordination challenges. Integrated care organizations such as the Mayo Clinic, Intermountain Healthcare, and Kaiser Permanente are often held up as

[1] For example, see Seru (2014); Pierce (2012); Stroebel (2016); Forbes and Lederman (2010); Forman and Gron (2011).

models of clinical efficiency and coordinated care (Enthoven, 2009). Yet empirical evidence on how organizational boundaries affect healthcare productivity is limited.

In this paper, we investigate how organizational boundaries affect healthcare utilization. Existing evidence has shown that when coordination of care is more difficult, healthcare utilization tends to be higher. These coordination challenges can emerge when healthcare for an individual patient is spread across many individual providers (Agha, Frandsen and Rebitzer, 2019; Frandsen et al., 2015), or when provider teams have fewer repeat interactions (Agha et al., 2018; Kim, Song and Valentine, 2020; Chen, 2020). Cebul et al. (2008) argue that fragmentation across *organizations* may also be an important source of healthcare inefficiency. Organizational boundaries can affect the coordination costs; e.g., healthcare firms often restrict information transmission to external providers by limiting transfer across electronic medical record systems. Providers may invest in firm-specific relationships and infrastructure that improve productivity (Huckman and Pisano, 2006). Finally, organizational fragmentation can affect incentives for clinical process improvement and other efficiency-enhancing investments due to common agency problems and spillovers that prevent firms from reaping the full benefit of their investments (Frandsen, Powell and Rebitzer, 2019).

We introduce the concept of "organizational concentration," which measures the distribution of a patient's outpatient visits across organizations. A patient's healthcare has maximal organizational concentration if all of their outpatient care is billed by the same organization. This construct builds on earlier work studying provider concentration (Pollack et al., 2016; Agha, Frandsen and Rebitzer, 2019). Organizational concentration describes the realized experience of a given patient, and so is distinct from market concentration measures used in antitrust research, which instead measure provider market power for pricing. Patients who receive all their healthcare from one firm will have high organizational concentration even if there are many firms in the market. Conversely, a patient may have low organizational concentration in a highly concentrated market if they receive healthcare from many different specialty practices, even if each practice has a monopoly in that specialty.

To our knowledge, we are the first paper to measure organizational concentration systematically, so we begin with a detailed descriptive analysis. Using a 20% sample of insurance claims for Medicare fee-for-service enrollees from 2007-2016, we construct a measure of each patient's experienced organizational concentration. There is substantial heterogeneity across regions in organizational concentration, even conditional on the spread of patient care across providers. Studying patients who move across regions, we find that moving to a location with a higher level of organizational concentration is associated with lower healthcare utilization. While these results suggest that organizational concentration leads to lower healthcare spending, they should be interpreted with caution because other attributes of regional practice style and place effects may be correlated with the level of organizational concentration.

To isolate variation in organizational concentration from other aspects of the local practice environment, we exploit quasi-experimental variation in patient assignment to physicians generated by physician exits. We examine the experiences of patients whose primary care provider (PCP) exits the local market, either due to a move or retirement, following recent work by Fadlon and Van Parys (2020) and Kwok (2019). Since patients may endogenously sort to new PCPs on the basis of changes in their health status, we use an instrumental variable strategy that leverages mean reversion to predict the change in a patient's assigned PCP's average organizational concentration, adapting the approach used by Laird and Nielsen (2017) and Abaluck et al. (2020). When PCPs with low organizational concentration exit the market, their patients switch to more typical PCPs with higher average concentration and subsequently experience lower healthcare utilization. Using this variation, we estimate that patients who switch to a PCP with 1 SD higher organizational concentration have 10% lower healthcare utilization in our preferred, most controlled specification. This finding is robust to controlling for the number and types of providers that the patient visits.

Our results indicate that organizational boundaries contribute additional frictions that lower the efficiency of healthcare provision, and this pattern does not simply reflect the challenges of spreading care across multiple providers. Although we cannot fully isolate a PCP's tendency for organizational concentration from every other possible dimension of PCP practice style, our estimated effect remains large in specifications that control for the spread of patient care across providers, the size of the PCP's practice group, as well as other PCP characteristics (residency training, experience, gender). To the extent that observable variables are informative about selection on unobservables, this supports the claim that organizational concentration is an important independent contributor to spending variation (Oster, 2019).

Finally, we investigate how organizational concentration influences quality of care. We use several measures related to distinct dimensions of healthcare quality, spanning gaps in primary care, appropriate management of chronic conditions, and repeated testing. We find no strong evidence that changes in PCP organizational concentration predict changes in inpatient or emergency department visits, or labs. However, for patients with a chronic condition (diabetes), switching to a PCP with higher levels of organizational concentration leads to better adherence to recommended care guidelines. This finding from diabetes care provides suggestive evidence that greater organizational concentration may facilitate improved management of chronic conditions. We also find suggestive evidence that spending higher organizational concentration reduces claims for diagnostic imaging.

High levels of organizational concentration arise when most of the providers a patient consults are integrated within the same firm. Our research is motivated by earlier work finding the effects of firm integration on productive efficiency are theoretically ambiguous. Bringing transactions into the same firm could improve communication (Arrow, 1975) and reduce contracting barriers (Hart and Moore, 1990; Hart and Holmstrom, 2010). On the other hand, integration may also lead resources within the firm to be allocated less efficiently (Alonso, Dessein and Matouschek, 2008; Friebel and Raith, 2010). Moreover, integration may improve coordination in stable environments but lead to worse adaptation to change (Dessein, 2014).

Empirical evidence from other industries on how integration affects firm performance has found mixed results. Mullainathan and Scharfstein (2001), Seru (2014), and Pierce (2012) document downsides to integration including less efficient capacity management, lower innovation, and insufficient knowledge sharing. By contrast, Stroebel (2016), Forbes and Lederman (2010), Forman and Gron (2011) find benefits of firm integration including superior information, better performance, and faster technology adoption. Atalay, Hortaçsu and Syverson (2014) argue that integration facilitates the efficient intrafirm transfer of intangible inputs, such as high quality managerial oversight and planning. We build on this literature by studying how firm boundaries affect health care delivery, a setting where the potential benefits of improved coordination, knowledge-sharing, and management are high, and rich insurance claims data allows us to track the production process.

Within healthcare, there is limited evidence on how the integration of healthcare providers affects care delivery. Although large consolidated practice groups argue they can deliver lower cost, higher quality healthcare through improved coordination, leveraging returns to specialization, and facilitating fixed cost investments, empirical evidence of these benefits is limited (Cutler and Scott Morton, 2013). Recent work suggests that hospital mergers and acquisitions of physician practices do not spur improvements in clinical quality or health outcomes (Beaulieu et al., 2020; Koch, Wendling and Wilson, 2018).[2] We build on this research by studying changes in the extent to which individual patient care crosses firm boundaries, rather than focusing on short-run effects of mergers and acquisitions. Care coordination depends on the ease of communication across multiple providers who treat the same patient, but mergers may simply bring competing providers—who rarely would have treated the same patient—into the same firm. Further, the process of organizational transformation is often slow. Because this paper does not focus on short-run effects of mergers, the effects we study may reflect longer-run operational changes associated with integration.

This paper is also related to a growing literature investigating differences in practice patterns across individual physicians. Across a variety of care contexts, individual physician quality and practice style have important effects on care outcomes.[3] Recent work by Kwok (2019) and Fadlon and Van Parys (2020) documents that primary care physicians in particular have substantial influence on patients' healthcare spending. We build on this insight by investigating one important dimension of PCP practice environment and referral patterns, i.e. the PCP's tendency to concentrate patient care within organizations.

The paper is organized as follows. Section 2.2 introduces our measure of organizational concentration. Section 2.3 describes our data and sample selection. Section 2.4 reports descriptive statistics on regional variation in organizational concentration and uses movers between regions to explore how regional variation in organizational con-

[2] These acquisitions may even raise healthcare spending, as physicians shift the site of care from doctors' offices to hospital outpatient settings (Koch, Wendling and Wilson, 2017) and exploit reimbursement rules that allow hospital-owned physician practices to charge additional facility fees (Capps, Dranove and Ody, 2018).

[3] For example, see Gowrisankaran, Joiner and Léger (2017); Molitor (2018); Chan, Gentzkow and Yu (2019); Currie and MacLeod (2017); Currie, MacLeod and Van Parys (2016); Sahni et al. (2016).

centration may contribute to regional variation in healthcare utilization. Section 2.5 lays out our main empirical strategy exploiting PCP exits to explore the impact of organizational concentration. Section 2.6 presents the results on how healthcare utilization and quality outcomes change when a patient switches to a PCP with a different level of organizational concentration. Section 2.7 concludes.

2.2 Defining Organizational Concentration

In this project, we study the coordination frictions that arise when healthcare is spread across organizational boundaries. To do so, we define *organizational concentration*, adapting a concentration index that has been used in prior literature to measure the spread of patient care across providers.[4] Specifically, we use a Herfindahl-Hirschman Index (HHI) that calculates how outpatient healthcare received by a patient is spread across organizations. We measure organizational concentration using outpatient care, following previous literature defining continuity of outpatient care across individual physicians (Nyweide and Bynum, 2017; Nyweide et al., 2013). This allows us to consider the impact of outpatient organizational concentration on the likelihood that a patient requires an emergency department visit or hospitalization.

We calculate patient i's share of outpatient visits at each organization j, in a year t. Organizational concentration is then defined as the sum of squared shares across all the organizations:

$$OrgConc_{it} = \sum_j share_{ijt}^2. \tag{2.1}$$

In general, organizational concentration is higher when a patient visits fewer organizations. When a patient's outpatient visits are uniformly distributed across N organizations, this measure is simply $1/N$. When a patient receives all the visits

[4] Pollack et al. (2016) provides an overview and comparison of commonly used measures of care continuity.

from one organization, this concentration measure will be 1. Lower values correspond to patient care that is spread more diffusely across organizations.

For some analyses, we aggregate organizational concentration up to at the hospital referral region (HRR) level. In our primary empirical strategy, we aggregate organizational concentration up to the PCP level.

Defining provider concentration

To distinguish our findings from prior analyses, we study variation in organizational concentration conditional on provider concentration: the spread of patient healthcare across providers. Following Agha, Frandsen and Rebitzer (2019), we construct a measure of provider care concentration where the $share_{ipt}$ measures the share of patient i's outpatient visits in year t for each provider p:

$$ProviderConc_{it} = \sum_{p} share_{ipt}^2. \tag{2.2}$$

This measure will capture the challenges of coordinating healthcare across many providers, thus allowing us to distinguish them from the frictions that are specific to crossing organizational boundaries.

2.3 Data and Sample Construction

2.3.1 Patient sample selection

Our primary source of data is a 20% sample of Medicare Fee-For-Service (FFS) Part A and Part B claims data from 2007-2016. The 10-year panel data allows us to observe both patient moves and PCP exits. We use the Carrier, Inpatient, and Outpatient claims files to measure care utilization and spending.[5] Patient demographics (age, sex,

[5]The Inpatient file contains institutional inpatient claims, and the Outpatient file contains claims from institutional outpatient providers such as hospital outpatient departments or community mental health centers. The Carrier file contains non-institutional claims billed by individual providers such

zip code) and chronic conditions are extracted from the Master Beneficiary Summary file with the Chronic Condition segment. In the remainder of this section, we describe the sample restrictions implemented to construct our main analytic samples.

Initial sample restrictions

We restrict our sample to Medicare beneficiaries who are 66–99 years old (inclusive) and continuously enrolled in Medicare FFS. After these restrictions, our data covers 9,356,144 beneficiaries. Our organizational concentration measure is defined based on outpatient site of care visits billed in the Carrier claims files, so we drop 223,822 beneficiaries who did not have any visits of this type. This comprises our Broad Sample. From this broad sample, we define two separate analytic samples for different purposes. First, we define a "Patient Mover Sample" for a descriptive analysis studying regional variation in organizational concentration. Second, we define a "PCP Exit Sample" for our primary analysis studying the relationship between PCPs' organizational concentration and patient care utilization. We describe each of these samples below.

Patient Mover Sample

We construct a Patient Mover Sample for our initial descriptive analysis. Sample restrictions defined here follow the construction process outlined in Agha, Frandsen and Rebitzer (2019). We assigned each patient to a hospital referral region (HRR) on an annual basis, using the zip code reported in the Beneficiary Summary File. Further, we require that the patient received at least 75% of billed claims within that HRR; we drop beneficiaries who do not meet this requirement. To be included as a mover, the patient's HRR must have changed once (and only once) in our 10-year period. Further, the beneficiary must be continuously in the sample from two years before

as physicians, and these claims can result from services provided at either outpatient or inpatient settings.

their move to two years after. Our sample includes all moving patients who meet these criteria as well as a 25% random sample of non-movers (whose HRR never changed during this time period); non-movers contribute toward covariate identification. The final Patient Mover Sample includes 25,592 mover beneficiaries and 1,364,198 non-mover beneficiaries.

PCP Exit Sample

Next, we construct our PCP Exit Sample for our main analysis. This analysis focuses on beneficiaries who change their attributed PCP due to the original PCP's relocation or retirement. We use provider taxonomies to distinguish primary care specialties from other types of providers. The provider taxonomy codes used for this categorization are reported in Table A1 and include codes for internal medicine, family medicine, pediatrics, geriatrics, and general practice. Provider taxonomy codes are the primary specialty code from the National Plan and Provider Enumeration System (NPPES), which is linked to our sample by providers' National Provider Identifier (NPI). We attribute each patient to their plurality PCP in each year, defined as the provider who bills a plurality of the patient's Evaluation & Management (E&M) visits that year; ties are broken randomly. We exclude patients who have no E&M visits and thus cannot be matched to a provider, as well as patients whose plurality provider does not report a primary care specialty. If a patient cannot be matched to a PCP according to this algorithm, they will be excluded from the PCP Exit Sample.

We limit this analysis to patients whose initial attributed PCP either moved (i.e. relocated once to a different HRR) or retired (i.e. bills no further Medicare claims). We also exclude patients who move across HRRs themselves or who have ever changed their PCP in our sample period prior to the exit of their assigned original PCPs. The PCP Exit Sample includes 62,924 beneficiaries and 335,868 beneficiary-year observations. These patients are initially attributed to one of 4365 relocating PCPs or

11,437 retiring PCPs; including both the exiting PCPs and the destination PCPs, this sample covers 52,981 PCPs.

2.3.2 Measuring organizational concentration

Measuring Organizations

The next step is to construct our measure of organizational concentration. We begin by identifying provider organizations delivering outpatient care to each patient. We limit to provider services billed in the Carrier claims file and provided in an outpatient setting. The outpatient setting is identified using the place of service code listed on the Carrier file claims; a complete list of places of service codes is in Appendix Table A2. We then define a visit by aggregating claims to a unique provider-date pair. About 85% of visits measured in the Carrier claim file are classified as outpatient visits.

We use the federal tax ID numbers (TINs) associated with each Carrier file claim to identify provider organizations. Our sample covers 447,009 TINs. TINs provide a measure of financial organization, with integrated physician practices typically billing under a unique TIN, although some large provider groups may organize themselves into subsidiaries, billing under separate TINs (Baker, Bundorf and Royalty, 2016). In these cases, TINs may still delineate organizational boundaries within the firm, even though they are not a perfect measure of firm boundaries.

We calculate organizational concentration at the patient-year level following the definition in equation 2.1 and the descriptions in Section 2.2. To construct these regional and shrunk PCP-level averages, we include our full initial sample of Medicare beneficiaries before implementing any of the specialized restrictions for the Mover or PCP Exit analysis samples.

We find that our baseline TIN-based measure of organizational concentration is highly correlated with an alternative definition based on physicians' reported organizational ties in the CMS Physician Compare database. Physician Compare data is only available for the final three years of our sample (2014-2016), so we cannot use it as our baseline analysis which tracks organizational concentration over a longer time period. In years where both measures are available, we use the affiliations reported in Physician Compare to construct an alternative measure of organizational ties, and compare this to our baseline TIN-based definition. The organizational concentration measures are correlated at 0.95 when averaged at the HRR level, and are correlated at 0.85 when averaged at the PCP level (prior to any shrinkage).

Earlier work by Baker et al. (2014), Austin and Baker (2015) and Baker, Bundorf and Kessler (2020*a*) has also used TINs to measure local competition across physician provider groups. This research has shown that areas with higher market concentration pay higher prices for physician services. While this prior work suggests that providers sharing the same TINs are able to leverage oligopoly power in areas with high market concentration, our paper will test whether TIN-based measures of business organization are predictive of clinical integration that may yield offsetting benefits for patients and payers.

Aggregating organizational concentration to the HRR-level and PCP-levels

To characterize the pattern of organizational concentration at the hospital referral region (HRR) level we average the patient-level measures across all patients within the relevant region.

Our primary empirical strategy exploits variation in PCPs' tendencies towards organizational concentration. The average patient in our sample is seen by a PCP who has 35 other attributed patients in the same year. To account for statistical noise in PCP organizational concentration, we apply a conventional empirical Bayes correction (Morris, 1983). This correction shrinks the estimated PCP concentration towards the year-specific mean, in proportion to the amount of estimation error.[6]

To investigate the degree of shrinkage, we calculate "pseudo shrinkage coefficients" for organizational concentration, defined as each physician's demeaned Bayesian posterior divided by the demeaned raw (not shrunk) estimate. A coefficient of one implies no shrinkage. The median coefficient is 0.89, with the 10th percentile at 0.63 and the 90th percentile at 0.99. This distribution suggests modest shrinkage, consistent with the high correlation (0.97) between the raw and shrunk measures.

For regression analyses at the HRR and PCP level, we exclusively rely on jackknifed versions of these organizational concentration measures that omit the index patient to avoid bias driven by an individual patient's need for more specialized care.

Paralleling the procedure for organizational concentration, we calculate regional and provider level measures of provider concentration to include as a control in some regressions. This measure is also jackknifed, and the PCP level provider concentration is also shrunk with an empirical Bayes procedure.

2.3.3 Outcome measures

Our primary outcome variable is a patient's annual healthcare utilization, which aggregates a patient's spending across the Medicare Inpatient, Outpatient and Carrier claim files. Utilization measures are constructed using a fixed set of annual Medicare prices expunged of regional price adjusters.[7]

We also study the relationship between organizational concentration and several utilization-based measures of healthcare quality. We study two measures related to

[6] To implement the empirical Bayes correction, we estimate a random effects model where patient-level organizational concentration depends on year fixed effects and PCP-year random effects. To achieve jackknifing, we omit the index patient from this regression. We recover empirical Bayes estimates of PCP-year organizational concentration as the sum of the year fixed effect and the best linear unbiased predictor of the PCP-year random effect.

[7] Medicare prices include some regional adjustments on the basis of local wage indices, and we do not want this source of regional variation in wage indices to confound the relationship between organizational concentration and spending. Following Finkelstein, Gentzkow and Williams (2016), we adjust total spending to strip away variation that is due to regional price adjustments.

the use of hospital care: a binary indicator for any inpatient hospitalization, and a binary indicator for any emergency department (ED) visit. Following Venkatesh et al. (2017), we define ED visits as any Carrier claim with a HCPCS code for E&M care in an ED setting. One potential cost of poorly coordinated care is additional low-value or duplicative imaging tests. We define an imaging test as repeated if it follows a prior test on the same body part with the same imaging modality within 30 days. Lastly, we examine the effects of organizational concentration on the indicators of healthcare quality for patients with diabetes: HbA1c test, and LDL test. These outcomes are only defined for the sub-sample of patients with diabetes, as defined by the Chronic Condition Warehouse; tests are identified with HCPCS codes.

2.4 Descriptive Evidence on Organizational Concentration

2.4.1 Summary statistics

Table 2.1 provides summary statistics for the Broad Sample (column 1), the Patient Mover Sample (column 2), and the PCP Exit Sample (column 3). Summary statistics suggest these samples are broadly similar. The average level of organizational concentration is 0.45, demonstrating that most patients regularly seek outpatient care across multiple organizations. The average level of provider concentration is lower than average organizational concentration at 0.38, as expected given that patients will often see multiple providers within the same organization. Average care utilization is $8641 per year; utilization is lower in the PCP exit sample, perhaps in part due to the disruptive impact of PCP exits.

Appendix Table A3 further reports the mean and standard deviation of the patient-level, PCP level, and HRR level measures of organizational concentration, provider concentration, and total utilization. We use the standard deviations reported here to interpret the scale of our regression results. The empirical Bayes procedure recovers an estimate of the true standard deviation of organizational concentration across PCPs; as expected the adjusted standard deviation of 0.13 is slightly lower than that of the raw means (0.16).

Large variation between regions in healthcare usage suggests that some regions may be inefficient (Skinner, 2011), and prior research has sought to explain why this variation exists (e.g. Cutler et al. (2019); Molitor (2018); Frakes (2013); Finkelstein, Gentzkow and Williams (2016)). We examine how organizational concentration varies across regions in Figure 2·1. This map displays residual variation in organizational concentration across regions, after accounting for the role of provider concentration, age, sex, and race. As shown in the map, the West and Upper Midwest have higher organizational concentration than would be predicted by their provider concentration and demographics, while the South and Mid Atlantic have lower organizational concentration.

Figure 2·2 shows binned scatter plots relating organizational concentration and total healthcare spending. In Panel A, the observation is the regional (HRR) average, while in Panel B, the observation is the average of patients attributed to the same PCP. Panel A illustrates that regions with higher organizational concentration have lower levels of care utilization on average; we will investigate this relationship in more detail with our analysis of patients who move across regions, while Panel B shows that patients of PCPs with higher organizational concentration have lower levels of healthcare utilization. These patterns motivate our study of PCP exits in Section 2.5.

The patterns uncovered in these descriptive graphs motivate our analytic approach. First, they suggest a link between organizational concentration and care utilization, which we will investigate for the remainder of this paper. Second, it will be important to separate organizational concentration from variation in provider concentration; we focus on residual variation in organizational concentration conditional on provider concentration. Finally, given the possible endogenous link between patient health status and organizational concentration, we focus on econometric strategies which allow us to plausibly isolate the supply-side variation in organizational concentration from variation in patient demand for care.

2.4.2 Regional variation in organizational concentration and patient moves

Previous work has examined patients who move between regions to identify the effect of regional practice variation on spending (Finkelstein, Gentzkow and Williams, 2016; Agha, Frandsen and Rebitzer, 2019). Here, we use the same mover design to examine how regional organizational concentration correlates with the care received by moving patients. When moving between regions, patients are exposed to a change in the local pattern of organizational concentration. We provide descriptive evidence on the possible role of organizational concentration in shaping regional differences in care. Following prior work, we run regressions of the form:

$$\begin{aligned} Y_{it} = \quad & \delta_1 \Delta OrgConc_{region(i)} \times post_{it} + \delta_2 \Delta ProviderConc_{region(i)} \times post_{it} \\ & + x_{it}'\beta + \alpha_i + \gamma_t + \tau_{(i,t)} + \epsilon_{it} \end{aligned} \tag{2.3}$$

where Y_{it} is the outcome of interest, $\Delta OrgConc_{region(i)}$ is the change in regional organizational concentration experienced when patient i moves, and $\Delta ProviderConc_{region(i)}$ is the change in regional provider concentration experienced when the patient moves. We also include: x_{it}, a vector of age fixed effects (in 5 year bins); α_i, an individual fixed effect; γ_t, a year fixed effect; and $\tau_{(i,t)}$, a vector of event-time fixed effects indicating the year relative to the patient move.

Figure 2·3 presents event study graphs and shows that when patients move to a region with higher average organizational concentration, they experience an immediate and persistent increase in their individual organizational concentration. Table 2.2

reports the regression results, summarizing how changes in regional average organizational concentration translate into individual patients' experiences when they move. If all regional variation were due to differences in the types of patients that lived in each region, then we would expect zero pass-through, while if movers fully adopted the average patterns of care in each region they lived, we would expect 100% pass-through. The regression in column 1 shows that about 80% of the regional difference in organizational concentration translates into patient-level changes in organizational concentration.

The final columns of Table 2.2 show how moving to a region with a different level of average organizational concentration is associated with changes in total care utilization. Column 2 shows that moving to a region with 1 standard deviation (SD) greater regional organizational concentration (an increase of 0.05) is associated with a 4.6% decline in total utilization. However, we know that changes in regional organizational concentration are also correlated with changes in regional provider concentration. Column 3 adds a control for the region's provider concentration, and finds that the relationship between organizational concentration and total utilization diminishes only slightly: a 1 SD increase in regional organizational concentration is associated with a 3.7% decline in total utilization. These results suggest that the spread of patient care across distinct organizations is an important predictor of regional variation in health care utilization.

2.5 Identification Strategy: PCP Exits

In the previous section, we described how regional variation in organizational concentration predicts spending outcomes. The hurdle for interpreting these findings is that regional organizational concentration may also be correlated with other features of the local healthcare environment. To address this concern, we turn to our study of

PCP exits. When a PCP exits a local market, due to a retirement or long-distance move, that PCP's patients must find new care providers within their local market. This natural experiment allows us to study exogenous variation in PCP practice style holding constant many features of the local healthcare market.

Organizational concentration may depend on a patient's PCP for a few reasons. First, PCPs may deliberately choose to refer preferentially to other providers within a multispecialty practice. In addition, PCPs themselves may be affiliated with a large organization that is tied to many local specialists, increasing the organizational concentration that would occur even without preferential referrals. We characterize each PCP's practice pattern with their average organizational concentration. We then test what happens to patient-level organizational concentration and healthcare utilization when a PCP exit forces the patient to switch to a new PCP with a different level of organizational concentration.

Our study of PCP exits thus analyzes how *changes* in the organizational concentration of a patient's assigned PCP affects the patient's outcomes. Because we observe patients who switch PCPs, we can include patient fixed effects in our regression model to control for any fixed patient attributes that influence their healthcare utilization. However, patients may endogenously sort to new PCPs on the basis of changes in their demand for care. For instance, patients who have gotten sicker may deliberately seek out multispecialty practices or well-known health systems when their original PCP exits. This type of sorting would bias our estimation of how organizational concentration affects healthcare spending within a difference-in-differences framework, since patient fixed effects would not adequately capture changes over time. As a result, we focus our analysis on an instrumental variables strategy adapted from Laird and Nielsen (2017) and Abaluck et al. (2020).

Our instrumental variables (IV) approach exploits the statistical property of mean reversion to predict the change in the organizational concentration of a patient's assigned PCP after their original PCP exits. Patients whose initial PCP was highly concentrated will on average experience a decrease in their PCP's organizational concentration when they switch providers. Patients whose initial PCP had low concentration will on average experience an increase in their PCP's organizational concentration.

The exclusion restriction for this identification strategy requires that *changes* in patient demand for care are not endogenously related to the *level* of organizational concentration of the original PCP. While we cannot test this assumption directly, we investigate event-study graphs to assess whether patients with different original PCP organizational concentration are on differential trends prior to that PCP's exit. The monotonicity assumption for this strategy requires that having an original PCP with high organizational concentration can only increase the probability that the patient experiences a decline in the PCP organizational concentration after the original PCP exits. This should hold when patients use a similar approach to selecting their second PCP as they applied when searching for the original PCP. We discuss these IV assumptions in more detail in the next section.

2.5.1 Estimating equations

To fix ideas, we consider first a simple difference-in-difference regression, noting that the change in PCP organizational concentration is potentially endogenous. We then lay out our IV regression equations. Letting i index patients, t index calendar years, and τ index years relative to the exit of a patient's PCP, the difference-in-difference equation we estimate is:

$$Y_{it} = \delta_1 \Delta OrgConc_{PCP(i)} \times post_{it} + x_{it}'\beta + \alpha_i + \gamma_t + \tau_{(i,t)} + \epsilon_{it} \tag{2.4}$$

where Y_{it} denotes a patient-level, time-varying outcome; in our baseline specifications, we consider two outcomes, the patient's total healthcare utilization and the patient's experienced organizational concentration. We define $\Delta OrgConc_{PCP(i)}$ as the difference between the destination PCP's organizational concentration in the year after the move minus the origin PCP's organizational concentration in the year before the move : $\Delta OrgConc_{PCP(i)} = OrgConc_{destinationPCP(i,\tau+1)} - OrgConc_{originPCP(i,\tau-1)}$. The new PCP is defined as the patient's plurality provider in the year following his original PCP's exit. This is interacted with the indicator variable, $post_{it}$, equal to 1 in periods after a patient's original PCP has exited, and zero otherwise. As a result, the coefficient δ_1 identifies how changes in care utilization before and after PCP exit relate to changes in PCP organizational concentration practice style. [8]

The regression controls for individual patient fixed effects α_i and year fixed effects γ_t, as well a time-varying patient characteristic (age) in x'_{it}. The regression also includes a vector of event time fixed effects $\tau_{(i,t)}$ indicating the year relative to the PCP exit event; these controls will account for any differential trends or disruption in care when PCPs exit that are experienced uniformly by all patients whose physician exits, regardless of the exiting physician's specific practice style.

The challenge to interpreting this difference-in-differences regression is that patients may endogenously sort to new PCPs on the basis of *changes* in their health status. To overcome this identification challenge, we do not estimate the difference-in-differences regression directly, but instead focus on an instrumental variables strategy.

When a patient's PCP exits the market due to a retirement or long-distance move, the patient is forced to find a new provider. On average, patients tend to switch to more typical providers. This pattern implies that a patient's lagged PCP exit will predict their care utilization differentially depending on the organizational

[8]Recall, the PCP's organizational concentration measures are defined in a jackknifed manner that omits the index patient from the calculation to avoid mechanical endogeneity.

concentration of their exiting PCP. This insight underlies our instrumental variables strategy, which builds on recent work with similar instruments by Abaluck et al. (2020) and Laird and Nielsen (2017). Our first stage equation uses the initial PCP's organizational concentration, denoted $OrgConc_{PCP(i),initial}$, to predict the change in organizational concentration when the initial provider exits:

$$\begin{aligned} \Delta OrgConc_{PCP(i)} \times post_{it} = \quad & \delta_1^o OrgConc_{PCP(i),initial} \times post_{it} \\ & + x'_{it}\beta^o + \alpha_i^o + \gamma_t^o + \tau_{i,t}^o + \epsilon_{it}^o. \end{aligned} \tag{2.5}$$

With the fitted values from this first stage equation, we construct a two-stage least squares estimate of equation 2.4.

Interpreting δ_1 from our instrumental variable estimates as the average causal impact of the PCP's organizational concentration on individual outcomes requires several assumptions, which we describe here. Under the assumption of constant treatment effects, assumptions 1 and 2 below suffice to recover treatment effects of being treated by a PCP with higher organizational concentration. If there are heterogeneous treatment effects, then assumptions 3 and 4 are needed to ensure that we recover average treatment effects.[9] Finally, assumption 5 is needed to interpret PCP organizational concentration (rather than another correlated dimension of PCP practice style) as the underlying reason for the differences in patient care utilization.

1. **First stage**: The original PCP's level of organizational concentration must predict the patients' change in PCP organizational concentration after the original PCP exits, conditional on included covariates. This assumption is directly testable; we report first stage F-statistics along with our IV results.

2. **Exclusion restriction**: This assumption requires parallel trends among pa-

[9] Assumptions 3 and 4 together are similar to the fallback condition described in Abaluck et al. (2020).

tients with different initial exposure to PCP organizational concentration. Specifically, patients who are initially attributed to PCPs with high levels of organizational concentration must be on the same counterfactual utilization trajectory as patients whose initial PCP has a lower level of PCP organizational concentration. We assess the plausibility of this assumption with event study graphs.

3. **Monotonicity**: Having an origin PCP with high organizational concentration can only increase the probability that the patient experiences a decline in the PCP organizational concentration after the original PCP exits. This is satisfied if patients use similar selection strategies to find a replacement PCP as they used to find their original PCP. For example, this assumption would be violated if some patients of high organizational concentration PCPs deliberately seek out a PCP with an even higher concentration due to their experience with the original PCP.

4. **No differential selection on gains**: Conditional variation in the original PCP's organizational concentration must not predict the degree of selection on gains in choosing a new provider. The treatment effect of switching PCPs is independent of the exit timing and the practice styles of the exiting PCP.

5. **Organizational concentration selection on observables only**: Other factors that influence a PCP's effect on patient care utilization must be uncorrelated with organizational concentration, after controlling for observed patient and provider characteristics. Without randomized manipulation of referral patterns, this is a strong assumption, and we discuss it in more detail below. When this assumption is violated, our estimate can be interpreted as the causal effect of switching to a higher organizational concentration PCP, rather than isolating the effect of organizational concentration from other dimensions of practice

style.

Although the PCP exit strategy approach holds the regional practice environment fixed, PCP practice style is still multidimensional. A PCP's organizational concentration may be correlated with other aspects of the PCP's practice style, which would violate assumption 5 (selection on observables only) described above. In particular, physicians who make more referrals, ceding more of their patients' care to other internists and specialists, will have more opportunities to reduce the organizational concentration. Prior research has documented that concentrating patient care within a narrow set of providers (*provider concentration*) is associated with lower levels of utilization (Agha, Frandsen and Rebitzer, 2019; Hussey et al., 2014; Frandsen et al., 2015).

To establish that the impact of organizational concentration is distinct from the well-studied phenomenon of provider concentration, our main regression specifications include both measures. Moreover, we instrument for the change in provider concentration using an analogous approach to how we instrument for the change in organizational concentration: with the provider concentration practice style of the exiting PCP. Defining $\Delta ProviderConc_{PCP(i)}$ as the difference between the new PCP's provider concentration and old PCP's provider concentration, we estimate a new first stage for organizational concentration as follows:

$$\begin{aligned} \Delta OrgConc_{PCP(i)} \times post_{it} = \quad & \delta_1^o OrgConc_{PCP(i),initial} \times post_{it} \\ & + \delta_2^o ProviderConc_{PCP(i),initial} \times post_{it} \\ & + x_{it}'\beta^o + \alpha_i^o + \gamma_t^o + \tau_{i,t}^o + \epsilon_{it}^o. \end{aligned} \tag{2.6}$$

We also estimate a parallel first stage equation for $\Delta ProviderConc_{PCP(i)}$. Finally, we estimate the second stage equation, instrumenting for both endogenous variables:

$$\begin{aligned} Y_{it} = \quad & \delta_1 \Delta OrgConc_{PCP(i)} \times post_{it} + \delta_2 \Delta ProviderConc_{PCP(i)} \times post_{it} \\ & + x'_{it}\beta + \alpha_i + \gamma_t + \tau_{i,t} + \epsilon_{it}. \end{aligned} \tag{2.7}$$

Further, we test the robustness of our findings to adding controls for PCP characteristics and practice environment. These specifications control for PCP gender, experience, residency training, and the size of the PCP's practice organization. Larger firms may hire higher quality staff, have greater capital investment, or different managerial quality; by controlling for the size of the PCP's practice organization, we can separate any general benefits of having a PCP who is employed by a large firm from the effects of organizational concentration.

2.6 Results

This section uses our instrumental variables strategy to show how PCP organizational concentration affects healthcare utilization. After discussing our baseline findings, we consider several alternative specifications, and then explore the relationship between organizational concentration and care quality.

2.6.1 PCP organizational concentration and utilization

To analyze how care patterns respond when a patient's PCP exits, we begin by examining Figure 2·4. These graphs exploit the same variation underlying our instrumental variables approach, but instead of including a single indicator variable for the post period, they include a vector of fixed effects for each year relative to the PCP exit event. The endogenous variables of interest are the interaction of these relative event time fixed effects with the change in PCP organizational concentration, and the instrumental variables are the vector of interactions between these relative event time fixed effects and the original PCP's organizational concentration.

The figure illustrates that when a patient's PCP exits the local market, the patient's care outcomes shift sharply towards the practice style of their new PCP. In Panel A, we show that if the new provider is predicted to have higher organizational concentration (so their patients receive care at fewer distinct organizations), the patient's experienced organizational concentration also increases. This establishes that PCP organizational concentration plays an important role in shaping patient-level organizational concentration, even when the patient remains in the same geographic location. In Panel B, we show that if the new provider is predicted to have greater organizational concentration, the patient's total healthcare utilization declines.

In both panels of this graph, we note an absence of pre-trends prior to the move. This demonstrates that patients whose original PCPs have different levels of organizational concentration are not on differential trends of care utilization prior to the original PCP's exit. This pattern supports the exclusion restriction, described as assumption 2 above. We also see that in year 1, the first full calendar year after their PCP has exited, patients have the largest year-over-year change on both experienced organizational concentration and utilization. The new PCP's influence may grow over time, as she gradually shapes the set of referred providers that the patient consults. In subsequent years 2 through 5, patients' care evolves to conform more closely to the practice style of their new PCP.

Our IV regressions in Table 2.3 show that the effects of organizational concentration on utilization that are large and robust to accounting for other dimensions of PCP practice style, training, and practice setting. We instrument for the change in organizational concentration with the level of organizational concentration at the original PCP. The estimated first stage equation in column 5 is strong, and shows that coming from an origin PCP with a 0.1 higher organizational concentration predicts a 0.043 greater decrease in the new PCP's organizational concentration. The associated second stage with this specification in Column 1 finds that about 29% of the variation in PCP organizational concentration practice style translates into the patient's individually experienced organizational concentration.

Columns 2-4 contain our main results relating organizational concentration to spending, while columns 6-8 below each second stage result contain the associated first stage equation for that set of controls.[10] Column 2 shows that a 0.1 instrumented for increase in organizational concentration leads to an 11% decline in healthcare utilization. The estimated standard deviation of organizational concentration across PCPs after applying Bayesian shrinkage is 0.13 (see Appendix Table A3), suggesting that a 1 SD increase in PCP organizational concentration leads to a 14% decline in utilization. Column 3 shows that this effect persists and is attenuated only slightly by the inclusion of provider concentration as an additional endogenous variable. Though the standard error on the estimate doubles, the relationship between organization concentration and care utilization remains statistically significant at the 1% level. This result shows that the frictions that arise when care crosses firm boundaries are distinct from previously studied concepts of provider concentration.

The main hurdle to interpreting this relationship as the causal effect of organizational concentration is that PCPs with more concentrated practice styles may differ along other dimensions besides their organizational concentration. By focusing on PCP exits experienced by patients who are not themselves moving, we are able to hold constant many features of the local healthcare environment. Nevertheless, PCPs' training, practice environment, and taste for aggressive care may covary with the PCP's tendency to concentrate care within an organization. To address this concern, we introduce controls for PCP gender, residency training, and experience

[10] Note that columns 5 and 6 share a common first stage since they differ only in the choice of the dependent variable, so column 6 simply repeats column 5.

(based on medical school graduation year). Further, we control for the size of the PCP's practice organization, as measured by the number of distinct providers billing to the TIN, as well as the number of claims billed to the TIN. By controlling for the organization size, we can account for the possibility that physicians working in larger practice groups have different quality, practice style, or access to capital inputs.

Reassuringly, we find no attenuation of the relationship between the PCP's organizational concentration and patient utilization once we account for PCP characteristics and practice size. Our preferred, most controlled specification (Column 4) shows that a 1 SD increase in PCP organizational concentration is predicted to reduce health care spending by 10%. The robustness of our findings to these controls provides evidence that our results are driven by differences in organizational ties, and are not an artifact of different practice settings, physician training, or experience.

2.6.2 Robustness and alternative specifications

Difference in differences results. These findings can be contrasted with the difference in differences specifications reported in Appendix Table A4. Without the instrumental variable approach, we estimate a smaller effect of PCP organizational concentration on care utilization. We believe these results are attenuated due to confounding. Patients who find themselves in worsening health are more likely to seek out care at large, integrated practices that include a wide array of specialists. PCPs affiliated with these practices are likely to have higher organizational concentration, but the patients who endogeneously select them may have increasing demand for health care services. This comparison highlights the motivation behind the instrumental variables approach. Specifically, a patient's choice of new PCP after their original PCP exits is likely to be endogenous to changes in the patients' demand for care. By isolating the variation in PCP organizational concentration that is predictable due to mean reversion, the IV approach avoids relying on these endogenous selection patterns to estimate the impact of organizational concentration.

Exploring the role of PCP provider concentration. Appendix Table A6 provides more detail on our results, specifically reporting our instrumental variable results on how PCPs' *provider* concentration practice style affects care utilization. In column 1, we estimate an alternative specification that only includes PCP provider concentration as an endogenous variable, omitting organizational concentration from the model. As expected, patients whose PCPs tend to concentrate their patients' care within a smaller set of providers also have lower spending. This finding corroborates the pattern found in the earlier literature on provider fragmentation (Agha, Frandsen and Rebitzer, 2019; Frandsen et al., 2015; Austin and Baker, 2015), and shows that the finding holds under a new identification strategy– our instrumental variables approach. However, once we add PCP organizational concentration as an additional endogenous variable in our IV framework, the estimated effect of provider concentration attenuates and becomes statistically insignificant, as seen in columns 2 and 3. These results suggest that some of the spending previously attributed to the spread of care across providers may have actually reflected the challenges of coordination across organizations. Accounting for the role of organizational coordination diminishes the role of provider concentration.

Accounting for patient demand for specialized care. Appendix Table A5 establishes that the relationship we uncover is also robust to including detailed controls for the number and type of providers the patient consults. Specifically, we extend our instrumental variables specification to include additional controls for the number of generalist providers the patient sees, as well as the number of specialist providers the patient sees. The estimated effect of organizational concentration remains large and statistically significant; the point estimate is actually larger than that reported in Table 2.3. The larger coefficient suggests these results may in fact overstate the relationship between organizational concentration and care utilization. Specifically, patients with high organizational concentration PCPs who consult many doctors may have less underlying demand for care than patients who see more doctors with a low organizational concentration PCP. This could occur, for example, if large practices with greater organizational concentration (because they cover a wider breadth of specialists) also tend to rotate patients across providers more commonly.

Decomposing the source of utilization changes Appendix Table A7 disaggregates our findings on care utilization to identify how different types of care respond. Specifically, we consider three categories of utilization: Carrier file claims, which cover professional billings; Outpatient file claims, which cover institutional claims for outpatient care; and Inpatient file claims, which cover hospital billings. Patients treated by PCPs with higher organizational concentration have lower spending on professional services (carrier) and outpatient institutional care. Taken together, these results confirm that outpatient care utilization is lower when the PCP has high organizational concentration. The estimated effect on inpatient spending (conditional on having an inpatient admission) is also negative, but has a large standard error and is not statistically significant.

2.6.3 Organizational concentration and quality of care

In this section, we explore the relationship between organizational concentration and quality of care. While the quality of ambulatory care is multidimensional and difficult to quantify empirically, we present evidence on a variety of measures related to the provision of low-value care (duplicate imaging), high-value care (recommended monitoring of patients with diabetes), and use of intensive care settings (inpatient or emergency department) which may signal deficiencies in outpatient care. Results are reported in Table 2.4. In this table, we report our most controlled specification from Table 2.3, including PCP provider concentration as an endogenous variable and controlling for the full set of PCP characteristics and PCP organization size.

An important pathway by which organizational concentration could reduce total spending is by reducing the use of inpatient care. Recall that we define organizational concentration solely using outpatient provider interactions. As a result, there is no direct, mechanical relationship between organizational concentration and the PCP's propensity to recommend hospitalization, since care delivered in the hospital setting will not contribute to the concentration measure. We do not find statistically significant effects of changes in organizational concentration on hospital-related outcomes, though standard errors are large.

Next, we investigate process of care measures for patients with diabetes. We rely on two quality of care measures, adapted from the HEDIS guidelines: receiving a regular HbA1c test and LDL test. Switching to a physician with 0.1 higher organizational concentration leads to a 4.5 percentage point increase in HbA1c testing and a 5.8 percentage point increase in LDL tests; these relationships are statistically significant at the 5% and 1% level, respectively. Patients with diabetes are more likely to receive guideline-concordant care when their PCP has greater organizational concentration. Recall that this specification does not simply reflect the benefits of being treated in a large practice group (which might proxy for investment in clinical decision support or other electronic reminder system), because we control for the size of the PCP's practice organization. Rather, this finding suggests that keeping the patient's primary and specialty care integrated may lead to fewer gaps in care for chronically ill patients.

Finally, we turn to testing and imaging. Using BETOS codes, we identify Carrier claims for laboratory tests and diagnostic imaging. Changes in organizational concentration do not lead to statistically significant changes in the use of lab tests. By contrast, switching to a more concentrated PCP decreases the number of claims for imaging tests, with a 0.1 increase in organizational concentration reducing imaging claims by 5% (a decline of 0.2 claims from a base of 4.4, $P = 0.056$). We also specifically investigate a measure of repeated imaging, which we define as imaging of the same body part with the same imaging modality repeated within 30 days. While some duplication of this sort is clinically indicated, the measure will be sensitive to repeated imaging that occurs when patients seek care across different organizations that lack seamless systems for image transfer. The coefficient on repeated imaging is very imprecisely estimated relative to the mean and not significantly different from zero. These findings suggest that while reduced imaging may contribute to the utilization reductions, these reductions are not primarily driven by changes to repeated imaging tests.

Appendix Table A8 further investigates the relationship between organizational concentration and measures of preventive care provision. We find no consistent pattern between organizational concentration and preventive care. Higher organizational concentration predicts increases in mammogram and prostate cancer screenings, declines in colorectal cancer screenings, and little change in the provision of pap smears, pelvic exams, flu shots, and cardiovascular screenings.

2.7 Conclusion

In this paper, we explore the coordination challenges that arise when clinical care is split across firm boundaries. Firms may facilitate both informal relationships among care providers, as well as firm-specific investment in coordination technology. In the healthcare setting, coordination technology could include messaging systems, investments in health information technology, and established norms for passing off patient information across providers.

Studying patients who move regions, we document that regions with higher levels of organizational concentration also have lower levels of care utilization. This pattern suggests a role for organizational concentration in explaining regional variation in healthcare spending.

Our main analysis studies patients who stay in the same area after their PCP exits the local market due to a retirement or move. Patients who switch to a PCP with higher organizational concentration experience reductions in care utilization, relative to patients who switch to a PCP with lower organizational concentration. These relationships persist after conditioning on detailed measures of how many generalist and specialist providers the patient sees, and how concentrated the patient's care is across those providers. This evidence indicates that the organizational ties between a patient's healthcare providers have an impact on their total healthcare utilization.

Our estimated effect (10% decrease in utilization from a 1 SD increase in PCP organizational concentration) is large relative to other healthcare interventions. By way of comparison, Agha, Frandsen and Rebitzer (2019) find that moving to a *region* with 1 SD higher provider fragmentation increases care utilization by 10%. Clemens and Gottlieb (2014) estimate that a 2 percent increase in payment rates leads to a 3 percent increase in healthcare utilization. The introduction of a major policy initiative, Accountable Care Organizations and the Medicare Shared Savings Program, led to comparatively small reductions (less than 5%) in spending (McWilliams et al., 2018).

Although switching to a PCP with greater organizational concentration is associated with lower total utilization of physician services, we see no evidence that higher organizational concentration reduces quality of care. In fact, PCPs with greater organizational concentration perform better on these measures of effective care for patients with diabetes.

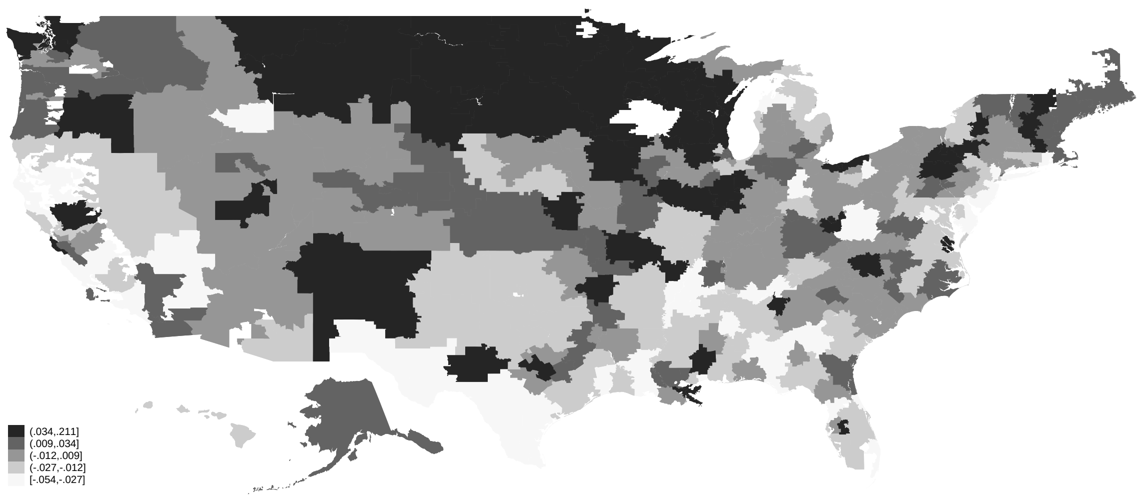
Taken together, these findings point to a potential mechanism by which higher organizational concentration lowers utilization. When providers share an organizational affiliation, they are likely to have lower barriers to information sharing and greater trust. These benefits may reduce gaps in care—e.g. resulting in better monitoring of diabetes patients—and improve hand-offs between providers. In turn, these improvements may allow providers to avoid unnecessary referrals, ensure that referred patients have already completed the requisite workup, and centralize follow-up care with the patient's PCP. Each of these effects may reduce low-value visits that generate repeated contact with specialists.

It is also worth considering alternative explanations of these findings. Large organizations may hire higher-quality physicians. If this were the case, we would expect that our result would attenuate when we account for the size of the PCP's organizational affiliation, but our empirical estimates show no such attenuation. Another possibility is that it may be more difficult to get a timely appointment in a large, multi-specialty practice, leading to lower care utilization. If this were the primary explanation, we might expect patients to substitute to more intensive forms of care that do not require appointments, such as emergency department visits; but, we find no evidence of substitution along this margin.

While our results suggest potential savings associated with care delivered at integrated multispecialty practices, any gains from reduced utilization would need to be weighed against the higher prices likely paid by private insurance providers to larger practices that have more bargaining power. The Medicare claims we study are paid at administratively set prices, so an investigation of countervailing price effects is beyond the scope of this paper. These results also raise the question of whether horizontal mergers that create multispecialty physician practices generate the savings from reduced utilization described here. If these gains occur, they may take time to develop as providers adapt to changing communication systems and adopt new referral patterns.

Our findings illuminate the role that firm boundaries play in organizing economic activity. Future research examining the detailed mechanisms of how these boundaries affect teamwork and care coordination may be able to show how some of the benefits of organizational concentration could be replicated without financial integration– for example, through better integration of health information technology systems, or by co-locating distinct provider groups.

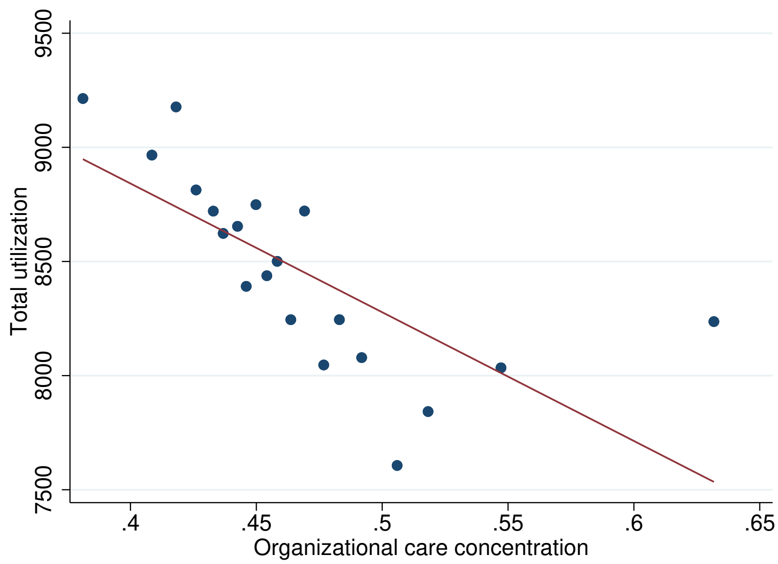
Figure 2·1: Residual of organizational concentration



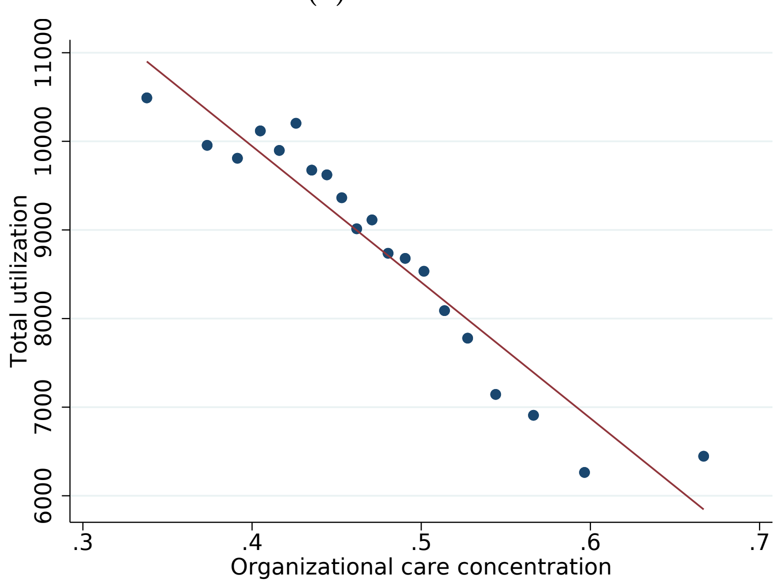
Notes: This map shows the mean residuals of patients' organizational concentration after regression adjustment for regional differences in average provider concentration, age, sex, and race. Organizational concentration and provider concentration are calculated as Herfindahl-Hirschman Index based on patients visits across healthcare organizations and providers, respectively. Hospital Referral Regions (HRRs) in darker gray have higher residual organizational concentration. Data is from the initial analytic sample, covering 9,132,322 beneficiaries.

Figure 2·2: Relationship between Organizational Concentration and Healthcare Utilization.

(a) HRR level



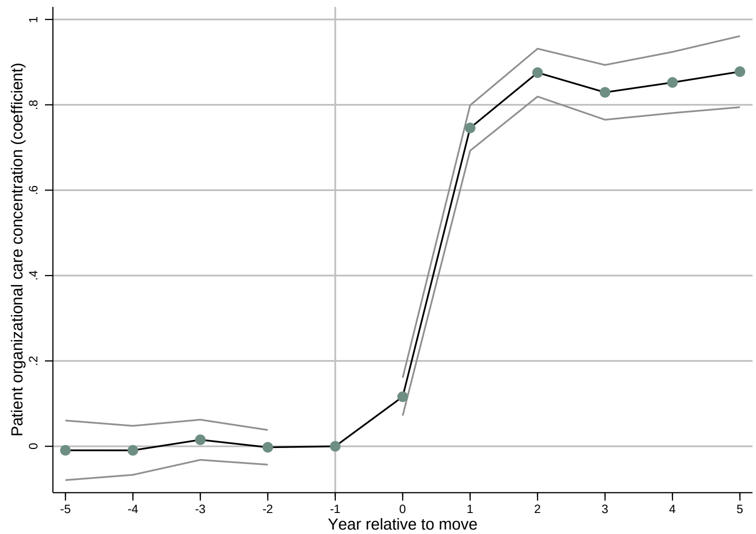
(b) PCP level



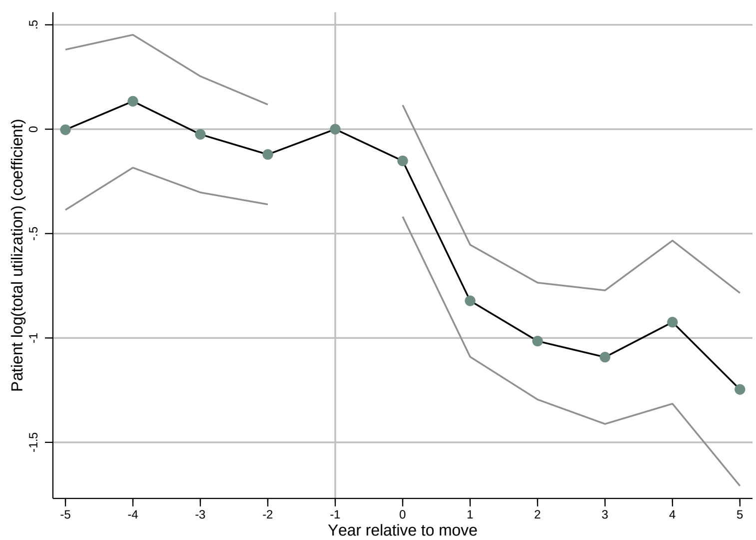
Notes: These binned scatterplots show the relationship between organizational concentration and total healthcare utilization. Panel (A) shows the relationship between these measures averaged at the Hospital Referral Region level, while Panel (B) shows the relationship between these measures averaged at the PCP level.

Figure 2·3: Event study figures. Based on patient movers.

(a) Response of patients' organizational concentration to changes in regional organizational concentration



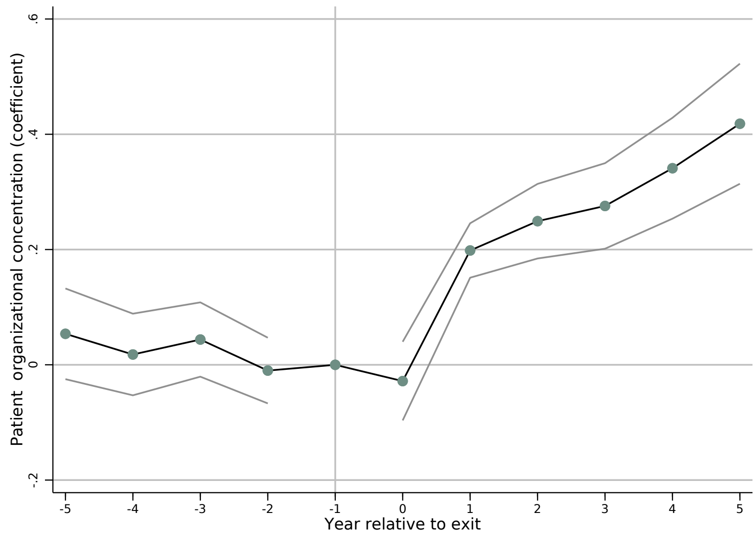
(b) Response of patients' total utilization to changes in regional organizational concentration



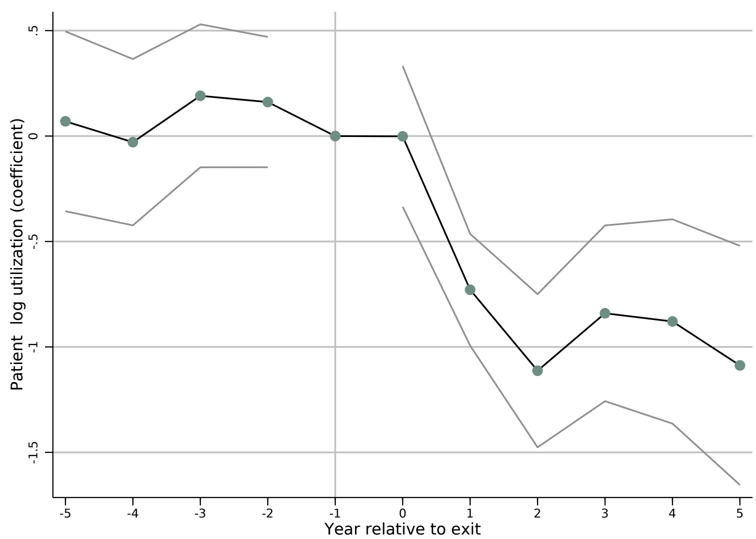
Notes: The two subplots show the estimates and 95% confidence intervals from two separate regressions. The dependent variables of subplot A and B are patients' organizational concentration and log utilization, respectively. Plots coefficient on the change in regional organizational concentration interacted with relative year. Both regressions control for patient age (five-year binned), calendar year fixed effects, and patient fixed effects. Standard errors are clustered at HRR and patient level.

Figure 2·4: Event study figures. Based on PCP exit.

(a) Response of patients' organizational concentration to changes in PCP organizational concentration



(b) Response of patients' total utilization to changes in PCP organizational concentration



Notes: The two subplots show the estimates and 95% confidence intervals from two separate regressions. The dependent variables of subplot A and B are patients' organizational concentration and log utilization, respectively. Regression specification matches the instrumental variable regressions in Table 2.3 column 1 (for Panel A) and column 2 (for Panel B), except that the post variable is now a vector of fixed effects for relative year. Both regressions control for patient age (five-year binned), calendar year fixed effects, and patient fixed effects. Standard errors are clustered at PCP and patient level.

Table 2.1: Summary statistics of different samples

	(1) Broad Sample	(2) Patient Mover Sample	(3) PCP Exit Sample
Organizational concentration	0.45	0.42	0.46
	(0.27)	(0.25)	(0.25)
Provider concentration	0.38	0.34	0.38
	(0.27)	(0.24)	(0.25)
Total utilization ($)	8641	8673	6512
	(17,487)	(17,127)	(12,722)
Age	76.1	76.34	77.19
	(7.48)	(7.38)	(7.18)
Sex: Female	0.59	0.59	0.63
Race: White	0.86	0.87	0.86
Has Diabetes	0.28	0.29	0.33
Has Hypertension	0.62	0.65	0.73
Has Heart disease	0.32	0.34	0.3
N patient-year obs	48,436,521	7,576,900	335,868
N patients	9,132,322	1,389,790	62,924
N assigned PCPs			52,981

Notes: This table reports summary statistics for the various analytic subsamples. Column 1 describes the Broad Sample. Column 2 reports the sample underlying our mover analysis, including both patients who move and the 25% random sample of non-movers. Column 3 reports summary statistics only for patients who move. Column 4 reports summary statistics for the analytic sample underlying our analysis of PCP exits. This sample restricts to patients whose PCP exits the local market. The number of assigned PCP in column 4 includes exiting PCPs as well as the PCPs patients switched to.

Table 2.2: Patient movers and regional organizational concentration

	(1)	(2)	(3)
	$OrgConc_{it}$	$Log(total\ utilization)_{it}$	
$\Delta OrgConc_{region(i)} \times post_{it}$	0.797***	-0.916***	-0.735***
	(0.021)	(0.099)	(0.113)
Regional provider concentration			X

Notes: All regressions control for patient age (five-year binned), calendar year fixed effects, relative year fixed effects, and patient fixed effects. Regional organizational concentration is jackknifed. Standard errors are clustered at HRR and patient level. Sample: Movers Analysis Sample, N=7,576,900 patient-year observations.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.3: Organizational concentration and spending, identified from PCP exits

Instrumental Variables				
Second stage	(1)	(2)	(3)	(4)
	$OrgConc_{it}$	$Log(total\ utilization)_{it}$		
$\Delta OrgConc_{PCP(i)} \times post_{it}$	0.293***	-1.058***	-0.729***	-0.794***
	(0.021)	(0.118)	(0.251)	(0.246)
First stage	(5)	(6)	(7)	(8)
	$\Delta OrgConc_{PCP(i)} \times post_{it}$			
$OrgConc_{PCP(i)t-1} \times post_{it}$	-0.432***	-0.432***	-0.295***	-0.299***
	(0.006)	(0.006)	(0.007)	(0.007)
F-test	$1.0 * 10^5$	$1.0 * 10^5$	20,703	23,845
PCP provider concentration			X	X
PCP characteristics				X
PCP organizational size				X

Notes: Each column represents an instrumental variables regression, where instrumental variable is the exiting PCP's jackknifed organizational concentration multiplied by a post indicator. In specification (1), the outcome variable is the individual patient's realized organizational concentration and in specifications (2)-(4) the outcome variable is the patient's log of total utilization. All regressions control for patient age (five-year binned), calendar year fixed effects, relative year fixed effects, and patient fixed effects. Specifications (3) and (4) include PCP provider concentration as an additional endogenous variable, instrumented by the original PCP's provider concentration multiplied by a post indicator. Specification (4) controls for PCP characteristics: gender, experience quartile indicators, residency training indicators (internal medicine vs. family practice), and the PCP's organization size (log total number of claims billed to the PCP's TIN, and the log number of unique providers billing to the PCP's TIN). Standard errors have two-way clustering at PCP and patient levels. Cragg-Donald Wald F-test reported for first-stage. The PCP Exit Sample has 335,868 patient-year observations.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.4: Organizational concentration and measures of quality

	(1) Mean of dependent variable	(2) Coefficient on $\Delta OrgConc_{PCP(i)} \times post_{it}$
Dependent variable:		
A. Hospital outcomes		
Any inpatient visit	0.155	-0.001
		(0.072)
Any emergency department visit	0.259	-0.022
		(0.083)
B. Diabetes care outcomes		
Any HbA1C test	0.631	0.452**
		(0.189)
Any LDL test	0.590	0.578***
		(0.195)
C. Imaging use outcomes		
Number of lab test claims	14.245	0.358
		(3.356)
Number of imaging test claims	4.417	-2.127*
		(1.112)
Number of repeated imaging tests	0.263	0.163
		(0.294)

Notes: Each row corresponds to a regression. The specifications match that reported in column (4) of Table 2.3, but with alternative dependent variables. Specifically, all regressions control for changes in PCP provider concentration, PCP characteristics, PCP organization size, as well as patient age (five-year binned), calendar year fixed effects, relative year fixed effects, patient fixed effects. Both changes in PCP organizational concentration and changes in PCP provider concentration are instrumented for using the exiting PCP's practice style. Standard errors are clustered at PCP and patient level. Panels A and C use the full PCP Exit Sample (335,868 patient-year observations). Panel B uses the subset of the PCP Exit Sample of patients identified with diabetes as chronic condition (105,940 patient-year observations).
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Chapter 2 Appendix Additional Tables and Figures

Table A1: Mapping from provider taxonomy codes to specialties

Specialty	Provider taxonomy codes
PCP	207Q00000X, 207QA0000X, 207QA0505X, 207QG0300X, 207R00000X, 207RA0000X, 207RG0300X, 208000000X, 2080A0000X, 208D00000X

Notes: These codes are used to define primary care specialties from the National Plan and Provider Enumeration System (NPPES).

Table A2: List of place of service codes included as outpatient care

Place of Service Code	Place of Service Name
05	Indian Health Service Free-standing Facility
07	Tribal 638 Free-standing Facility
11	Office
17	Walk-in Retail Health Clinic
20	Urgent Care Facility
22	On Campus-Outpatient Hospital
49	Independent Clinic
50	Federally Qualified Health Center
53	Community Mental Health Center
57	Non-residential Substance Abuse Treatment Facility
58	Non-residential Opioid Treatment Facility
62	Comprehensive Outpatient Rehabilitation Facility
65	End-Stage Renal Disease Treatment Facility
71	Public Health Clinic
72	Rural Health Clinic

Notes: These codes are used to identify claims in the Medicare Carrier File for services that take place in an outpatient facility.

Table A3: Summary stats of key variables at different levels

	(1) Mean	(2) Std. Dev.
Patient level (N=9,132,322)		
Organizational concentration	0.50	0.24
Provider concentration	0.43	0.25
Total utilization	9116	14,800
PCP level (N=190,616)		
Organizational concentration (raw)	0.49	0.16
Organizational concentration (adjusted for statistical noise)	0.48	0.13
Provider concentration (raw)	0.39	0.15
Provider concentration (adjusted for statistical noise)	0.38	0.11
Total utilization	9377	11,263
Regional level (N=306)		
Organizational concentration	0.47	0.05
Provider concentration	0.38	0.03
Total utilization	8465	918

Notes: This table summarizes provider concentration, organization concentration, and utilization outcomes at different levels of aggregation. The top panel has one observation per patient, and reports the means and standard deviations across all patients. The middle panel has one observation per PCP (averaged across patient-year observations), and reports the mean and standard deviation across PCPs. The bottom panel has one observation per Hospital Referral Region (averaged across patient-year observations) and reports the mean and standard deviation across regions.

Table A4: Difference in differences analysis of PCP exits

	(1)	(2)	(3)	(4)
	$OrgConc_{it}$	$Log(total\ utilization)_{it}$		
$\Delta OrgConc_{PCP(i)} \times post_{it}$	0.340***	-0.623***	-0.088	-0.132
	(0.010)	(0.057)	(0.084)	(0.086)
PCP provider concentration			X	X
PCP characteristics				X
PCP organizational size				X

Notes: This table shows the difference in differences estimates of equation 2.4 without using the instrumental variable strategy to predict variation in the change in organizational concentration after a PCP exit. In specification 1, the outcome variable is the individual patient's realized organizational concentration and in specifications 2-4 the outcome variable is the patient's log of total utilization in specifications. All regressions control for patient age (five-year binned), calendar year fixed effects, relative year fixed effects, and patient fixed effects. Specifications 3 and 4 include PCP provider concentration as an additional endogenous variable, instrumented by the original PCP's provider concentration multiplied by a post indicator. Specification 4 controls for PCP characteristics: gender, experience quartile indicators, training indicators (internal medicine vs. family practice), and the PCP's organization size (log total number of claims billed to the PCP's TIN, and the log number of unique providers billing to the PCP's TIN). There are 335,868 patient-year observations. Standard errors have two-way clustering at PCP and patient levels.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A5: Instrumental variable analysis of PCP exits, controlling for number of physicians the patient consults

Instrumental Variables		
Second stage	(1)	(2)
	$Log(total\ utilization)_{it}$	
$\Delta OrgConc_{PCP(i)} \times post_{it}$	-0.794***	-1.451***
	(0.246)	(0.195)
First stage	(3)	(4)
	$\Delta OrgConc_{PCP(i)} \times post_{it}$	
$OrgConc_{PCP(i)t-1} \times post_{it}$	-0.299***	-0.310***
	(0.007)	(0.007)
F-test	23,845	23,733
PCP provider concentration	X	X
PCP characteristics	X	X
PCP organizational size	X	X
Spline N generalists seen by patient		X
Spline N specialists seen by patient		X

Notes: See notes to Table 2.3. For reference, specifications (1) and (3) replicate the results reported in (4) and (8) of Table 2.3. In specification (2) and (4), the regression adds new control variables that account for the number of distinct providers each patient sees. Specifically, these specifications control for a 4-knot spline in the number of generalist providers (as defined in Table A1: family practice, internal medicine training, or geriatrics training) and a 4-knot spline in the number of specialist providers (with any other training type). Standard errors have two-way clustering at PCP and patient levels.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A6: Impact of organizational concentration and provider concentration

	A. Baseline estimates		
Instrumental Variables			
Second stage	(1)	(2)	(3)
	Log(total utilization)$_{it}$		
$\Delta OrganizationConc_{PCP(i)} \times post_{it}$		-0.729***	-0.794***
		(0.251)	(0.246)
$\Delta ProviderConc_{PCP(i)} \times post_{it}$	-1.072***	-0.453	-0.279
	(0.108)	(0.248)	(0.243)
PCP characteristics			X
PCP organization size			X

Notes: This table reports the results of instrumental variables regressions similar to those reported in Table 2.3, but now providing further detail on the relationship between PCP provider concentration and care utilization. Column 1 reports a specification similar to that in column 2 of Table 2.3, but replacing the endogenous and instrumental variables related to PCP organizational concentration with analogous variables describing PCP provider concentration. Columns 2 and 3 are identical to the specifications reported in columns 3 and 4 of Table 2.3, which include both PCP organizational concentration and PCP provider concentration as endogenous variables, but here we report the coefficient on PCP provider concentration. There are 335,868 patient-year observations. Standard errors have two-way clustering at PCP and patient levels. See notes to Table 2.3 for further details.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A7: Instrumental variable analysis of PCP exits, spending decomposition

	(1) Mean of dependent variable (not log)	(2) Sample size	(3) Coefficient on $\Delta OrgConc_{PCP(i)} \times post_{it}$
Dependent variable:			
Log of carrier spending (professional)	2663	335,868	-0.426**
			(0.189)
Log of outpatient spending (institutional)	1364	335,868	-1.397**
			(0.586)
Log of inpatient spending (hospital, if > 0)	16,507	35,002	-0.402
			(0.538)

Notes: See notes to Table 2.3. This table replicates the instrumental variable specification reported in Table 2.3 columns (4) and (8) with alternative outcome variables that decompose Medicare billing depending on the type of bill. Inpatient billings are only defined among patients with at least one hospitalization. Sample size is 335,868 for carrier and outpatient claims; sample size is 35,002 for inpatient claims.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A8: Organizational concentration and preventive care

	(1) Mean of dependent variable	(2) Coefficient on $\Delta OrgConc_{PCP(i)} \times post_{it}$
Dependent variable:		
A. Preventive care for women ($N = 211,823$)		
Mammogram	0.567	0.423**
		(0.174)
Pap smear	0.165	-0.209
		(0.129)
Pelvic exam	0.142	-0.006
		(0.114)
B. Preventive care for men ($N = 124,042$)		
Prostate cancer screening	0.273	0.648***
		(0.227)
C. Preventive care for full sample ($N = 335,868$)		
Flu shot	0.671	0.048
		(0.119)
Colorectal screening	0.157	-0.581***
		(0.097)
Cardiovascular screening	0.909	-0.239
		(0.258)

Notes: Each row corresponds to a regression. The specifications match that reported in column (4) of Table 2.3, but with alternative dependent variables. Specifically, all regressions control for changes in PCP provider concentration, PCP characteristics, PCP organization size, as well as patient age (five-year binned), calendar year fixed effects, relative year fixed effects, patient fixed effects. Both changes in PCP organizational concentration and changes in PCP provider concentration are instrumented using the exiting PCP's practice style. Standard errors are clustered at PCP and patient level.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Chapter 3

Does Physician Affiliation Matter? The Effects of Physician Vertical Integration on Referral Patterns, Patient Welfare, and Market Dynamics

with Christopher Whaley

3.1 Introduction

Approximately 1 million physicians currently practice in the United States. The practice environment for U.S. physicians is rapidly changing. Rather than practice in traditional models of independent or solo practices, many physicians are now employed by hospitals or health systems. From 2012 to 2018, the share of physician practices employed by a hospital or health system increased from 46.8% to 54.1% (Kane, 2019). From a clinical perspective, vertical integration of physicians and hospitals is designed to enable care coordination, investments in technology, and improve patient care. At the same time, vertical integration has also raised important market competition questions. Physicians and hospitals that vertically integrate may do so to increase bargaining leverage with private insurers. Internalizing the bargaining leverage of larger hospitals allows for smaller physicians to obtain higher reimbursement rates (Peters (2014); Dafny, Ho and Lee (2019*a*)). Consistent with this model, several existing studies find that vertical integration leads to increases in prices (Capps, Dranove and Ody, 2018; Baker et al., 2014; Baker, Bundorf and Kessler, 2020*b*; Chernew et al., 2018).

A potentially more consequential impact for the market competition is the impact of vertical integration on patient demand and allocation of patients across providers. Unlike many other markets, physicians often act as gatekeepers for patients (Chernew et al., 2018). Other studies have raised the potential that hospitals may acquire "upstream" physician practices in the hopes of steering patient demand for "downstream" services to the hospital (Baker, Bundorf and Kessler, 2016*a*). This type of behavior follows common models of input foreclosure, but with few exceptions, how hospitals use vertical integration to foreclose inputs has not been fully examined. While broad acknowledgment of potential for input foreclosure exists, measuring the impacts of physician vertical integration on patient demand for hospitals remains limited. In addition, the existing and extensive literature on health care market competition has primarily focused on horizontal competition between hospitals (Dafny, Ho and Lee, 2019*b*), but has not fully considered how vertical integration can, in turn, impact horizontal competition. Related work examines the impact of vertical integration on self-referrals for skilled nursing facilities Cutler et al. (2020) and for outpatient procedures in Florida (Richards, Seward and Whaley, 2020). Fully understanding these impacts is especially important given the size of the U.S. health care market and how rapidly physician practices are vertically integrating.

In this paper, we use data covering 2013 to 2017 from a 100% sample of the U.S. fee-for-service Medicare population to test how vertical integration between physicians and hospital groups changes patient demand for downstream hospital services.[1] We focus on two procedures of particular interest—arthroscopy surgeries and colonoscopies. Both procedures are extremely common among the Medicare population—our data includes 1.5 million arthroscopy and 13 million colonoscopy

[1]Medicare is the U.S. insurance system that provides health benefits to elderly (above age 65) and disabled individuals.

procedures. In addition, both procedures are commonly performed in two distinct provider organizations—hospital outpatient departments (HOPDs), which are a unit of a traditional hospital, and ambulatory surgical centers (ASCs), which are free-standing centers that are often independent from hospitals. Several other studies have compared ASCs and HOPDs, and generally found that ASCs are less costly and have equivalent or higher quality (e.g., Gardner et al. (2005); Grisel and Arjmand (2009); Munnich and Parente (2014)).

Examining how vertical integration impacts competition between ASCs and HOPDs is particularly relevant in the Medicare population, as Medicare imposes strict fee schedules that pay each provider organization different rates. Unlike private insurance markets, where prices are negotiated, vertical integration does not allow providers to negotiate higher prices with the Medicare system. However, Medicare reimburses ASCs at approximately one-half to two-thirds the reimbursement rate for HOPDs.[2] For example, in a common procedure we examine, a colonoscopy with a biopsy, Medicare pays \$715 for procedures performed in an ASC and \$1,212 for procedures performed in HOPDs. From a purely financial standpoint, performing the same procedure in an HOPD instead of an ASC creates an arbitrage opportunity to increase Medicare payment by nearly 70% (Chernew (2021)). If vertical integration leads to increases in the use of HOPDs, this payment differential based on site of care potentially serves as both a motivation for and a consequence of hospital acquisition of physician practices. Through this channel, moving patient volume from ASCs to HOPDs can lead to increased revenue for acquirer organizations. Related work finds that increases in site-based payment differentials contribute to physician vertical integration (Post et al. (2021)).

To address these questions, we leverage changes in primary care physician affili-

[2]Procedure-specific prices by site-of-care are available at https://www.medicare.gov/procedure-price-lookup/.

ation to examine how patients choose between HOPDs and ASCs for "downstream" procedures. Starting with an aggregated-level analysis, we find reduced-form evidence that hospital employment of physician practices shifts patient demand to hospitals and away from ASCs. Our group-level analysis shows that for both procedures, following vertical integration, there is a 3.7 percentage point increase in the use of HOPDs instead of ASCs, which translates to a 6.3% relative increase for arthroscopy, and a 7.5% relative increase for colonoscopies. We also find increases in the travel distance between the patient and the provider (1.6 miles for arthroscopies and 1.1 miles for colonoscopies). Importantly, we find no change in clinical quality measures.

We then use a nested logit discrete choice model to describe the patient-physician joint decision of healthcare delivery organization for the "downstream" procedures. Our model estimates suggest that vertical integration leads to a 14.2% and 32% higher probability of choosing an HOPD rather than an ASC for arthroscopy and colonoscopy, respectively. At an aggregated level, our results imply that for just these two services, changing from status-quo vertical relationships for all physicians will lead to an increase of 309,090 cases in the use of HOPDs, a $38.2 million increase in patients' out-of-pocket payment, and a $316 million increase in Medicare spending.

At the same time, while existing studies on health care market competition consider horizontal competition between hospitals, few studies have considered horizontal competition between different delivery organizations, for example, ASCs and HOPDs. Existing studies have also not fully examined how changes in vertical relationships for upstream providers, in this case, primary care physicians, can lead to changes in horizontal competition for downstream providers. We extend our results and examine how the changes in referral patterns caused by vertical integration can lead to changes in horizontal market competition. We find that vertical integration leads to a meaningful increase in horizontal concentration among PCP groups, but does not meaningfully change horizontal concentration for downstream surgical providers.

Our results have implications for the economics of understanding provider competition in an important market and also important policy implications. Medicare and other insurers commonly pay health care providers different amounts for care performed in different settings. To the extent to which site of care payment differentials motivate vertical integration, our results show the impacts of allocation changes following vertical integration for both the Medicare system and patients. At the same time and independent of changes to payment policies, regulators have faced questions of how to address vertical integration. Our results show a clear link between vertical integration and horizontal competition, which regulators can apply when monitoring consolidation activity.

3.2 Data

We use data from a 100% sample of Medicare claims data that covers the 2013 to 2017 period. This data includes all inpatient, outpatient, physician, and other types of care for Medicare enrollees who are enrolled in Traditional Medicare (e.g., fee-for-service Medicare). We do not have information on Medicare Advantage enrollees or non-Medicare patients. These data include approximately 38 million enrollees per year. The breadth of this data allows us to examine the market structure of the entire provider market in the US.

With the data, we identify two specific procedures of interest—joint arthroscopy and colonoscopy. Joint arthroscopy is a minimally-invasive surgical procedure used to diagnose and repair minor joint damage. Colonoscopies are a procedure used to screen for and to treat early-stage colorectal cancer.[3] We restrict our sample to

[3]The CPT codes for joint arthroscopy procedures are 29830, 29831, 29832, 29833, 29834, 29835, 29836, 29837, 29838, 29862, 29870, 29871, 29872, 29873, 29874, 29875, 29876, 29877, 29878, 29879, 29880, 29881, 29882, 29883, 29884, 29885, 29886, 29887, 29805, 29806, 29807, 29808, 29809, 29810, 29811, 29812, 29813, 29814, 29815, 29816, 29817, 29818, 29819, 29820, 29821, 29822, 29823, 29824,

beneficiaries between ages 65 and 100 and who are continuously enrolled in Medicare for all 12 months in a year. We also exclude cases where the patient and provider are located more than 100 miles apart. After imposing all those restrictions, we obtain two full samples, one for each procedure. The joint arthroscopy full sample has 1,998,620 observations from 709,154 beneficiaries, 53,823 primary care groups, and 8,747 joint arthroscopy providers. For colonoscopy procedures, the full sample has 17,765,556 observations from 7,230,249 beneficiaries, 79,653 PCP groups, and 21,745 colonoscopy providers.

For each procedure, we identified if the procedure was performed in an Ambulatory Surgical Center (ASC) or a Hospital Outpatient Department (HOPD), using the site of care codes contained in the Medicare claims. We limited the procedures to those performed in outpatient settings and excluded the small number procedures performed in other settings, such as physician office, inpatient, and emergency department settings. Of these, 42% of arthroscopies and 51% of colonoscopies were performed in an ASC instead of an HOPD. As secondary outcomes, we also measured the total Medicare reimbursement amount for each procedure and patient payments, and the distance between the centroid of the patient's zip code and the zip code as the provider.[4] Finally, an advantage of these procedures is the existence of clinical complication measures. For each procedure, we identified both procedure and post-operative complications using the approach applied in previous studies, which we measure as the *z-score* of complication rates (Robinson, Brown and Whaley (2015), Whaley and Brown (2018), Whaley, Guo and Brown (2017), Aouad, Brown and Whaley (2019)). These complications, which are fully listed in Appendix Tables A1

29825, 29826, 29846. The CPT codes for colonoscopy procedures are 44388, 44389, 44390, 44391, 44392, 44393, 44394, 45378, 45379, 45380, 45381, 45382, 45384, 45385, G6019, G0104, G0120, G0121, G2204, G6020, G6024, G6025, G9252, G9253, G9659, G9660, G9661, G9933, G9935, G9936, G9937.

[4]The internal point is defined by the Census Bureau. We used the NBER zip code distance database in which the distance is calculated using the Haversine formula based on internal points in the zip code area.

and A2 , include adverse outcomes like intestine perforations, cardiac events, and readmissions. Because procedural complications are rare, we use the z-score of the fraction of patients having any complications as the complication rate. The complication rates are thus measured as the standardized share of patients that have any procedure-related complications.

3.2.1 Group identification and integration status

In addition to information on procedure types, we also used the data to collect information on patient primary care. Existing studies on vertical integration and referrals have examined market-level trends in vertical integration Scheffler, Arnold and Whaley (2018) and integration status of referring papers (e.g., Baker, Bundorf and Kessler (2016*b*); Chernew et al. (2018)). A challenge with market-level vertical integration or referring-level measures is the potential that changes in patient allocation occur before a patient is referred to a specific provider. We instead examine the impact of changes in the referral status of the patients' primary care provider. Doing so allows us to measure the impacts of vertical integration on patient allocation using the patient's initial contact with the health care delivery system. To construct patient vertical integration measures, we identify the physician group that provides the majority of primary care services to the patient. A full description of the attribution model used is described in Appendix A.

Consistent with existing studies, in our sample, we find an increasing trend in the share of providers that are vertically integrated. This trend is similar to results from surveys conducted by the American Medical Association (Kane, 2019). Figure 1 shows this trend from 2013-2017. As shown in the top panel of Figure 1, the number of physicians working at an integrated group increased from 40.6% to 44.4%. The increase in physician group vertical integration is accompanied by a 19.5% increase (from 33.9% to 40.5%) in the number of Medicare beneficiaries who receive primary care from a vertically integrated group. This larger increase in the number of beneficiaries provides descriptive evidence that the groups that became vertically integrated during this period are larger than groups that did not integrate and that the PCP market became more concentrated over the 2013 to 2017 period.

3.2.2 Market definition

We define markets based on a combination of geographical region and patient flow. Traditional fixed geographical market definitions do not vary from year to year and prohibit patients from choosing from other regions, even though some patients may live close to a geographical boundary. To relax these restrictions, we use a geography and patient flow combined measure. First, we use a fixed geographic market of Hospital Referral Regions (HRRs), which have been constructed by the Dartmouth Atlas of Care to measure hospital market areas (Atlas et al. (2009)). Starting from an individual HRR, we include all the options into the choice set for an HRR if they have ever been chosen by patients in this region and the distance to the patient is shorter than 100 miles. Using this market definition, we allow market boundaries to vary across years and allow patients to travel across HRRs if the options are popular in their HRRs and not too far from their homes. Following the definition, the average market has 365 arthroscopy patients (596 observations) and 4482 colonoscopy patients (5586 observations). Our market definition also contains a mean of 26 arthroscopy providers, of which 16 are hospital-based, and 76 colonoscopy providers, of which 45 are hospital-based.

3.2.3 Sample construction

We construct two separate analytic samples, one for our reduced-form analysis that measures physician-group level changes in referral patterns following vertical integration, and another separate sample for our patient choice model. First, for our physician group-level reduced-form analysis, we start by assigning each PCP group a home state, which is the one from which most of the patients reside. This assignment is necessary for groups operating in multiple states because the local environment, the Medicare fee schedule system, and traveling distance vary across states. We also exclude groups who churned in and out of systems between 2013 and 2017. After that, we aggregate visits to PCP group-state-month-procedure code level. We count the number of visits as the total volume and the number of visits in hospital-based facilities as hospital-based volume. For the other outcome variables, we use the average value within each cell. The reduced-form sample has 780,078 observations of 50,393 PCP groups for joint arthroscopy, and 2,665,339 observations of 74,235 PCP groups for colonoscopy.

The second sample is constructed for our choice model. We first apply the aforementioned market definition. Then, we restrict patients' choices to popular choices. For each market, we drop providers whose number of visits is below the median, keep the most popular five providers, and group the rest of them into two "others" options—one for hospital-based choices and the other one for independent choices. The last step of sample construction is dealing with repeated choices. In our sample, less than 50% of arthroscopy patients and 25% colonoscopy patients have more than one visit. For the main analysis, we keep all the visits and assume that each visit decision is made independently. [5] The final arthroscopy choice model sample includes 273,485 beneficiaries, 20,444 primary care groups, and 3,968 joint arthroscopy providers. The colonoscopy sample has 4,576,543 beneficiaries, 41,687 PCP groups, and 7,875 colonoscopy providers. Due to computational limitations, we use a 25% random sample of the final choice model samples for estimation. Note that the most important contribution of a 100% sample is to have a complete choice set and ac-

[5] We also tried a more restricted sample which only includes patients who have only one visit. The results are robust to this sample selection.

curate characteristics for each option in the choice set. In our 25% random sample, beneficiaries' choice sets are constructed based on the 100% sample.

Descriptive characteristics of each sample are presented in Table 1. Panel A presents characteristics of the group-level sample used for the reduced-form analysis, and Panel B presents characteristics of the patient-level sample used for the patient choice analysis.

3.3 Relationship between vertical integration and group-level outcomes

In this section, we use group-level aggregated data to examine the effects of vertical integration on site-of-care choice and other outcomes that patients care about.

3.3.1 Reduced Form Effects of Vertical Integration On Allocation and Patient Outcomes

To measure the reduced form effect of vertical integration on each of our five outcomes, we use a difference-in-differences approach. We estimate a regression model of the form

$$Y_{jt} = \theta VI_{jt} + \delta_t + \zeta_j + \varepsilon_{jt} \tag{3.1}$$

where Y_{jt} is the outcome of interest (e.g, ASC vs. HOPD, Medicare spending, travel distance, and procedure complication rate), VI_{jt} is an indicator of being vertically integrated for group j in time t, δ_t and ζ_j are time fixed effect and group fixed effect, respectively. The coefficient θ identifies how changes in dependent variables relate to changing to vertically integrated. We estimate this model using OLS regressions and cluster standard errors at the group (j) level.

Our primary identification assumption is that groups who select into vertical integration arraignments do so independently of changes to the outcomes we measure. We include physician group fixed effects, which account for any time-invariant differences between groups (e.g., the referral pattern preferences or patient composition). A potential threat to our identification is if PCP groups select into vertical integration due to factors that impact our outcomes (e.g., changes in underlying patient preferences for ASCs vs. HOPDs). In such a case, trends in our outcome variables of interest will likely change prior to vertical integration events. To test for these trends, we also estimate event study regressions that test for the “parallel trends” assumption common in difference-in-differences designs.

Table 3.2 reports our reduced-form regression results. The top panel shows results for arthroscopy procedures, and the bottom panel presents results for colonoscopy procedures. Column (1) shows that becoming vertically integrated leads to a 3.7 percentage-point increase in the probability of receiving an arthroscopy procedure at an HOPD, which translates to an approximately 6.3% relative increase. Along with the increase in the probability, the average Medicare spending increased by \$6.5. This effect on the Medicare allowed amount outcome is insignificant and relatively small. Due to the reasons discussed in the previous section, the effect on the allowed amount is underestimated. Columns (3) and (4) show that patient payments increase by \$15.15 more OOP and travel distance increases by 1.6 miles more for an arthroscopy procedure, which is a 10% increase in patient’s travel distance.

In the bottom panel, the reduced form effects of vertical integration are similar for colonoscopy procedures. Columns (6) to (9) suggest that primary care provider vertical integration leads to an increase of 7.5% in the probability of receiving a colonoscopy in an HOPD, an increase of \$7.1 in Medicare spending, a \$9.5 increase in patients’ OOP payment, and an additional 1.1 miles of traveling distance (a 9% relative increase). For both arthroscopy and colonoscopy, we do not find differences between vertical integration and our quality measures. The lack of a relationship suggests that the quality of arthroscopy and colonoscopy procedures has a very limited relationship with the place of service.

As an additional test of our reduced form effects, Figure 3·2 presents event study graphs that measure how each outcome changes in the months before and after vertical integration. Consistent with our difference-in-difference results, following vertical integration of primary care groups, patients are more likely to go to HOPDs and travel further for both downstream surgical procedures. Importantly, we do find evidence that these trends in these changes occur prior to vertical integration, which supports the assumptions necessary for a causal interpretation of our results.

3.4 Effect of Vertical Integration on Downstream Provider Choice

Our reduced-form approach measures the effects of vertical integration at an aggregated level. A limitation of this strategy is that the vertical integration indicator might be correlated with time-variant group characteristics, such as characteristics of patients. To mitigate this concern, we employ a discrete choice model for the choice of downstream providers. In this model, we include all three parties–patients, PCP groups, and downstream surgeons. Thus, this model enables us to incorporate patients' characteristics into the referral decision.

3.4.1 Patient and Provider Utility Function

To model patient choice of provider, we start with the patients' perspective. In the most straightforward approach, patients select surgeons based on how much they will pay for the procedure, the quality of the surgeon, and how long they need to travel. In addition to provider characteristics, patients' own characteristics also affect their utility, For example, patients with different age and gender may obtain different level of utility from a given provider. Therefore, we assume that patient i's utility from choosing surgeon k at time t is

$$v_{ikt} = X_{it}^{pat\prime}\beta^{X} + d_{ik}\beta^{d} + X_{k}^{surg\prime}\beta^{S} + X_{t}^{\prime}\beta^{W} + \delta_{t} + \xi_{ikt} \tag{3.2}$$

where X_{it}^{pat} is a vector of patient's characteristics—age and gender, d_{ik} is the travel distance between patient i and provider k, X_{k}^{surg} is a vector of surgeon's characteristics such as average patient payments and complications rate, W_t is a vector that contains region information, and δ_t is a time fixed effect.

However, the selection of downstream providers is not driven just by patient preferences, but instead incorporates provider decision-making and referral patterns. Several papers expand upon the model of patient choice in equation (2). Healthcare providers are altruistic—they take account their patients' utility in addition to their own preferences (Arrow (1963); McGuire (2000)). Rather than an isolated patient decision, the selection of surgeons for a downstream service is a joint decision that depends on both the patients' utility and the PCPs' own preference. When PCP groups are vertically integrated with hospitals or health systems, they will be strongly incentivized, if not required, to refer patients to their owning hospitals or systems, while independent PCP groups do not have such incentive.

We model this joint decision by expanding equation (2) to include both patient and provider predictors of choice. We estimate the choice utility of referring patient i to provider k at time t as

$$u_{ij(i)kt} = X_{it}^{pat\prime}\beta^{X} + X_{j(i)t}^{PCP\prime}\beta^{P} + VI_{j(i)t}\beta^{V} + d_{ik}\beta^{d} + X_{k}^{surg\prime}\beta^{S} + X_{t}^{\prime}\beta^{W} + \delta_{t} + \varepsilon_{ij(i)kt}. \tag{3.3}$$

In this model, X_{ji}^{pcp} is a vector of PCP characteristics such as group size, $VI_{(j(i)t)}$ is an indicator of whether PCP j is integrated at time t. $\varepsilon_{(ij(i)kt)}$ is the idiosyncratic error that follows the Gumbel distribution. In our model, patients' PCP are assigned annually. Thus, for a given patient i in market t, the patient's PCP j is given. To simplify the notation, we will eliminate the subscript j thereafter.

3.4.2 Model Structure

As discussed above, our two downstream services, colonoscopies and arthroscopies, can be performed in two separate delivery organizations–hospital-based facilities (e.g., HOPD) and independent facilities (e.g., ASC). The preference of patients and PCPs over these two facility types are not completely captured by the observed characteristics. Because of the obvious difference between the two types, the unobserved provider characteristics could be systematically different. Therefore, we assume that patients and PCPs have different sensitivities for the two types. In the model, we group surgeons based on their facility types and use a nested-logit model to describe the joint choice of surgeon (McFadden (1980); Goldberg (1995)). Our model has two virtual levels—top-level for facility type and bottom-level for surgeons. The top-level includes the two aforementioned facility types. For each facility type, at the bottom-level, the choice set includes all the popular surgeons with that facility type and an "others" option constructed by grouping all the less popular surgeons with that facility type. A surgeon is classified as popular in a market if the number of her/his patients ranks at the top 5 among all surgeons in that market. The outside option is not having any procedure.

Let $C_1 \in \{H, I\}$ denotes the top level choice and $C_2 \in \{K_H, K_I\}$ denotes the bottom level choice. At the bottom-level, we model the choice by a multinomial logit model. The conditional probability of patient i choosing surgeon $k \in K_f$ given facility type f is

$$Pr(C_2 = k | C_1 = f) = \frac{exp((d_{ik}\beta^d + X_k^{surg}\beta^S + \delta_t)/\tau_f)}{\sum_{l \in K_t} exp((d_{il}\beta^d + X_l^{surg}\beta^S + \delta_t)/\tau_f} \tag{3.4}$$

where τ_f is the dissimilarity parameter, indicating how similar the two types are. The more closer to one the τ_f is, the more similar the two types.

At the top-level, the probability of patient i choosing facility type f is

$$Pr(C_1 = f) = \frac{exp(X_i^p \beta_f + \tau_f I_f)}{\sum_{s \in \{H,I\}} exp(X_i^p \beta_s + \tau_s I_s)} \tag{3.5}$$

where X_f^p represents is a vector of patient-PCP group characteristics. Linking to the utility representation, X_f^p includes X_{it}^{pat}, $X_{(j(i)t)}^{pcp}$, and $VI_{j(i)t}$. I_f is the inclusive value associated with choice f,

$$I_f = ln\left(\sum_{k \in K_f} exp(\frac{d_{ik}\beta^d + X_k^{surg}\beta^S + \delta_t}{\tau_f})\right). \tag{3.6}$$

Since the integration decision needs a relatively long preparation time and it is made at the PCP group level instead of the PCP level, we assume away the endogeneity of the VI indicator in our model.

3.5 Choice Model Results

Table 3.3 presents the estimates from the choice model with different specifications for arthroscopy and colonoscopy. At the bottom of each column, we report the average marginal effect of vertical integration on the top-level choice probability. For joint arthroscopy procedures, column (1) shows that when a patient's PCP group changes from independent to vertically integrated, this patient will be 2.84 percentage points more likely to go to a hospital-based facility for the arthroscopy procedure. Compared to the average hospital-based volume share of 56%, this increase translates to a 5.1% relative increase. In column (2), we add three market characteristics—per capita income, the number of non-federal PCPs, and the number of non-federal MDs—to control for the geographical income variation and the market concentration. After controlling for the market characteristics, the effect of vertical integration reduces slightly, from 2.84 percentage points to 2.82 percentage points. In Column (3), we allow for the interaction between OOP payment and complication rate. In this way, the trade-off between OOP payment and quality is flexible. This modification does not meaningfully change the estimates. The model corresponding to column (4) allows for the interaction between surgeon characteristics and the PCP group integration status. This specification relaxes the assumption that patients and PCPs respond to the procedure price and quality in the same way, regardless of whether the PCP group is integrated or independent. When this interaction is included, the effect of vertical integration increases to 3.24 percentage points, a 5.6% relative increase.

For colonoscopy procedures, the results in column (5) show that becoming vertically integrated leads to a 6.70 percentage points increase in the probability of referring to a hospital-based facility, which translates to a 14.6% relative increase in the average hospital-based volume share. The average marginal effect reduces to 6.59 percentage points when we control for market characteristics. When the model includes the interaction between OOP payment and complication rate, the effect of vertical integration is 6.63 percentage points. In our final specification that allows for the interaction between surgeon characteristics and PCP group integration status, we estimate a 7.1 percentage point increase in HOPD use, which translates to a 14.4% relative increase.

3.5.1 Counterfactual analysis

Our estimates show that vertically integrated groups have a significantly higher probability of referring patients to hospital-based facilities rather than free-standing (ASC) facilities. In this section, we compare the observed scenario to two counterfactual market structures—a fully vertically integrated market and a fully independent market. By comparing with the two counterfactual scenarios, we estimate the welfare effects of vertical integration and the effects on non-monetary dimensions. To do so, we change the integration indicator to either one or zero for all PCP groups to reflect the counterfactual settings. Based on the estimates obtained from the choice model estimation, we predict patients' choice probability and assess the effect of vertical integration on the volume, the payment, traveling distances, and quality. The results are reported in the table 3.5.

Starting with the changes in site-of-care volume, we estimate that for arthroscopy, when all the PCP groups are vertically integrated, an estimated 16.7 thousand more visits to hospital-based facilities will occur–an 11.4% increase in volume. However, this estimate does not mean that all the hospital-based facilities will receive more patients. If changing to the other scenario, where all the PCP groups are independent, then the total volume of hospital-based visits will decrease by 8.8 thousand visits. The changes for colonoscopy are similar but with a larger magnitude. The number of visits at hospital-based facilities will increase by 341.9 thousand (15.8%) if all PCP groups are integrated and decrease by 195.6 thousand if all PCP groups are independent.

Those changes in volume lead to changes of millions of dollars in Medicare payments. To assess the effect on payments, we assume that the average level of payment for each surgeon remains the same in the counterfactual settings. This assumption is likely valid because PCP group integration has little influence, if any, on the Medicare fee schedule for downstream providers. Our estimates imply that if all PCP groups are vertically integrated, then patients payments will increase by \$4.7 million for joint arthroscopy procedures and by \$35.5 million for colonoscopy procedures. Converting to the total payment by Medicare and beneficiaries, the estimated increase in payments is \$23.5 million for arthroscopy and \$177.7 for colonoscopy procedures. In the other counterfactual setting, we estimate savings of \$75.4 million for the two procedures together.

Similar to the assumptions for Medicare and patient payments, we assume that locations and the average complication rates for a surgeon are unchanged. This assumption could be strong for long-run analysis, but it is unlikely that the vertical integration induces changes in provider locations or provider quality in the short-run. Based on this assumption, we calculate the average change in traveling distance and that in complication rates. The change in distance is not significant for patients seeking joint arthroscopy. This result is expected because traveling is really costly for patients who need joint arthroscopy. The effect on distance is larger for colonoscopy patients; however, the magnitude is still not significant economically.

3.6 Vertical integration and market concentration

As both our reduced form and choice model results find, vertical integration of primary care providers impacts the market share of downstream surgical providers. When a PCP group is vertically integrated with a hospital, more patients will be referred to the owning hospital and, thus, the owning hospital will gain market share. However, the extent to which this change in ”upstream” vertical market structure impacts the ”downstream” market structure of horizontal competition between outpatient surgical providers (e.g., ASCs and HOPDs) is unclear. At the same time, vertical integration can lead to changes in the horizontal market structure of PCP groups for two reasons. First, mechanically, consolidation of PCP groups into systems will reduce the number of PCP group owners. Second, vertical integration could lead to more investment in integrated PCP groups, making integrated PCP groups more attractive to patients. The extent to which vertical integration leads to horizontal consolidation among upstream providers is also not clear.

In this section, we examine the relationship between vertical integration and market concentration. For each market, defined at the HRR-year level, we measure market concentration using the Herfindahl-Hirschman Index (HHI), a widely used measure for market concentration. We measure HHIs for both upstream PCP groups and downstream outpatient surgery providers. The mean HHIs of PCP group markets are 925.7 and 924.2 in the arthroscopy sample and colonoscopy sample, respectively. The HHI of the arthroscopy provider market is 2025.8 and the HHI of the colonoscopy provider market is 1831.1. According to the standard used by the DOJ and FTC, the PCP market is not concentrated, while the arthroscopy and colonoscopy surgeon markets are moderately concentrated.[6]

With these definitions, we estimate the effect of market-level vertical integration on each HHI measure. We define market-level vertical integration as the share of physicians working at a vertically-integrated PCP group, VI_{gt}. We estimate

$$HHI_{gt} = \theta VI_{gt} + \delta_t + \gamma_g + \varepsilon_{gt} \tag{3.7}$$

Because the HHIs are highly skewed, we use log of HHI as the dependent variables. We estimate separate regressions for PCP group HHI (using both procedure samples) and HHIs for arthroscopy and colonoscopy procedures. We cluster standard errors at the HRR level.

Table 4 presents the relation between market-level vertical integration and the HHI in PCP markets and surgeon markets. In the columns with odd numbers, the estimates drop the year and HRR fixed effects. These estimates suggest that when the upstream vertical integration rate is 10 percentage points higher, the HHI of the upstream PCP market will increase around one-fifth, and that of the downstream surgeon market will increase around one-twentieth. For the columns with even numbers, the estimation relies on within-market variation. The limited variation yields limited results. For PCP markets, in both the arthroscopy and colonoscopy samples,

[6]The DOJ and FTC generally consider markets in which the HHI is between 1,500 and 2,500 points to be moderately concentrated, and markets in which the HHI is in excess of 2,500 points to be highly concentrated.

vertical integration still significantly affects market concentration. An increase of 10 percentage points in the integration rate is associated with a 2.1% increase in the market HHI for PCP groups. However, for the downstream provider market, the effect of integration is not statistically significant and is modest in magnitude.

Based on our choice model, we extend the counterfactual analysis to examining the effect of the integration on the market concentration. Using the estimated parameters, we simulated the surgeon market in the two counterfactual scenarios set up in the previous section—all the PCP groups are integrated, and all are independent, and compute the HHI of each market. Figure 3·3 shows the distribution of changes in surgeon market HHIs following the (dis-)integration. If all the PCP groups are integrated, as presented by the red contoured bars, most of the markets will be more concentrated for both joint arthroscopy and colonoscopy. Like the estimates from the above regressions, the average magnitude of changes is small—7.61 (SD 13.53) and 38.56 (SD 57.44) points increase for joint arthroscopy and colonoscopy, respectively. Moving to the other end—all the PCP groups are independent, the changes are even smaller. On average, the arthroscopy market and the colonoscopy market will be 2.45 (SD 9.83) and 7.6 (SD 31.88) points less concentrated, respectively.

3.7 Discussion and conclusion

The market environment for U.S. physicians is rapidly changing, as a growing share of physician practices have vertically integrated with hospitals and health systems. Changes in physician group ownership have led to concerns about how changes in physician financial incentives impact patient care. This concern is particularly noteworthy in the market for outpatient surgeries for Medicare patients. Due to the large difference in payments based on the site-of-care where a procedure is performed, vertical integration can create incentives to perform procedures in Hospital Outpatient

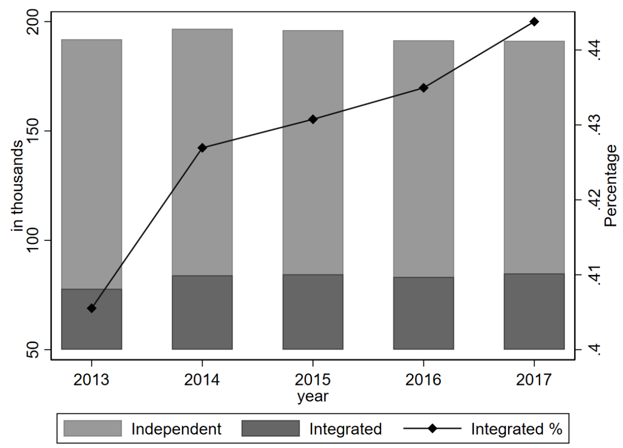
Departments rather than in Ambulatory Surgical Centers. This paper shows that vertical integration of primary care physician groups leads to allocation changes in outpatient care delivery settings. It increases the probability of referring patients to hospital-based facilities. These allocation changes, in turn, lead to increased spending and increased travel time for patients. Importantly, we did not find any quality improvement brought by the vertical integration for the two procedures. Future studies are needed to draw a broader conclusion for the total utilization and the overall health outcomes. However, the lack of a measurable change in quality suggests that efficiency gains from vertical integration may not materialize to patients.

This paper is not without limitations. For one, we restrict our analysis to the Medicare fee-for-service population. Medicare sets prices administratively, which is in stark contrast to the negotiated-prices system among private insurers. Other work has shown that vertical integration leads to an increase in prices for private insurers (Baker, Bundorf and Kessler, 2014). Thus, it is likely that the changes in procedure prices and spending are under-estimated in the Medicare system, relative to the private insurance system. In addition, we only examined two common outpatient procedures. However, the same site-of-care payment differentials exist for other outpatient surgeries and procedures like diagnostic imaging and laboratory tests. We also did not distinguish between within-system referral and cross-system referral among referrals to hospital-based facilities. Finally, we do not examine how increased Medicare payments are allocated between physician groups and hospitals.

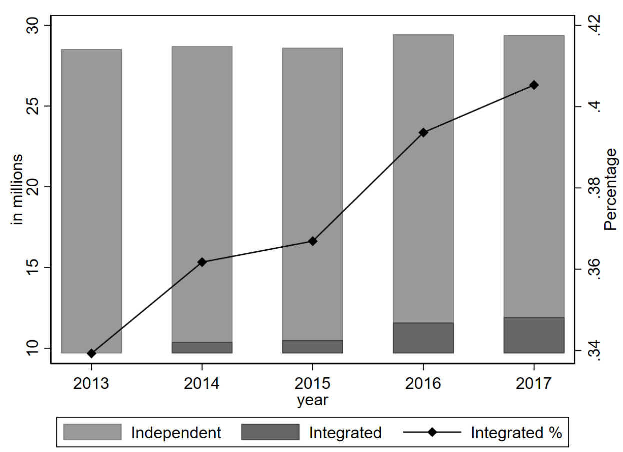
Despite these limitations, this paper's results indicate how payment incentives can impact the allocation of patients to providers. These results have several important policy implications. First, the extent to which Medicare and other insurers site-of-care differentials contribute to the incentives for health care systems to vertically consolidate suggests that reducing the "arbitrage" opportunities in payment models could limit the impacts of vertical integration. Second, the extent to which changes in patient allocation across providers represent allocation inefficiencies, such as requiring patients to travel further or increases in patient cost-sharing payments, suggests that more regulatory oversight of vertical consolidation may be warranted. While the Department of Justice and the Federal Trade Commission have provided guidelines for monitoring vertical consolidation, the impacts of regulatory activity on vertical consolidation of health care providers are uncertain (DoJ and FTC, 2020).

Figure 3·1: Trend of vertical integration

(a) Number of PCPs



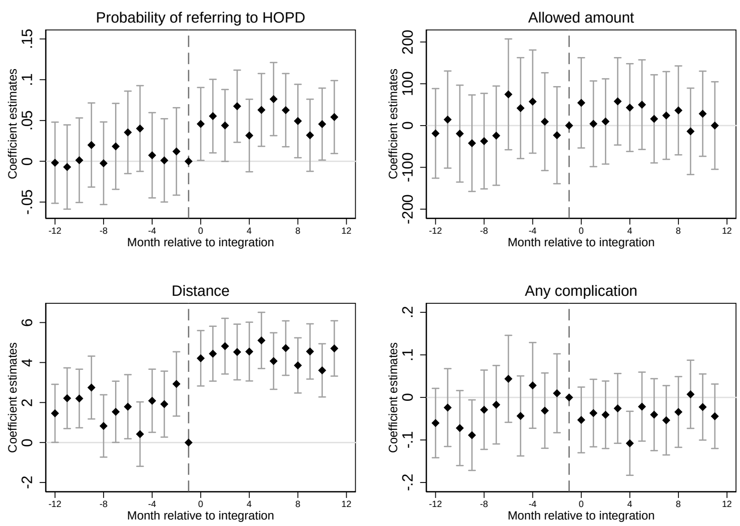
(b) Number of beneficiaries



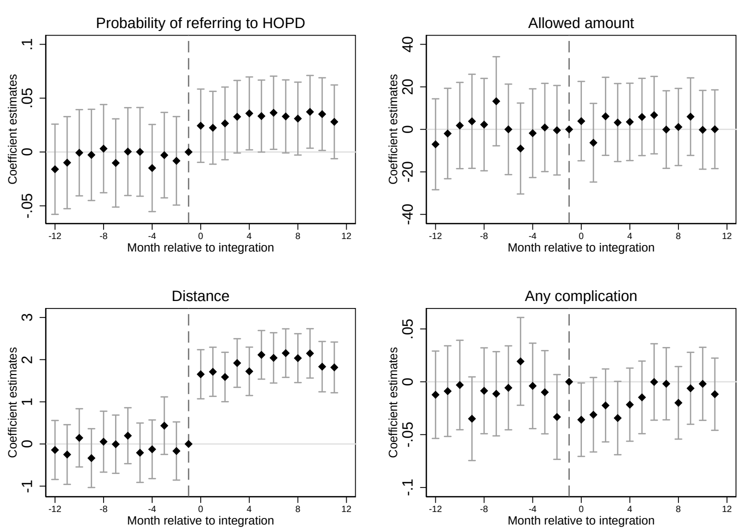
Notes: The two subplots show the trend of vertical integration from 2013 to 2017. The sample is the 100% Medicare Fee-For-Service sample. For the top figure, the integration is measured by the number of physicians working at integrated PCP groups. For the bottom figure, the integration is measured by the number of Medicare beneficiaries whose assigned PCP groups are integrated.

Figure 3·2: Event study figures

(a) Joint arthroscopy



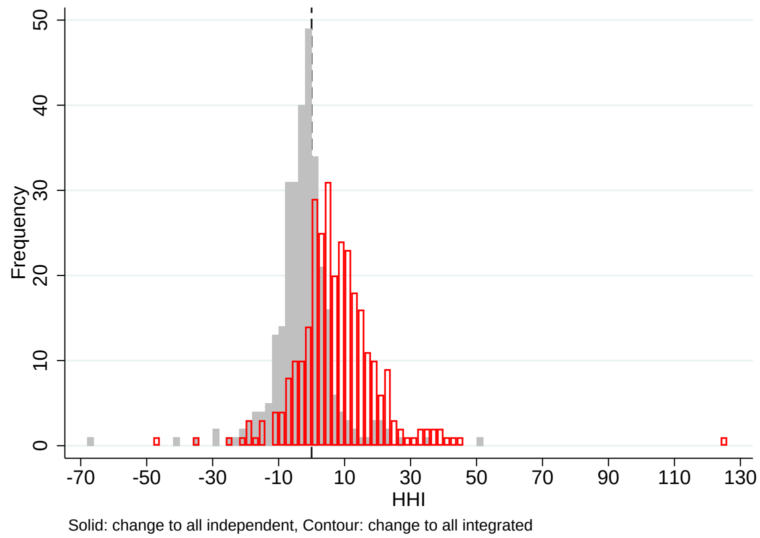
(b) Colonoscopy



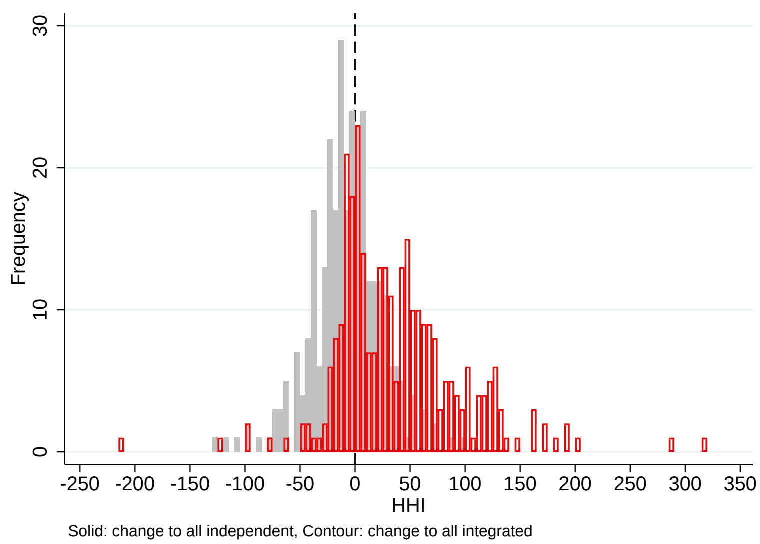
Notes: The eight subplots show the estimates and 95% confidence intervals from the event study regressions. The top and bottom panel correspond to arthroscopy and colonoscopy procedures, respectively. The titles indicate the dependent variables. In all the regressions, we control for procedure fixed effects, year fixed effects, and group fixed effects.

Figure 3·3: Changes in the market concentration following change of integration status

(a) Joint arthroscopy



(b) Colonoscopy



Notes: The two histograms show the distribution of HHI changes following a full dis-integration (grey solid) and a full integration (red contour) for joint arthroscopy and colonoscopy, respectively. The unit of observation is market.

Table 3.1: Summary statistics of reduced-form samples and choice model samples

	Arthroscopy		Colonoscopy	
	Mean	Std. Dev.	Mean	Std. Dev.
A. PCP group sample				
Integrated	0.272	0.445	0.167	0.373
Hospital-Based Share	0.577	0.478	0.492	0.471
Allowed amount ($)	1290.57	1249.93	587.29	310.6
OOP ($)	406.42	693.54	158.8	197.85
Distance (miles)	16.509	16.65	12.772	13.148
Any complication	0.002	0.966	0.075	0.998
N observations	773,374		2,659,752	
B. Individual choice sample				
Hospital-based facility	0.56	0.5	0.46	0.5
Avg. OOP payment	380.66	188.98	170.24	85.85
Complication rate	0.05	0.13	0.23	0.17
Distance	33.2	21.13	33.51	20.72
Age	71.75	4.95	72.96	5.55
Female	0.53	0.5	0.53	0.5
PCP group size	156.52	299.35	169.57	341.15
PCP integrated	0.39	0.49	0.39	0.49
Market avg N PCP	577.51	1043.43	699.27	1220.46
Market avg N MD	2253.93	4386.26	2850.51	5234.84
Market avg income	47475.56	14585.11	48460.72	15115.83
N Obs	1,366,823		13,193,080	

Notes: The top panel shows the summary statistics of the two reduced-form samples. The unit of observation is PCP group-state-month-procedure. The bottom panel shows the summary statistic of the 25% random sample of the choice model samples. This sample includes the full choice set for each beneficiary.

Table 3.2: Relation between vertical integration and group outcomes

	HOPD vol. share	Allowed amount	OOP	Distance	Complication
Arthroscopy (N=773,374)					
	(1)	(2)	(3)	(4)	(5)
Integrated	0.0366***	6.458	15.15*	1.623***	-0.000984
	(0.0141)	(15.13)	(8.859)	(0.390)	(0.00151)
R-squared	0.360	0.360	0.347	0.353	0.122
Colonoscopy (N=2,659,752)					
	(6)	(7)	(8)	(9)	(10)
Integrated	0.0369*	7.140	9.454***	1.148***	-0.000224
	(0.0189)	(5.736)	(2.876)	(0.265)	(0.00159)
R-squared	0.531	0.300	0.470	0.318	0.097

Notes: The eight columns report estimates from eight different regressions. The column heads show the dependent variables for each column. The top panel focus on arthroscopy while the bottom panel focus on colonoscopy. All the regressions include procedure code fixed effects, year and month fixed effects, and provider group fixed effects. Regressions are weighted by volume. The standard errors are clustered by provider group.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3.3: Nested-logit regression results for provider choice

Arthroscopy (N = 1,366,823)				
	(1)	(2)	(3)	(4)
Integrated	0.136***	0.136***	0.135***	0.156***
	(0.0218)	(0.0221)	(0.0218)	(0.0333)
Market characteristics		X	X	X
OOP X complication			X	
Surgeon characteristics X integrated				X
AME (VI)	0.0284	0.0282	0.0281	0.0324
	(0.0072)	(0.00705)	(0.00714)	(0.00768)
Colonoscopy (N =13,193,080)				
	(5)	(6)	(7)	(8)
Integrated	0.310***	0.307***	0.308***	0.330***
	(0.0171)	(0.0157)	(0.0156)	(0.0177)
Market characteristics		X	X	X
OOP X complication			X	
Surgeon characteristics X integrated				X
AME (VI)	0.0670	0.0659	0.0663	0.0708
	(0.0016)	(0.00158)	(0.00202)	(0.00139)

Notes: The eight columns report the estimates and the average marginal effect (AME) from eight different regressions. The four columns in the top panel focus on arthroscopy claims and the four in the bottom panel focus on colonoscopy claims. All the regressions control for patient characteristics (age and gender), PCP group size, surgeon characteristics (average OOP and complication rate), travelling distance, and year fixed effects. Column (2)-(5), and (6)-(8) also include market characteristics (per capita income, and number of non-federal PCPs). Column (3) and (7) allows for the interaction between OOP payment and complication rate. Column (4) and (8) allows for the interaction between surgeon characteristics and PCP groups' integration status. Standard errors of estimates are clustered by year. The standard errors of AME are calculated by the Delta method.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3.4: Relation between horizontal market concentration and vertical integration

Arthroscopy (N = 1,530)				
	PCP market HHI (mean = 925.73)		Surgeon market HHI (mean = 2025.78)	
	(1)	(2)	(3)	(4)
PCP integration rate	2.005***	0.212**	0.484***	-0.0496
	(0.103)	(0.0893)	(0.0955)	(0.0681)
R-squared	0.198	0.963	0.016	0.950
Year FE		Yes		Yes
HRR FE		Yes		Yes
Cluster SE		HRR		HRR
Colonoscopy (N = 1,530)				
	PCP market HHI (mean = 924.21)		Surgeon market HHI (mean = 1831.09)	
	(5)	(6)	(7)	(8)
PCP integration rate	2.216***	0.213**	0.503***	0.0268
	(0.157)	(0.101)	(0.139)	(0.0997)
R-squared	0.116	0.977	0.008	0.967
Year FE		Yes		Yes
HRR FE		Yes		Yes
Cluster SE		HRR		HRR

Notes: The table shows the relation between vertical integration percentage in PCP market and the Herfindahl Hirschman Index (HHI) in the PCP market and the surgeon market. The observations unit is Hospital Referral Regions (HRR)-year. The HHIs are calculated based on number of claims. The dependent variables in the regressions are the log of the HHIs. The top and bottom panel focus on arthroscopy and colonoscopy claims, respectively. Columns of odd number show the raw correlation between the percentage of integrated physicians and the market HHIs. In columns of even number, we control for year fixed effects and HRR fixed effects. The standard errors are clustered by HRR.

Table 3.5: Effects of vertical integration on patient welfare

	Arthroscopy		Colonoscopy	
	All integrated	All independent	All integrated	All independent
	(1)	(2)	(3)	(4)
N decisions (million)	3.15		7.59	
Δ HOPD volume (thousand)	16.71	-8.84	341.89	-195.58
Δ OOP (million)	4.7	-1.98	35.54	-13.1
Δ Medicare total payment (million)	23.48	-9.91	177.71	-65.52
Δ Distance per patient	0.00941	-0.00129	0.15328	-0.04725
Δ Complication rate (z-score)	0.00025	0.00004	0.00819	-0.00316

Notes: The table shows the welfare analysis based on the choice model estimation

Chapter 3 Appendices

A Construction of Vertical Integration Measures

Attribution of Physicians and Physician Groups to Hospitals and Health Systems

To measure vertical integration, we linked physician groups to hospitals and health systems. This process involved several data sources and steps.

1. First, we used data from the Medicare Data on Provider Practice and Specialty (MD-PPAS) data to physician-level National Provider Identifiers (NPIs) to organization-level Tax Identification Numbers (TINs). This mapping constructed our preliminary physician organizations.

2. Next, we combined TINs for academic physician organizations into a single entity.

3. We then combined TINs for non-academic physician organizations by evaluating how frequently two TINs were listed by the same physician in MD-PPAS. If two TINs had a sufficient number of physicians in common and/or had similar enough names in MD-PPAS, we considered the TINs to belong to the same physician organization. After an initial round of combining TINs into physician organizations, a second round of physician organizations were created considering the physician overlap between a TIN and the physician organizations from the first round.

4. After mapping physicians to groups, we mapped groups to hospitals and health systems using the Provider Enrollment and Chain Ownership System (PECOS) data. We supplemented this data with data from IRS Form 990 reports for non-profit hospitals.

Attribution of Medicare Beneficiaries to Physician Groups

After linking physician groups to hospital and health systems, we linked Medicare beneficiaries to physician groups. We developed eight attribution rules that consider patient visits and expenditures. Out of a total Medicare fee-for-service (FFS) population of 38.4 million, our mapping successfully matched 26,442,987 beneficiaries to a physician group. Of the 38.4 million beneficiaries during the 2013 to 2016 time period, 12.0 million did not have an evaluation and management visit with a primary care provider (step 4) and were thus not mapped to a primary care provider in their respective year.

Our mapping uses the following steps:

1. Assign beneficiaries to the physician group where the beneficiary receives the most evaluation and management visits (CPT codes 99201- 99499) from a primary care provider. Primary care providers were defined as general internal medicine, family medicine, and geriatrics specialties as well as nurse practitioners and physician assistants (88.1% of attributed beneficiaries).

1. a. In cases of ties, we only considered the three physician specialties. b. For additional ties, we looked at all evaluation and management visits performed by any type of provider. c. For unresolved ties, we calculated total standardized spending on evaluation and management services to any type of provider and used the provider with the largest spending amount for a beneficiary's care. 2. For beneficiaries with no evaluation and management visits from a primary care provider, we assigned the beneficiary to the group that performed the most evaluation and management visits among internal medicine subspecialists (2.2% of attributed beneficiaries). a. In cases of ties, we used all evaluation and management visits performed by any type of provider (0.3% of attributed beneficiaries). b. For additional ties, we used standardized evaluation and management spending attributable to any type of provider (0.7%

of attributed beneficiaries). 3. If steps 1 and 2 result in ties, we use the location of both the annual Medicare Wellness visit and the last visit (1.2% of attributed beneficiaries). 4. We did not attribute beneficiaries with no evaluation and management services to primary care providers or internal medicine subspecialists.

B Additional Tables and Figures

Table A1: ICD Codes for Complication Identification following joint arthroscopy procedures

Panel A. ICD-9	
30 days	
Bleeding	998.1, 719.10, 719.16, 719.17, 39.98
Post-operative deep vein thrombosis	453.40–453.42, 453.50–453.52,453
Pulmonary embolism	415.1
90 days	
Mechanical failure	996.40, 996.4, 996.49
Wound infection	682.1–682.9, 686.9, 998.6, 998.7, 998.83, 998.3,998.5, 996.66, 996.67, 86.22, 86.28, 86.04, 81.53,81.55, 81.59, 00.70, 00.71, 00.72, 00.73, 00.80, 00.81, 00.82, 00.84,80.05, 80.06, 80.09
Postoperative nerve injury	955, 956, 957.8, 957.9

Panel B. ICD-10	
7 days	
Cardiovascular Complications	I2109, I2119, I2111, I2129, I214 , I213 , I219, I21A1 , I21A9, I973, I20, I240, I248,
Pneumonia & Influenza	J13, J181, J120, J121, J122, J1281, J1289, J129, J150, J151, J14, J154, J153, J1520, J15211, J15212, J1529, J158, J155, J156, A481, J159, J157, J180, J189, J1100, J09X1, J1008, J690,
Shock	R571, R578, R6521, T8110A,
Sepsis	A409, A412, A4101, A4102, A411, A403, A414, A415, A413, A4151, A4152, A4153, A4159, A4189, A419 , R6521, R6520, R7881
30 days	
Bleeding complications	D7801, D7821, D7822, E3601, E3602, G9731, G9732, G9751, G9752, H59111, H59112, H59113, H59119, H59121, H59122, H59123, H59129, H59311, H59312, H59313, H59319, H59321, H59322, H59323, H59329 , H9521, H9522, H9541, H9542, I97410 , I97411, I97418, I9742, I97610, I97611, I97618, I9762, J9561, J9562, J95830, J95831, K9161, K9162, K91840, K91841, L7601, L7602, L7621, L7622, M96810, M96811, M96830 , M96831, N9961, N9962, N99820, N99821, D7831, G9762, H59341, H59342, H59343, H59349, H9551, H9552, I97621, L7602, L7632, M96840, M96841, N99841, G9764, I97622, L7634, M96842, M96843, N99843, T888XXA
Hemarthrosis	M2500, M25069, M25061, M25062, M25011, M25012, M25019, M25073, M25076
Control Bleeding	0W3Q3ZZ, 0W3Q4ZZ, 0W3Q7ZZ, 0W3Q8ZZ, 0X320ZZ, 0X323ZZ, 0X324ZZ, 0X330ZZ, 0X333ZZ, 0X334ZZ, 0X340ZZ, 0X343ZZ, 0X344ZZ, 0X350ZZ, 0X353ZZ, 0X354ZZ, 0X360ZZ, 0X363ZZ, 0X364ZZ, 0X370ZZ, 0X373ZZ, 0X374ZZ, 0X380ZZ, 0X383ZZ, 0X384ZZ, 0X390ZZ, 0X393ZZ, 0X394ZZ, 0Y390ZZ, 0Y393ZZ, 0Y394ZZ, 0Y3B0ZZ, 0Y3B3ZZ, 0Y3B4ZZ, 0Y3C0ZZ, 0Y3C3ZZ, 0Y3C4ZZ, 0Y3D0ZZ, 0Y3D3ZZ, 0Y3D4ZZ, 0Y3F0ZZ, 0Y3F3ZZ, 0Y3F4ZZ, 0Y3G0ZZ, 0Y3G3ZZ, 0Y3G4ZZ, 0Y3H0ZZ, 0Y3H3ZZ, 0Y3H4ZZ, 0Y3J0ZZ, 0Y3J3ZZ, 0Y3J4ZZ
Post-operative DVT/PE	I742, I743, I8010, I80209, I803, I808, I809, I82220, I82290, I823, I82479, I82499, I82609, I82629, I82890, I82A19, I82B19, I82C19, I2690, I2692, I2699, T800XXA, T81718A, T8171XA , T8172XA, I82409, I82419, I82429, I82439, I824Y9, I82449, I82499, I824Z9, I82509 , I82549, I8291
ABO incompatibility	T8030XA, T80311A, T8039XA
Pulmonary embolism	I2690 , I2699, T800XXA, T81718A, T8172XA, I2692
90 days	
Mechanical Complications	T84498A, T84039A, T84029A, T84019A, M979XXA, M9711XA, M9712XA, T84033A, T84032A, T84059A, T84069A, T84099A, T84119A, T84129A, T84199A,
Cellulitis & Infection	L03221 , L03319, L03119, L03129, L03317, L03811, L03818, L0390, L0391, L089, T8183XA, T8169XA, T8189XA
Wound Disruption	T8130XA, T8132XA, T8131XA, T8133XA, K6811, T8450XA, T8460XA, T847XXA
Postoperative nerve injury	S4430XA, S4410XA, S4400XA, S4420XA, S4440XA, S4450XA, S6430XA, S448X9A, S4490XA, S7400XA, S7410XA, S8400XA, S8410XA, S7420XA, S8420XA, S84809A, S84809A, S8490XA, S149XXA , S149XXA

Table A2: ICD Codes for Complication Identification following colonoscopy procedures

ICD-9	
arrhythmia	427.0–427.4, 427.6–427.9
congestive heart failure	428.0–428.9
cardiac or respiratory arrest	427.5, 799.1, 997.1
syncope, hypotension, or shock	453.29, 458.8–458.9, 639.5,780.2, 785.50–785.51, 998.0, 995.4
perforation	569.83, 998.2
Lower gastrointestinal bleeding	558.9, 578.1, 995.2, 995.89,998.1–998.13, 286.5, 459, 562.02–562.03, 562.12, 562.13, 569.3,569.84–569.86, 578.9, 792.1
infection	780.66,790.7, 424.9–424.99
Paralytic ileus	560.1
Nausea, vomiting, dehydration	276.5, 536.2, 787.0-02
Abdominal pain	789
Diverticulitis	562.01, 562.03, 562.11, 562.13
Enterocolitis	555–556
ICD-10	
Arrhythmia	I471, I472, I479, I4891, I4892, I4901, I4902
Acute myocardial infarction	I2109, I2119, I2111, I2129, I214, I213, I219, I21A1, I21A9, I495, R001, I498, I499
Congestive heart failure	I50814, I509, I501, I5020, I5021, I5023, I5030, I5031, I5033, I5040, I5041, I5043, I50810, I50811, I50813, I5082, I5083, I5084, I5089, I509, I110, I130, I132, I255, I420, I425, I426, I427, I428, I429, I43x, I469, R092, I9788, I9789
Syncope, hypotension, or shock	45329, I9589, I959, R55, T882XXA, R579, R570, T8110XA, T81, T811, T8110, T8110XA, T8110XD, T8110XS, T8111, T8111XA, T8111XD, T8140XS, T8140, T8112XA, T8112XD, T8112XS, T8119, T8119XA, T8119XD, T8119XS
Disruption of wound, including perforation	T813, T8130, T8130XA, T8130XD, T8130XS, T8131, T8131XA, T8131XD, T8131XS, T8132, T8132XA, T8132XD, T8132XS, T8133, T8133XA, T8133XD, T8133XS, T814, T8140, T81 40XA, T8140XD, T8140XS, T815, T8150, T81504, T81504A, T81504D, T81504S, T81508, T81508A, T81508D, T81508S, T81509, T81509A, T81509D, T81509S, T8151, T81510, T81514, T81514A, T81514D, T81514S, T81518, T81518A, T81518D, T81518S, T81519, T81519A, T81519D, T81519S, T81524, T81524A, T81524D, T81524S, T81528, T81528A, T81528D, T81528S, T81529, T81529A, T81529D, T81529S, T8153, T81532, T81533, T81534, T81534, T81534D, T81534S, T81538, T81538D, T81538S, T81539, T81539A, T81539D, T81539S, T8159, T81590, T81594, T81594A, T81594D, T81594S, T81595, T81596, T81597, T81598, T81599, T816, T8160, T8161, T8169, T817, T8171, T81710, T81711, T81718, T81718A, T81718D, T81718S, T81719, T8172 , T8172 A, T8172 D, T8172 S, T818, T8181, T8182, T8183, T8189, T819
Perforation	K631, E3611, E3612, G9749, I9752, J9572, K9171, K9172, L7612, T888XXA,
Lower gastrointestinal bleeding	K921, T50905A, T8851XA, E3601, E3602, E89810, E89811, E89820, E89821, G9732, G9752, G9762, I9742, I97620, I97621, J9562, J95831, J95861, K9161, K9162, K91840, K91841, L7602, L7622, L7632, N9962, N99821, N99841, E89822, E89823, G9764, I97622, J95863, K91872, K91873, L7634, N99843, T888XXA,
Bleeding Complications	R58, K625, K5521, K6381, K922, R5084, J690, J698, J158, J159, 4838, J168, J189, R7881, I38, I39,
Ileus	K560, K567 ,
Nausea, vomiting, dehydration	E869, E860, E861, R1110, R112, R110,
Abdominal pain	R109,
Diverticulitis	K5712, K5713, K5732, K5733,
Central Nervous System Event	
Hemorrhage	I609, I619, I621, I6200, I629
Cerebral infarction	I6330, I6340, I6350,
Occlusion and stenosis	I669, I6609, I6619, I6629,
Other cerebrovascular diseases	I6789
Pulmonary embolism	I260, I2601, I2602, I2609, I269, I2690, I2692, I2693, I2694, I2699

References

Abaluck, Jason, Mauricio M. Caceres Bravo, Peter Hull, and Amanda Starc. 2020. "Mortality Effects and Choice Across Private Health Insurance Plans." National Bureau of Economic Research Working Paper 27578.

Abraham, Jean, Coleman Drake, Daniel W. Sacks, and Kosali Simon. 2017. "Demand for health insurance marketplace plans was highly elastic in 2014–2015." *Economics Letters*, 159: 69–73.

Abraham, Jean M, Pinar Karaca-Mandic, and Kosali Simon. 2014. "How has the Affordable Care Act's medical loss ratio regulation affected insurer behavior?" *Medical Care*, 370–377.

Acemoglu, Daron, and Amy Finkelstein. 2008. "Input and technology choices in regulated industries: Evidence from the health care sector." *Journal of Political Economy*, 116(5): 837–880.

Aghadadashli, Hamid, Markus Dertwinkel-Kalt, and Christian Wey. 2016. "The Nash bargaining solution in vertical relations with linear input prices." *Economics Letters*, 145: 291–294.

Agha, Leila, Brigham Frandsen, and James B Rebitzer. 2019. "Fragmented division of labor and healthcare costs: Evidence from moves across regions." *Journal of Public Economics*, 169: 144–159.

Agha, Leila, Keith Marzilli Ericson, Kimberley H Geissler, and James B Rebitzer. 2018. "Team Formation and Performance: Evidence from Healthcare Referral Networks." National Bureau of Economic Research Working Paper 24338.

Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2008. "When Does Coordination Require Centralization?" *American Economic Review*, 98(1): 145–179.

Aouad, Marion, Timothy T. Brown, and Christopher M. Whaley. 2019. "Reference pricing: The case of screening colonoscopies." *Journal of Health Economics*, 65: 246–259.

Arrow, Kenneth J. 1963. "Uncertainty and the Welfare Economics of Medical Care." *The American Economic Review*, 53(5): 941–973.

Arrow, Kenneth J. 1975. "Vertical Integration and Communication." *The Bell Journal of Economics*, 6(1): 173–183.

Atalay, Enghin, Ali Hortaçsu, and Chad Syverson. 2014. "Vertical integration and input flows." *American Economic Review*, 104(4): 1120–48.

Atlas, Steven J., Richard W. Grant, Timothy G. Ferris, Yuchiao Chang, and Michael J. Barry. 2009. "Patient-physician connectedness and quality of primary care." *Annals of Internal Medicine*, 150(5): 325–335.

Austin, Daniel R, and Laurence C Baker. 2015. "Less physician practice competition is associated with higher prices paid for common procedures." *Health Affairs*, 34(10): 1753–1760.

Averch, Harvey, and Leland L Johnson. 1962. "Behavior of the firm under regulatory constraint." *The American Economic Review*, 52(5): 1052–1069.

Baker, Laurence C., M. Kate Bundorf, and Anne Royalty. 2016. "Measuring Physician Practice Competition Using Medicare Data." In *Measuring and Modeling Health Care Costs*. , ed. Ana Aizcorbe, Colin Baker, Ernst R. Berndt and David M. Cutler, 351–377. University of Chicago Press.

Baker, Laurence C, M Kate Bundorf, and Daniel P Kessler. 2014. "Vertical integration: hospital ownership of physician practices is associated with higher prices and spending." *Health Affairs*, 33(5): 756–763.

Baker, Laurence C, M Kate Bundorf, and Daniel P Kessler. 2016*a*. "The effect of hospital/physician integration on hospital choice." *Journal of Health Economics*, 50: 1–8.

Baker, Laurence C., M. Kate Bundorf, and Daniel P. Kessler. 2016*b*. "The effect of hospital/physician integration on hospital choice." *Journal of Health Economics*, 50: 1–8.

Baker, Laurence C., M. Kate Bundorf, and Daniel P. Kessler. 2020*a*. "Does Multispecialty Practice Enhance Physician Market Power?" *American Journal of Health Economics*, 6(3): 324–347.

Baker, Laurence C, M Kate Bundorf, and Daniel P Kessler. 2020*b*. "Does Multispecialty Practice Enhance Physician Market Power?" *American Journal of Health Economics*, 6(3).

Baker, Laurence C, M Kate Bundorf, Anne B Royalty, and Zachary Levin. 2014. "Physician practice competition and prices paid by private insurers for office visits." *JAMA*, 312(16): 1653–1662.

Barrette, Eric, Gautam Gowrisankaran, and Robert Town. 2020. "Countervailing Market Power and Hospital Competition." National Bureau of Economic Research.

Beaulieu, Nancy D, Leemore S Dafny, Bruce E Landon, Jesse B Dalton, Ifedayo Kuye, and J Michael McWilliams. 2020. "Changes in Quality of Care after Hospital Mergers and Acquisitions." *New England Journal of Medicine*, 382(1): 51–59.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica*, 63(4): 841–890.

Callaghan, Sandra Renfro, Elizabeth Plummer, and William F Wempe. 2020. "Health Insurers' Claims and Premiums Under the Affordable Care Act: Evidence on the Effects of Bright Line Regulations." *Journal of Risk and Insurance*, 87(1): 67–93.

Capps, Cory, David Dranove, and Christopher Ody. 2018. "The effect of hospital acquisitions of physician practices on prices and spending." *Journal of Health Economics*, 59: 139–152.

CCIIO. 2017. "Summary of 2016 Medical Loss Ratio Results." Center for Consumer Information and Insurance Oversight.

Cebul, Randall D, James B Rebitzer, Lowell J Taylor, and Mark E Votruba. 2008. "Organizational fragmentation and care quality in the US healthcare system." *Journal of Economic Perspectives*, 22(4): 93–113.

Chan, David C, Jr, Matthew Gentzkow, and Chuan Yu. 2019. "Selection with Variation in Diagnostic Skill: Evidence from Radiologists." National Bureau of Economic Research Working Paper 26467.

Chen, Yiqun. 2020. "Team-Specific Human Capital and Team Performance: Evidence from Doctors." *Working Paper*.

Chernew, Michael E. 2021. "Disparities in payment across sites encourage consolidation." *Health Services Research*, 56(1): 5–6.

Chernew, Michael, Zack Cooper, Eugene Larsen-Hallock, and Fiona Scott Morton. 2018. "Are health care services shoppable? Evidence from the consumption of lower-limb MRI scans." National Bureau of Economic Research.

Cicala, Steve, Ethan M. J. Lieber, and Victoria Marone. 2019. "Regulating Markups in US Health Insurance." *American Economic Journal: Applied Economics*, 11(4): 71–104.

Clemens, Jeffrey, and Joshua D. Gottlieb. 2014. "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review*, 104(4): 1320–49.

Coase, Ronald. 1937. "The theory of the firm." *Economica*, 4(16): 386–405.

Cooper, Zack, Stuart V Craig, Martin Gaynor, and John Van Reenen. 2019. "The price ain't right? Hospital prices and health spending on the privately insured." *The Quarterly Journal of Economics*, 134(1): 51–107.

Cox, Cynthia, Rachel Fehr, and Larry Levitt. 2019. "Individual Insurance Market Performance in 2018." Kaiser Family Foundation.

Craig, Stuart V, Keith Marzilli Ericson, and Amanda Starc. 2018. "How Important Is Price Variation Between Health Insurers?" National Bureau of Economic Research.

Crawford, Gregory S, and Ali Yurukoglu. 2012. "The welfare effects of bundling in multichannel television markets." *American Economic Review*, 102(2): 643–85.

Currie, Janet, and W Bentley MacLeod. 2017. "Diagnosing expertise: Human capital, decision making, and performance among physicians." *Journal of labor economics*, 35(1): 1–43.

Currie, Janet, W Bentley MacLeod, and Jessica Van Parys. 2016. "Provider practice style and patient health outcomes: the case of heart attacks." *Journal of Health Economics*, 47: 64–80.

Curto, Vilsa, Liran Einav, Amy Finkelstein, Jonathan Levin, and Jay Bhattacharya. 2019. "Health care spending and utilization in public and private Medicare." *American Economic Journal: Applied Economics*, 11(2): 302–32.

Cutler, David, Jonathan S Skinner, Ariel Dora Stern, and David Wennberg. 2019. "Physician beliefs and patient preferences: a new look at regional variation in health care spending." *American Economic Journal: Economic Policy*, 11(1): 192–221.

Cutler, David M, and Fiona Scott Morton. 2013. "Hospitals, market share, and consolidation." *JAMA*, 310(18): 1964–1970.

Cutler, David M, Leemore Dafny, David C Grabowski, Steven Lee, and Christopher Ody. 2020. "Vertical Integration of Healthcare Providers Increases Self-Referrals and Can Reduce Downstream Competition: The Case of Hospital-Owned Skilled Nursing Facilities." National Bureau of Economic Research Working Paper 28305.

Dafny, Leemore. 2019. “Does It Matter if Your Health Insurer Is For Profit? Effects of Ownership on Premiums, Insurance Coverage, and Medical Spending.” *American Economic Journal: Economic Policy*, 11(1): 222–65.

Dafny, Leemore, Kate Ho, and Robin S. Lee. 2019*a*. “The price effects of cross-market mergers: theory and evidence from the hospital industry.” *The RAND Journal of Economics*, 50(2): 286–325.

Dafny, Leemore, Kate Ho, and Robin S Lee. 2019*b*. “The price effects of cross-market mergers: theory and evidence from the hospital industry.” *The RAND Journal of Economics*, 50(2): 286–325.

Dafny, Leemore, Mark Duggan, and Subramaniam Ramanarayanan. 2012. “Paying a premium on your premium? Consolidation in the US health insurance industry.” *American Economic Review*, 102(2): 1161–85.

Dessein, Wouter. 2014. “Incomplete Contracts and Firm Boundaries: New Directions.” *Journal of Law, Economics, & Organization*, 30: i13–i36.

DoJ, and FTC. 2020. “Vertical Merger Guidelines.” U.S. Department of Justice and the Federal Trade Commission.

Drake, Coleman. 2019. “What are consumers willing to pay for a broad network health plan?: Evidence from covered California.” *Journal of Health Economics*, 65: 63–77.

Duggan, Mark, Amanda Starc, and Boris Vabson. 2016. “Who benefits when the government pays more? Pass-through in the Medicare Advantage program.” *Journal of Public Economics*, 141: 50–67.

Efron, Bradley, and Robert Tibshirani. 1986. “Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy.” *Statistical Science*, 54–75.

Einav, Liran, Amy Finkelstein, and Maria Polyakova. 2018. “Private provision of social insurance: drug-specific price elasticities and cost sharing in Medicare Part D.” *American Economic Journal: Economic Policy*, 10(3): 122–53.

Enthoven, Alain C. 2009. “Integrated delivery systems: the cure for fragmentation.” *American Journal of Managed Care*, 15(12): S284.

Ericson, Keith M Marzilli, and Amanda Starc. 2015. “Pricing regulation and imperfect competition on the massachusetts health insurance exchange.” *Review of Economics and Statistics*, 97(3): 667–682.

Fadlon, Itzik, and Jessica Van Parys. 2020. "Primary care physician practice styles and patient care: Evidence from physician exits in medicare." *Journal of Health Economics*, 71: 102304.

Fehr, Rachel, Cynthia Cox, and Larry Levitt. 2019. "Data Note: Changes in Enrollment in the Individual Health Insurance Market through Early 2019." Kaiser Family Foundation.

Fehr, Rachel, Daniel McDermott, and Cynthia Cox. 2020. "Individual Insurance Market Performance in 2019." Kaiser Family Foundation.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. "Sources of Geographic Variation in Health Care: Evidence From Patient Migration." *The Quarterly Journal of Economics*, 131(4): 1681–1726.

Forbes, Silke J., and Mara Lederman. 2010. "Does vertical integration affect firm performance? Evidence from the airline industry." *The RAND Journal of Economics*, 41(4): 765–790. DOI: 10.1111/j.1756-2171.2010.00120.x.

Forman, Chris, and Anne Gron. 2011. "Vertical Integration and Information Technology Investment in the Insurance Industry." *Journal of Law, Economics, & Organization*, 27(1): 180–218.

Frakes, Michael. 2013. "The impact of medical liability standards on regional variations in physician behavior: Evidence from the adoption of national-standard rules." *American Economic Review*, 103(1): 257–76.

Frandsen, Brigham, Michael Powell, and James B Rebitzer. 2019. "Sticking points: common-agency problems and contracting in the US healthcare system." *The RAND Journal of Economics*, 50(2): 251–285.

Frandsen, Brigham R, Karen E Joynt, James B Rebitzer, and Ashish K Jha. 2015. "Care fragmentation, quality, and costs among chronically ill patients." *Am J Manag Care*, 21(5): 355–362.

Friebel, Guido, and Michael Raith. 2010. "Resource Allocation and Organizational Form." *American Economic Journal: Microeconomics*, 2(2): 1–33.

Gardner, Timothy F., CPT Michael U. Nnadozie, Barbara A. Davis, and Sharon Kirk. 2005. "Patient Anxiety and Patient Satisfaction in Hospital Based and Freestanding Ambulatory Surgery Centers." *Journal of Nursing Care Quality*, 20(3): 238–243.

Gaynor, Martin, Kate Ho, and Robert J Town. 2015. "The industrial organization of health-care markets." *Journal of Economic Literature*, 53(2): 235–84.

Goldberg, Pinelopi Koujianou. 1995. "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry." *Econometrica*, 63(4): 891–951. ArticleType: research-article / Full publication date: Jul., 1995 / Copyright © 1995 The Econometric Society.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town. 2015. "Mergers when prices are negotiated: Evidence from the hospital industry." *American Economic Review*, 105(1): 172–203.

Gowrisankaran, Gautam, Keith A Joiner, and Pierre-Thomas Léger. 2017. "Physician Practice Style and Healthcare Costs: Evidence from Emergency Departments." National Bureau of Economic Research Working Paper 24155.

Grennan, Matthew. 2013. "Price discrimination and bargaining: Empirical evidence from medical devices." *American Economic Review*, 103(1): 145–77.

Grennan, Matthew. 2014. "Bargaining ability and competitive advantage: Empirical evidence from medical devices." *Management Science*, 60(12): 3011–3025.

Griffith, Kevin, David K Jones, and Benjamin D Sommers. 2018. "Diminishing insurance choices in the Affordable Care Act Marketplaces: A county-based analysis." *Health Affairs*, 37(10): 1678–1684.

Grisel, Jedidiah, and Ellis Arjmand. 2009. "Comparing quality at an ambulatory surgery center and a hospital-based facility: Preliminary findings." *Otolaryngology - Head and Neck Surgery*, 141(6): 701–709.

Hart, Oliver, and Bengt Holmstrom. 2010. "A Theory of Firm Scope." *The Quarterly Journal of Economics*, 125(2): 483–513.

Hart, Oliver, and John Moore. 1990. "Property Rights and the Nature of the Firm." *Journal of Political Economy*, 98(6): 1119–1158.

Hendricks, Wallace. 1975. "The effect of regulation on collective bargaining in electric utilities." *The Bell Journal of Economics*, 451–465.

Ho, Kate, and Robin S. Lee. 2017. "Insurer Competition in Health Care Markets." *Econometrica*, 85(2): 379–417.

Ho, Kate, and Robin S Lee. 2019. "Equilibrium provider networks: Bargaining and exclusion in health care markets." *American Economic Review*, 109(2): 473–522.

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

Horn, Henrick, and Asher Wolinsky. 1988. "Bilateral monopolies and incentives for merger." *The RAND Journal of Economics*, 408–419.

Huckman, Robert S, and Gary P Pisano. 2006. "The firm specificity of individual performance: Evidence from cardiac surgery." *Management Science*, 52(4): 473–488.

Hussey, Peter S, Eric C Schneider, Robert S Rudin, D Steven Fox, Julie Lai, and Craig Evan Pollack. 2014. "Continuity and the costs of care for chronic disease." *JAMA internal medicine*, 174(5): 742–748.

Jaffe, Sonia, and Mark Shepard. 2017. "Price-linked subsidies and health insurance markups." *HKS Working Paper No. RWP17-002*.

Kane, CK. 2019. "Policy Research Perspectives. Updated data on physician practice arrangements: for the first time, fewer physicians are owners than employees." *Chicago, IL: American Medical Association*, 2019–07.

Karaca-Mandic, Pinar, Jean M Abraham, and Kosali Simon. 2015. "Is the medical loss ratio a good target measure for regulation in the individual market for health insurance?" *Health Economics*, 24(1): 55–74.

Kim, Song-Hee, Hummy Song, and Melissa Valentine. 2020. "Staffing Temporary Teams: Understanding the Effects of Team Familiarity and Partner Variety." *Available at SSRN 3176306*.

Koch, Thomas G, Brett W Wendling, and Nathan E Wilson. 2017. "How vertical integration affects the quantity and cost of care for Medicare beneficiaries." *Journal of Health Economics*, 52: 19–32.

Koch, Thomas G, Brett W Wendling, and Nathan E Wilson. 2018. "The Effects of Physician and Hospital Integration on Medicare Beneficiaries' Health Outcomes." *Review of Economics and Statistics*, 1–38.

Kwok, Jennifer H. 2019. "How Do Primary Care Physicians Influence Healthcare? Evidence on Practice Styles and Switching Costs from Medicare." *Working Paper*. `https://www.jenniferkwok.com/s/JenniferKwok_PrimaryCarePhysicians_JMP.pdf`.

Lafontaine, Francine, and Margaret Slade. 2007. "Vertical Integration and Firm Boundaries: The Evidence." *Journal of Economic Literature*, 45(3): 629–685.

Laird, Jessica, and Torben Nielsen. 2017. "Physician Prescribing Behaviors on Prescription Drug Use and Labor Supply : Evidence from Movers in Denmark." Working Paper.

Loewenstein, George, Joelle Y Friedman, Barbara McGill, Sarah Ahmad, Suzanne Linck, Stacey Sinkula, John Beshears, James J Choi, Jonathan Kolstad, David Laibson, et al. 2013. "Consumers' misunderstanding of health insurance." *Journal of Health Economics*, 32(5): 850–862.

McCue, Michael J, and Mark Hall. 2015. "Health insurers' financial performance and quality improvement expenditures in the affordable care act's second year." *Medical Care Research and Review*, 72(1): 113–122.

McCue, Michael, Mark Hall, and Xinliang Liu. 2013. "Impact of medical loss regulation on the financial performance of health insurers." *Health Affairs*, 32(9): 1546–1551.

McFadden, Daniel. 1980. "Econometric Models for Probabilistic Choice Among Products." *The Journal of Business*, 53(3): S13–S29. 02586 ArticleType: research-article / Issue Title: Part 2: Interfaces Between Marketing and Economics / Full publication date: Jul., 1980 / Copyright © 1980 The University of Chicago Press.

McGuire, Thomas G. 2000. "Chapter 9 Physician agency." In . Vol. 1, Part A, , ed. BT Handbook of Health Economics, 461–536. Elsevier. 00003.

McWilliams, J. Michael, Laura A. Hatfield, Bruce E. Landon, Pasha Hamed, and Michael E. Chernew. 2018. "Medicare Spending after 3 Years of the Medicare Shared Savings Program." *New England Journal of Medicine*, 379(12): 1139–1149. PMID: 30183495.

Molitor, David. 2018. "The evolution of physician practice styles: evidence from cardiologist migration." *American Economic Journal: Economic Policy*, 10(1): 326–56.

Morris, Carl N. 1983. "Parametric empirical Bayes inference: theory and applications." *Journal of the American Statistical Association*, 78(381): 47–55.

Mullainathan, Sendhil, and David Scharfstein. 2001. "Do Firm Boundaries Matter?" *American Economic Review*, 91(2): 195–199.

Munnich, Elizabeth L., and Stephen T. Parente. 2014. "Procedures Take Less Time At Ambulatory Surgery Centers, Keeping Costs Down And Ability To Meet Demand Up." *Health Affairs*, 33(5): 764–769.

Nevo, Aviv. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica*, 69(2): 307–342.

Nyweide, David J, and Julie PW Bynum. 2017. "Relationship between continuity of ambulatory care and risk of emergency department episodes among older adults." *Annals of Emergency Medicine*, 69(4): 407–415.

Nyweide, David J, Denise L Anthony, Julie PW Bynum, Robert L Strawderman, William B Weeks, Lawrence P Casalino, and Elliott S Fisher. 2013. "Continuity of care and the risk of preventable hospitalization in older adults." *JAMA Internal Medicine*, 173(20): 1879–1885.

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

Parys, Jessica Van. 2018. "ACA marketplace premiums grew more rapidly in areas with monopoly insurers than in areas with more competition." *Health Affairs*, 37(8): 1243–1251.

Peters, Craig. 2014. "Bargaining Power and the Effects of Joint Negotiation: The "Recapture Effect"." Department of Justice.

Pierce, Lamar. 2012. "Organizational Structure and the Limits of Knowledge Sharing: Incentive Conflict and Agency in Car Leasing." *Management Science*, 58(6): 1106–1121.

Pollack, Craig Evan, Peter S Hussey, Robert S Rudin, D Steven Fox, Julie Lai, and Eric C Schneider. 2016. "Measuring care continuity: a comparison of claims-based methods." *Medical care*, 54(5): e30.

Post, Brady, Edward C. Norton, Brent Hollenbeck, Thomas Buchmueller, and Andrew M. Ryan. 2021. "Hospital-physician integration and Medicare's site-based outpatient payments." *Health Services Research*, 56(1): 7–15.

Richards, Michael R., Jonathan Seward, and Christopher M. Whaley. 2020. "Treatment consolidation after vertical integration: Evidence from outpatient procedure markets." Publisher: RAND Corporation.

Roberts, Eric T, Michael E Chernew, and J Michael McWilliams. 2017. "Market share matters: evidence of insurer and provider bargaining over prices." *Health Affairs*, 36(1): 141–148.

Robinson, James C., Timothy Brown, and Christopher Whaley. 2015. "Reference-Based Benefit Design Changes Consumers Choices And Employers Payments For Ambulatory Surgery." *Health Affairs*, 34(3): 415–422.

Sahni, Nikhil R, Maurice Dalton, David M Cutler, John D Birkmeyer, and Amitabh Chandra. 2016. "Surgeon specialization and operative mortality in United States: retrospective analysis." *BMJ*, 354: i3571.

Saltzman, Evan. 2019. "Demand for health insurance: Evidence from the California and Washington ACA exchanges." *Journal of Health Economics*, 63: 197–222.

Scheffler, Richard M., Daniel R. Arnold, and Christopher M. Whaley. 2018. "Consolidation Trends In California's Health Care System: Impacts On ACA Premiums And Outpatient Visit Prices." *Health Affairs*, 37(9): 1409–1416.

Seru, Amit. 2014. "Firm boundaries matter: Evidence from conglomerates and R&D activity." *Journal of Financial Economics*, 111(2): 381–405.

Skinner, Jonathan. 2011. "Causes and Consequences of Regional Variations in Health Care." In *Handbook of Health Economics*. Vol. 2 of *Handbook of Health Economics*, , ed. Mark V. Pauly, Thomas G. Mcguire and Pedro P. Barros, 45–93. Elsevier.

Stein, Jerome L, and George H Borts. 1972. "Behavior of the firm under regulatory constraint." *The American Economic Review*, 62(5): 964–970.

Stroebel, Johannes. 2016. "Asymmetric Information about Collateral Values." *The Journal of Finance*, 71(3): 1071–1112. DOI:10.1111/jofi.12288.

Takayama, Akira. 1969. "Behavior of the firm under regulatory constraint." *The American Economic Review*, 59(3): 255–260.

Tebaldi, Pietro. 2017. "Estimating equilibrium in health insurance exchanges: Price competition and subsidy design under the aca." *Becker Friedman Institute for Research in Economics Working Paper*, , (2017-05).

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

Venkatesh, Arjun K, Hao Mei, Keith E Kocher, Michael Granovsky, Ziad Obermeyer, Erica S Spatz, Craig Rothenberg, Harlan M Krumholz, and Zhenqui Lin. 2017. "Identification of emergency department visits in medicare administrative claims: Approaches and implications." *Academic Emergency Medicine*, 24(4): 422–431.

Welch, W Pete, Alison Evans Cuellar, Sally C Stearns, and Andrew B Bindman. 2013. "Proportion of physicians in large group practices continued to grow in 2009–11." *Health Affairs*, 32(9): 1659–1666.

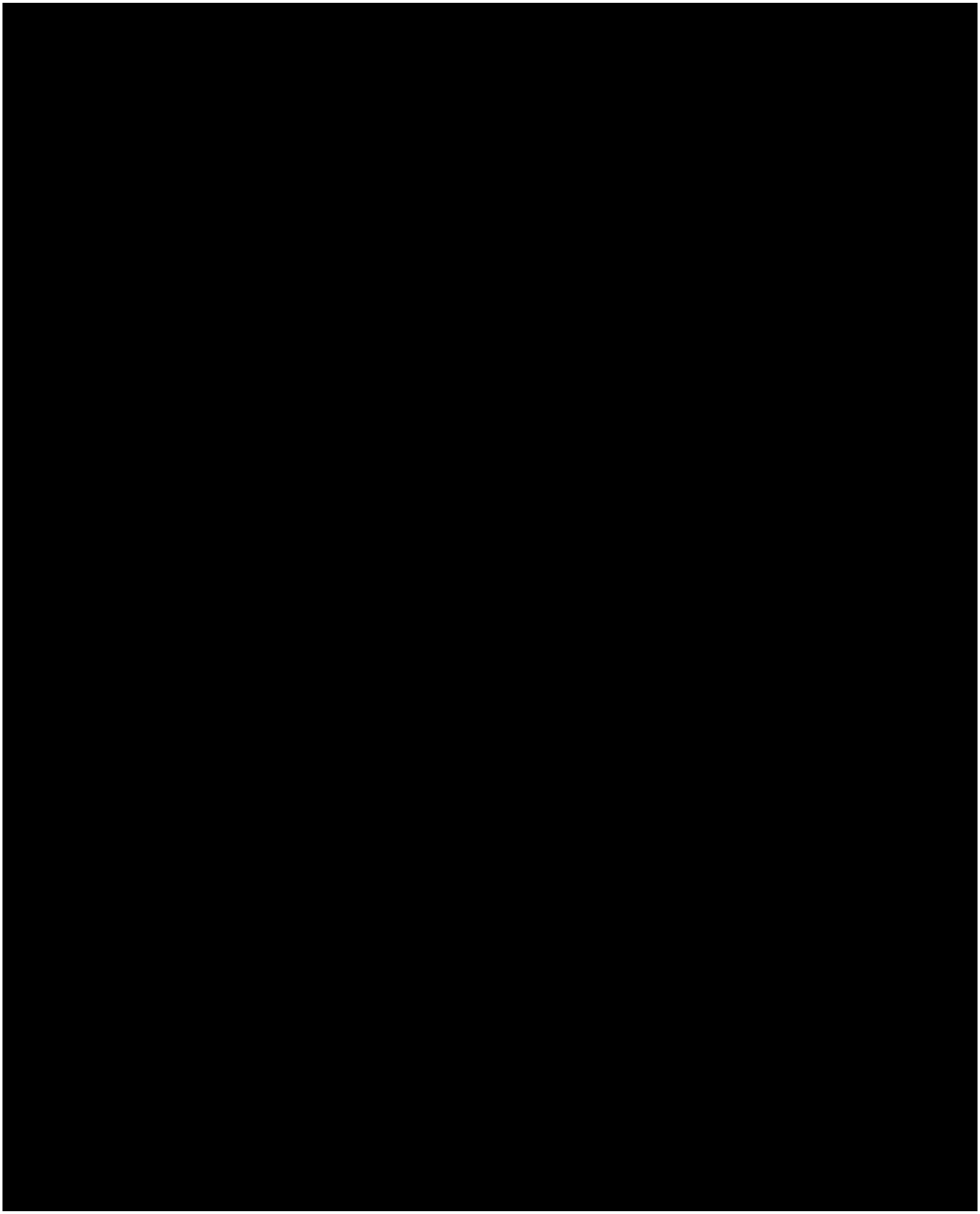
Whaley, Christopher M., and Timothy T. Brown. 2018. "Firm responses to targeted consumer incentives: Evidence from reference pricing for surgical services." *Journal of Health Economics*, 61: 111–133.

Whaley, Christopher M., Chaoran Guo, and Timothy T. Brown. 2017. "The Moral Hazard Effects of Consumer Responses to Targeted Cost-Sharing." *Journal of Health Economics*, 56: 201–221.

Williamson, O.E. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting.* Free Press.

CURRICULUM VITAE

[redacted]



[redacted]