Physician Practice Organization and Negotiated Prices: Evidence from State Law Changes*

Naomi Hausman
Hebrew University

Kurt Lavetti
Ohio State University

February 7, 2018

Abstract

We study the relationship between the organization of physician practices and prices negotiated with private insurers. Using variation from state-level judicial decisions, we show that changes in the enforceability of non-compete agreements (NCAs) in physician employment contracts alter the organization of physician practices and the service prices they charge. The effects of these NCA decisions are economically meaningful: an increase in NCA enforceability of 10% of the observed policy spectrum causes a 3.7% increase in average physician prices. Using two databases containing the universe of physician establishments and firms in the US between 1996 and 2007, linked to prices negotiated with private insurance companies, we show that this price effect is associated with reductions in practice sizes and market concentration that increase prices for services with high practice overhead costs. Using these judicial decisions as instruments, we estimate that a 100 point increase in the establishment-based Herfindahl Index (HHI) causes a 1.3% to 1.7% decline in prices, consistent with insurers extracting efficiency gains from larger establishments. In contrast, the same change in concentration caused by physically-distinct establishments negotiating jointly as a firm leads to price increases of 1.0% to 2.0%.

*We are grateful to Jay Bhattacharya, Jeff Clemens, Leemore Dafny, Michael Dickstein, Will Dow, Alon Eizenberg, Randy Ellis, Josh Gottlieb, Arthur Lewbel, Ian McCarthy, Jesse Rothstein, Ashley Swanson, and seminar participants at the AEA Annual Meeting, ASHE, Berkeley, Chicago Booth Junior Economics Summit, DOJ, Hebrew University, IDC, LSE, MIT, NBER Productivity Lunch, NBER Summer Institute, Northwestern Kellogg, NYU, Stanford, Tel Aviv University, and UGA for helpful comments, to Norman Bishara for sharing legal data, and to Eric Auerbach, Richard Braun, Akina Ikudo, and David Krosin for research assistance. This research was conducted while Lavetti was a Robert Wood Johnson Foundation Scholar in Health Policy at UC Berkeley, and their support is gratefully acknowledged. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. Correspondence: lavetti.1@osu.edu

1 Introduction

At 17.2% of GDP, the share of income devoted to healthcare in the US is over 90% higher than the OECD average.[1] Many studies, including Pauly (1993) and Anderson et al. (2003), have shown that this difference in spending is primarily due to differences in prices rather than quantities, which has led researchers to try to understand why prices are so much higher in the US. Much of this attention has focused on how competition affects prices in health insurance markets (Dafny (2010); Dafny et al. (2012); Ericson and Starc (2012); Ho and Lee (2016)) and in hospital markets (Gowrisankaran et al. (2014); Gaynor and Vogt (2003)). There is relatively less evidence on the determinants of physician prices, even though physician services account for a large and rising share (20%) of total U.S. medical spending.[2]

Research that has addressed the topic suggests that physician organizational structures can have effects on prices that vary by context. Kleiner et al. (2015) and Dunn and Shapiro (2014) find evidence consistent with market power among specialist physicians when they compare concentration levels across markets, while Baker et al. (2014) estimate that moving from perfect competition to monopolist physician practices would only increase prices by about 7%.[3] More recently, Clemens, Gottlieb, and Molnar (2017) show that when Medicare changes reimbursement rates, prices paid to physician practices track the Medicare benchmark for 90% of procedures at small practices, but only 40% of procedures at large practices. The relationship between practice size and negotiated prices is especially strong for capital-intensive procedures, for which average costs are more likely to differ from marginal costs; larger practices tend to receive prices that are adjusted downward relative to Medicare's average-cost-based prices for these procedures. Taken together, these studies suggest that physician organizational structures can have important effects on prices—potentially through cost efficiencies that may counteract the effects of increased bargaining leverage relative to insurers.

This paper provides new evidence on the impacts of physician practice organization and market concentration on prices negotiated with private insurers. Our research design builds upon and extends previous studies in two primary ways. First, we develop the most comprehensive known database to date on physician practices and negotiated prices, including two complete censuses of practices in all specialties and geographic markets in the US between 1996-2007. The Medicare Physician Identification and Eligibility Registry

[1]See OECD Health Statistics 2017

[2]National Health Expenditure Fact Sheet 2013, CMS.

[3]Clemens and Gottlieb (2016) also find evidence consistent with the presence physician market power, although they do not directly estimate the magnitude of the effect of market structure on prices.

(MPIER) from the Center for Medicare and Medicaid Services (CMS), which contains all practicing physicians in the US, allows us to aggregate physicians by practice location and measure establishment sizes by specialty and geography. In addition, we use confidential Census Bureau data from the Longitudinal Business Database (LBD), Economic Censuses (EC), and Business Register (SSEL) to observe firm-level linkages across establishments based on IRS tax IDs, and to measure total payroll and sales from all sources for each firm. We link these databases to Truven Health Analytics MarketScan data on ambulatory care (non-hospital) prices negotiated between physicians and a large sample of private commercial insurance companies covering every state in the US. Together, these data provide a uniquely comprehensive panel of virtually every physician market nationwide over twelve years.

Second, we address a fundamental challenge of potentially endogenous practice organization choices. We do this by constructing a new panel database of state-level law changes that affect physicians' organizational incentives and practice sizes. The database quantifies judicial decisions that change the enforceability of non-compete agreements (NCAs), which restrict an employee's ability to leave a firm and compete against it. As documented by Bishara (2011), NCA laws vary along seven dimensions across states and over time. Following Bishara's methodology, we measure each of these legal dimensions for every state-year during the sample period. We then trace the effects of these judicial decisions through changes in organizational incentives, practice structures, and market concentration to measure the impacts of these practice characteristics on negotiated prices.[4]

Our findings highlight an important role of state NCA policies in affecting healthcare markets. We show that a judicial decision decreasing NCA enforceability by 10% of the observed policy spectrum (about 0.39 standard deviations) causes physician prices to fall on average by 3.7%. Such a policy change at the national level would reduce aggregate medical spending by over $20 billion annually. Despite the important role of NCAs, 39 states have never comprehensively reviewed and legislated NCA policies, and instead the law itself is defined by the set of case-specific judicial decisions—following common law precedent—of the type we use for variation in this paper. Consistent with evidence from Clemens, Gottlieb, and Molnar (2017), we find that these price effects are driven overwhelmingly by procedures that require relatively high levels of capital equipment,

[4] NCA law has been used previously as a source of variation in important work by Fallick et al. (2006), Marx et al. (2009), and Garmaise (2009). These papers focus on a few specific law changes (in Michigan, Texas, Florida, and Louisiana) or cross sectional differences (Massachusetts vs. California) rather than using the full panel of judicial law changes on all seven legal dimensions and in all U.S. states, as we do. Lavetti et al. (2018) provide evidence from survey data that the use of NCAs in physician employment contracts is very common, with about 45% of primary care physicians in group practices bound by NCAs.

supplies, or non-physician labor, while there is little systematic effect for procedures that primarily require physician labor as an input.

To understand the mechanisms through which judicial decisions that alter NCA enforceability might generate these effects on prices, we test for evidence on how the law changes affect practice organization. We find that changes in NCA enforceability significantly affect the rate of physician-establishment job separations and the creation of new establishments, which in turn affects the distribution of establishment sizes. Our controlled event-study estimates suggest that an average law change increasing NCA enforceability causes a 165 point decline in the HHI within 2 years.

In addition to physician practice organization and associated costs, we evaluate several alternative mechanisms through which judicial decisions on NCAs could potentially generate these effects. Survey data from Lavetti et al. (2018)—which links information on whether physicians have signed NCAs to data on service prices and quality measures—provides evidence against the hypothesis that changes in NCA enforceability might generate quality differences between firms due to physician sorting. We also test whether changes in NCA enforceability affect the total number of physicians in a market through entry or exit and find no such evidence. Since health insurers do not tend to use NCAs, it is unlikely that changes in insurer organizational structure would be affected by these legal changes. Nonetheless, our main specifications control for changes in insurer concentration.

That these NCA policies appear to affect prices through a practice organization channel, without directly affecting other factors that may influence prices, suggests the law changes may satisfy the instrumental variables assumptions required to estimate the local average treatment effect of practice organization on negotiated prices. After discussing and evaluating the validity of the IV assumptions, we use these seven NCA law indices as instruments to estimate the effects of physician practice sizes and market concentration on prices. Our fixed effects specifications control for unobserved heterogeneity across geographic markets as well as for census-division-by-year effects, medical specialty effects, service procedure code effects, and medical facility type effects. In addition, our unique ability to observe both establishments and firms allows us to estimate the marginal effect on prices of increasing establishment concentration conditional on firm concentration, and vice versa.

We find that changes in concentration have heterogeneous effects on negotiated prices that depend on the structural nature of the changes. Increases in concentration caused by the growth of physician establishments lead to negative price effects, while increases in concentration due to the growth of firms that may have physically distinct establishments cause prices to rise. Specifically, we find that a 100 point increase in the *establishment-*

based HHI causes a reduction in negotiated prices of about 1.3% to 1.7% on average. In contrast, the same increase in concentration caused by firm-level consolidation holding fixed establishment concentration causes prices to increase by 1.0% to 2.0%.[5] OLS specifications imply very small (but statistically significant) positive price effects of 0.02% or less, consistent with evidence from Baker et al. (2014). [6]

Taken together, these results suggest that the effects of consolidation on prices depend on a tradeoff between the efficiency gains of larger establishments and the increased negotiating power associated with bargaining as a larger organization. To the extent that larger establishments have improved bargaining positions, any consequent positive effect on prices is outweighed by insurers extracting cost reductions due to economies of scale, resulting in a net negative price effect. These economies of scale could be due, for example, to shared nursing, laboratory, technological, and administrative resources among more physicians. However, when practices grow larger through multi-establishment expansion, the net effect on prices is positive, implying that any economies of scale from mergers of physically-distinct practices have smaller effects on prices than does the associated change in bargaining position. Although the variation in practice organization (caused by NCA law changes) underlying our estimated local average treatment effects may differ to some extent from the margin of variation occurring more broadly in physician markets, such as hospital acquisitions of physician practices, our estimates indicate that price effects come predominantly from the channel of establishment-level growth. The negative net relationship between concentration and prices suggests there may be important efficiency gains from physical consolidation of practices.

The paper is structured as follows. Section 2 provides background on non-compete laws and their usage by physicians. Section 3 describes the data sources. Section 4 presents reduced-form evidence on the effect of NCA laws on negotiated prices, and Section 5 elaborates on the potential mechanisms behind these effects. Section 6 describes our main empirical model, IV results, and a variety of robustness tests. Section 7 concludes and discusses the policy implications of our findings.

2 Background: Non-Compete Laws and Physicians

NCA Laws and Changes: Non-compete agreements are clauses of employment contracts that prohibit an employee from leaving a firm and competing against it. In the

[5]We define an establishment as a specific physical practice location, differentiated by mailing addresses. In contrast, firms may own multiple establishments, and we identify firms by IRS tax IDs.

[6]Baker et al. (2014) use Marketscan price data but estimate market structure using Medicare beneficiaries.

case of physicians, who compete in local geographic markets, NCAs prohibit practicing medicine within a specified geographic area and fixed period of time. Physicians bound by an NCA who leave their firm must either exit the geographic market, wait until the NCA has expired, or take a job outside of medicine.[7] Common physician NCAs restrict competition within 10-15 mile radii for 1-2 years. Allowable radii depend in part on how far patients generally travel to see a doctor, which can vary across urban and rural markets, and by physician specialty. However, since the enforceability of NCAs is determined by state law, there is also a large degree of variation across states in how restrictive these contracts can be. For example, some states do not allow employment-based NCAs to be enforced at all, while other states allow them to be easily enforced with broad market definitions and/or long durations.

The permissibility of NCAs dates back to at least 1621 under English common law, and 39 US states still follow common law in determining the enforceability of NCAs. Thus, historical precedent is the main determinant of NCA policies in most states. However, states that follow the same common law origins have diverged dramatically in their enforcement of NCAs. For example, Kansas has the second highest NCA enforceability measure while North Dakota has the lowest measure, despite the fact that both states follow legal traditions that were heavily influenced by English common law.

Common law requires judges to consider three specific questions when evaluating NCA contracts. First, does the firm have a legitimate business interest that is capable of being protected by an NCA? Second, does the NCA cause an undue burden on the worker? And third, is the NCA contrary to the public interest? Changes in the interpretation and relative importance of these questions have caused judicial decisions to break from precedent. Under common law, a judge's decision to deviate from precedent has the effect of changing the law going forward.

The vast majority of these policy changes involve legal cases unrelated to healthcare markets. For example, in Shreveport Bossier v. Bond (2001) a Louisiana construction company attempted to enforce an NCA against a carpenter. The state Supreme Court ruled that the NCA could only prevent the carpenter from establishing a new business, but not from joining a pre-existing firm. This decision abruptly changed the law in the state, allowing all workers, including employed physicians, who had previously signed NCAs to escape the restrictions and move to other firms.

To take advantage of the rich variation in the relevant legal environments, we quantify

[7] In some states contracts with NCAs are required to specify a buyout option. For example, Sorrel, AL (2008) describes a case in Kansas in which a physician had a buyout option of paying her former practice 25% of her earnings during the NCA restriction period.

variation in NCA laws across states and 52 law change events during our study period (28 that strengthen NCA enforceability, and 24 that weaken it) using the methodology developed by Bishara (2011). These data are described in detail in Section 3.4.

Physician Markets and the Use of NCAs: In order to understand the mechanism behind these instruments, it is useful to know what motivates physician practices to use NCAs. Lavetti, Simon, and White (2018) study this question and conclude that physician practices use NCAs primarily to deter physicians who exit a group practice from taking clients with them to another firm. In firms that provide skilled services, information asymmetries between clients and service providers make it costly for clients to search for new providers, generating loyalty towards providers. The loyalty of patients to their doctors is arguably the most valuable asset of most physician practices—the stock of patients is often the basis for determining a price when practices are sold—but firms have no direct property rights or control over these valuable assets. They are threatened by the possibility that steering patients to a new physician who joins the practice could lead to losing the patients if the physician were to exit the practice and the patients were to follow. NCAs can prevent this type of loss.

Lavetti et al. (2018) find that about 45% of primary care physicians in group practices are bound by NCAs on average, where use ranges in a five state sample from about 30% in California, a low enforceability state, to 66% in Pennsylvania. They also show that NCAs are used more frequently in practice settings where ongoing patient relationships are more valuable, such as in office-based practices as opposed to in hospitals, and in metro or micropolitan markets where the supply of physicians is larger relative to the population, making patient stocks more valuable.

Our empirical analyses suggest that NCA enforceability is generally negatively correlated with physician practice sizes and market concentration. Although explaining the nuances of all of the legal dimensions of NCAs is beyond our space constraints (we provide a brief overview in Appendix Table A2), an example of one dimension of the law called the 'Employer Termination Index' measures the extent to which state law allows a firm to fire a worker and still enforce the NCA. In some states this action would be legal, while in other states NCAs can only be enforced if the worker quits. An increase in this component of the law causes a spike in job separations and a significant decrease in establishment concentration as it becomes less costly for firms to fire workers, and as workers tend to move to smaller practices or start new practices. In contrast, another component of the law called the 'Blue Pencil Index' measures the extent to which NCA clauses that are overly restrictive to workers can be modified by judges ex post and thus still enforced. This dimension of the law is the only one that is positively correlated with concentration; this correlation could occur if increases in this dimension make it harder for physicians to escape pre-existing NCA agreements, leading practices to grow larger over time by deterring exits. Each of the seven dimensions of NCA law undergoes a number of state level judicial changes during our sample period (1996-2007), generating exogenously timed variation in physician concentration in the affected state relative to nearby states. In Sections 4 and 6.6 we present evidence supporting the exogeneity of the law changes, including a lack of pre-trends in either concentration or prices, and we show that there is no clear correlation between law changes and state-level economic or political measures.

3 Data

We use data from a variety of sources to construct a longitudinal database that includes physician market concentration measures, negotiated prices, and our 7 instrumental variables. The main sample, during which all of the data components are available, spans 1996-2007.

3.1 MPIER Physician Panel

The Medicare Physician Identification and Eligibility Registry (MPIER) is a database collected by the Center for Medicare and Medicaid Services (CMS). The database began in 1989 when the Health Care Financing Administration assigned unique identifying numbers to all physicians associated with Medicare. In 1996 the physician identification requirement was strengthened under HIPAA, which mandated every physician to receive an identifying number and be included in the MPIER regardless of their association with Medicare. The coding system used in MPIER was in place through 2007.

The MPIER data provide each physician's name, identifying number, the number of practices that the physician is associated with, the dates of any changes in practice affiliations, physician specialties, a group practice indicator, the practice billing address, and the practice's business location street address. Physicians can have multiple practice affiliations at the same time, and each location at which a physician treats patients is required to be recorded. Using the `soundex` fuzzy matching algorithm[8] we construct a longitudinal database of establishments by matching physicians to establishment street addresses. We allow slight differences that may be due to typographical errors in street addresses, but we require exact matches on street numbers and office numbers. In the appendix we examine the sensitivity of our results to the fuzzy matching tolerance parameter.

[8] See R. Russell US Patent 1261167 (1918).

There are two limitations with this database. First, we cannot observe connections between establishments, which could be important to the extent that multi-establishment firms negotiate as a single entity with insurers. Second, we cannot observe revenues or allocations of time for physicians that work in multiple establishments. To calculate concentration measures from these data we use the shares of the number of physicians in a given market. Each physician with multiple establishment associations is allocated in equal proportions to each of the establishments for as long as each establishment continues, so that each physician contributes exactly one to the total physician headcount at any time. Although it has some limitations, this dataset is to the best of our knowledge the only complete national census of individual physicians during our study period.

3.2 Longitudinal Business Database

Several of these data limitations can be overcome with data from the Census Bureau's confidential Longitudinal Business Database (LBD), which contains data on all non-farm employer establishments in the US and is available from 1976 to (nearly) the present. The LBD contains establishment employment, payroll, industry codes, and county locations with firm linkages via IRS Employer Identification Numbers. Physician practices are identified by NAICS industry code 621111, described as 'Offices of Physicians (Except Mental Health Specialists).' While the LBD solves the problem of observing firm-level information, it too has limitations since it does not contain the medical specialties of the physicians at each firm.

We also use the LBD to construct longitudinal measures of health insurance market concentration using data on sales from firms in NAICS code 524114, 'Direct Health and Medical Insurance Carriers'. We control for insurer concentration in our main specifications.

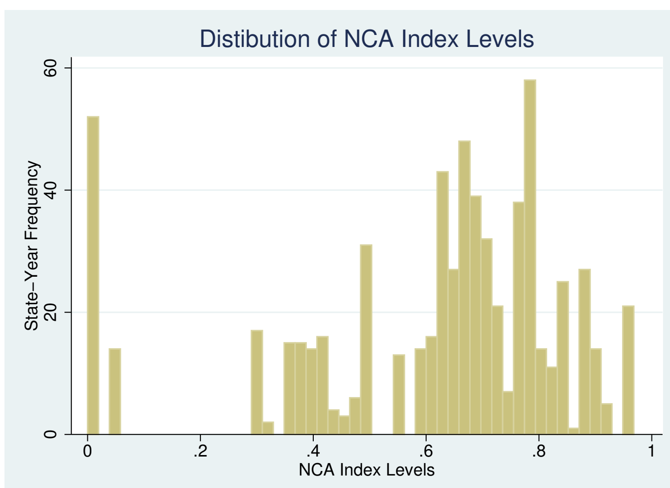
3.3 MarketScan Negotiated Prices Data

Data on prices negotiated between physicians and private commercial insurers come from the Truven Health Analytics Marketscan database. The database includes the medical claims for all active employees and their dependents from a sample of large firms. We use data between 1996-2007 on average negotiated prices, counts, and variances of negotiated prices by county, year, physician specialty, Current Procedural Terminology (CPT) code, and medical facility type (for example, physician office, urgent care facility, end-stage renal disease facility). This combination of dimensions gives about 10 million average negotiated prices, based on prices from about 550 million procedure claims, covering every state-year and nearly every county-year during the study period. Note that since the average is taken over claim level observations, it reflects the empirical distribution of prices, which naturally weights practices according to their sample market shares in each cell. Our analyses use the top 35 most common procedure codes to reduce imbalance across cells caused by infrequently used procedure codes. The sample contains only prices for ambulatory services that are not hospital-based.

3.4 NCA Law Data

We develop a new database quantifying the variation in state-level NCA laws systematically over time, following the measurement system developed by Bishara (2011). Bishara (2011) analyzes case law in each state and scores states along 7 different dimensions, following the framework from a series of legal texts by Malsberger (1991, 1996, 1997, 2000, 2001, 2003, 2004, 2006, 2008, 2009, and 2011). Each of the dimensions is assigned a weight, based on legal knowledge of their relative importance, to create a weighted index score. The 7 components and the scoring system are described in detail in Table A2.

Figure 1: Distribution of NCA Index Levels



Notes: Data points underlying the histogram are state-year observations of the NCA Index, a weighted sum of the 7 NCA law dimensions. The Index is scaled to range from 0 to 1, where 0 is the least restrictive state-year in the sample and 1 is the most restrictive.

The analysis by Bishara (2011) quantifies laws in 1991 and 2009. Using the same coding methodology, we code the timing and degree of the law changes, creating an annually-measured longitudinal dataset that spans the period 1991-2009 and matches the endpoint measures of Bishara (2011).[9] During the period we study, there were 52 law change events. Each event moved one or more of the seven legal dimensions. Previous work using NCA law changes for variation in organizational incentives in non-physician markets examined specific events in Michigan (Marx et al. (2009)) and in Texas, Florida, and Louisiana (Garmaise (2009)).

In the Bishara (2011) data, the weighted sum of scores for all seven components ranges from 0 to 470, where 470 (Florida) corresponds to policies under which NCAs are easiest to enforce, and 0 means that NCAs cannot be enforced in employment contracts. In our analyses we normalize the measures by dividing each component by its maximum value to create continuous measures that range from 0 to 1, representing the observed spectrum of each policy dimension, where 1 corresponds to the state-year policy in which NCAs are easiest to enforce. Figure 1 shows the frequencies of these NCA index values in all state-year pairs in our sample, and Table 1 presents summary statistics on the changes in legal indices by Census region, indicating that changes are geographically dispersed and move in both directions within each region. The average magnitude of law changes in our sample is 0.08 in absolute value, which is about one-third of a standard deviation of the overall policy variation.

Table 1: NCA Law Components: Descriptive Statistics by Census Region

Region	Northeast	Midwest	South	West	Total
Average Index	0.66	0.72	0.64	0.51	0.63
Standard Deviation of Index	0.28	0.25	0.22	0.27	0.26
Maximum Index	1.00	1.00	0.96	0.88	1.00
Minimum Index	0.00	0.00	0.00	0.00	0.00
Number of Law Changes	10	11	22	9	52
Number of States in Region	9	12	17	13	51
Number of Index Increases	7	7	9	5	28
Number of Index Decreases	3	4	13	4	24
Average Magnitude Positive Index Change	0.04	0.12	0.06	0.08	0.08
Maximum Positive Index Change	0.09	0.26	0.14	0.16	0.26
Average Magnitude Negative Index Change	–0.07	–0.07	–0.15	–0.05	–0.09
Maximum Negative Index Change	–0.09	–0.10	–0.63	–0.07	–0.63

Notes: Statistics in the table represent data from 1994-2007 for each state-year in which a legal precedent exists, and uses physician-specific laws whenever applicable. States that forbid NCAs either generally or for physicians specifically are CO, DE, MA, and ND. The minimum of each component is 0 and the maximum of each component is normalized to 1.

[9]We are grateful for legal expertise from Richard Braun, J.D., and for research assistance from Akina Ikudo, and David Krosin in the creation of this dataset.

4 Effects of NCA Law Changes on Physician Prices

4.1 Reduced-Form Effects

We begin by estimating the effect of NCA policies on negotiated prices using the following equation:

$$P_{mfpct} = \alpha + \beta NCA_{s(c),(t-1)} + \eta_m + \pi_f + \theta_p + \gamma_c + \nu_{dt} + \varepsilon_{mfpct} \quad (1)$$

The dependent variable is the average negotiated price of procedure code p performed by a physician with medical specialty m in facility type f, county c, and year t. The fixed effects specification controls for procedure code effects, specialty effects, facility type effects, county effects, and census division by year effects. The coefficient β therefore identifies the extent to which prices move differentially in counties in law change states relative to those in other states in the same census division (there are 4.6 within-division comparison states, on average). This specification, which we use in lieu of imposing functional form restrictions on time trends, allows the prices in each census division to have any arbitrary unobserved idiosyncratic variation over time. Since negotiations between physicians and insurers tend to occur annually, we use a lagged specification, consistent with previous studies of negotiated healthcare prices (Dafny et al. (2012), Dunn and Shapiro (2014), and Baker et al. (2014)).

Table 2 presents estimates from Equation 1. In the first row we use the weighted average NCA enforceability index created by Bishara (2011), and in the rows below we show estimates from separate regressions for each of the seven legal indices. The results suggest that prices increase by 4.9% in the year following a 0.1 unit increase in the weighted average NCA index, and a 3.7% effect persists after 2 years. We find that six of the seven individual indices have significant effects on prices, and five have positive coefficients ranging from 0.6% to 1.8% per 0.1 unit increase in the corresponding index.

4.2 Event Study Analyses of Price Trends

To evaluate whether these results may be affected by differential trends in states with changes in NCA laws, Figure 2 presents event study estimates of Equation 1. While the regressions report estimates from the full sample, the graphs depict event studies from a sample that is limited to treatment states with only one law change within the event window, and thus uncontaminated, and control states in the same census division with no law changes. The subfigures in column 1 are estimated using binary indicators of an

Table 2: Reduced-Form Price Effects, by NCA Index

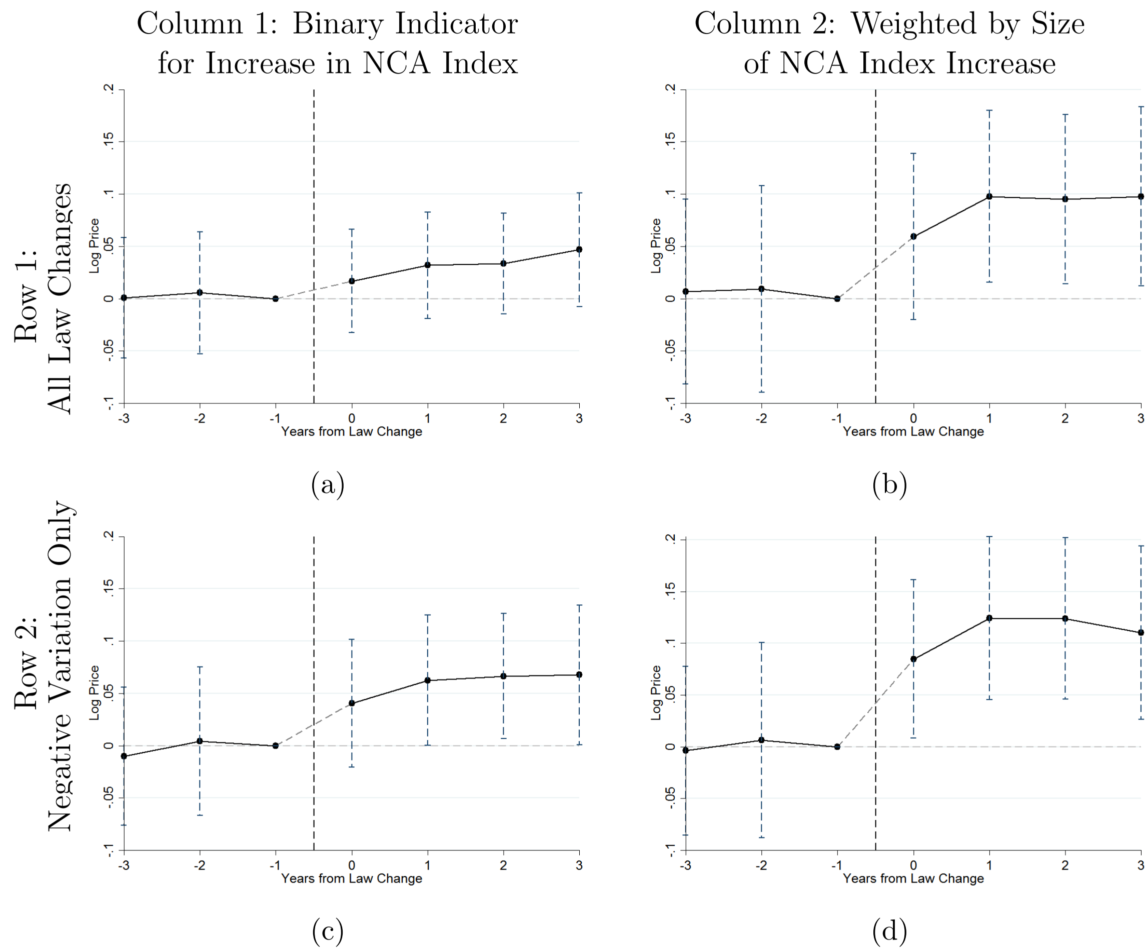
Dependent Variable:	ln($Price$)	
	$\text{NCA}_{(t-1)}$	$\text{NCA}_{(t-2)}$
NCA Index (Weighted Average)	0.491*	0.370*
	(0.094)	(0.095)
Statutory Index	0.103	0.118*
	(0.062)	(0.038)
Protectible Interest Index	–0.007	–0.094
	(0.121)	(0.105)
Consideration Index Inception	0.242*	0.179*
	(0.045)	(0.049)
Consideration Index Post-Inception	0.062*	0.059*
	(0.030)	(0.025)
Burden of Proof Index	0.193*	0.131*
	(0.036)	(0.039)
Blue Pencil Index	–0.093	–0.147*
	(0.078)	(0.061)
Employer Termination Index	0.260*	0.177*
	(0.082)	(0.075)
N (Each Model)	3,026,780	3,026,780

Notes: Each coefficient comes from a separate regression of log prices on either the first lag (column 1) or second lag (column 2) the corresponding legal index. Each legal index is scaled to range from 0 to 1, where 1 corresponds to the highest observed enforceability measure for that index. All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. All standard errors, in parentheses, are clustered by state-year. * indicates significance at the 0.05 level.

increase (+1) or decrease (-1) in NCA enforceability, while estimates in column 2 use the continuous magnitude of the law changes. Row 1 uses all law changes, while row 2 uses only variation from decreases in NCA enforceability.

There are several notable conclusions from these event studies. First, increasing NCA enforceability leads to higher prices on average. Figure 2b, for example, suggests that a 0.1 unit increase in NCA enforceability leads to about 9.8% higher prices on average within 2 years, a larger effect than is observed in the full sample in Table 2. Second, there is very little evidence of differential pre-period trends in states with law changes. We also test the common trends assumption for a broader set of law changes that includes the first law change in each state plus any subsequent law changes that occured at least three years after the previous change, providing an uncontaminated three-year pre-period. The parallel trends assumption is also satisfied in this broader sample—the p-value of an F-test that all three pre-period coefficients are equal to each other is 0.96. Third, decreases

Figure 2: Event Study Plots: Reduced-Form Price Effects



Notes: Sample includes treatment states with only one law change within the event window, and control states in the same Census division as the treatment state that had no law changes during the corresponding event window. Estimates are from fixed effects regressions including county effects, census division by year effects, procedure code effects, facility type effects, and specialty effects. Specialties included in sample are primary care and non-surgical specialists. Dashed lines represent 95% confidence intervals based on standard errors clustered by state-year. Year 0 is the calendar year during which the law change occurred, and the dependent variable is normalized to zero in year -1.

in enforceability have (negative) price effects that are similar to the overall estimates, suggesting that the effects of positive and negative law changes are symmetric. Finally, the price effects appear to flatten after about two years, suggesting that the law changes primarily impact price levels as opposed to rates of growth, and that the effects occur fairly quickly.

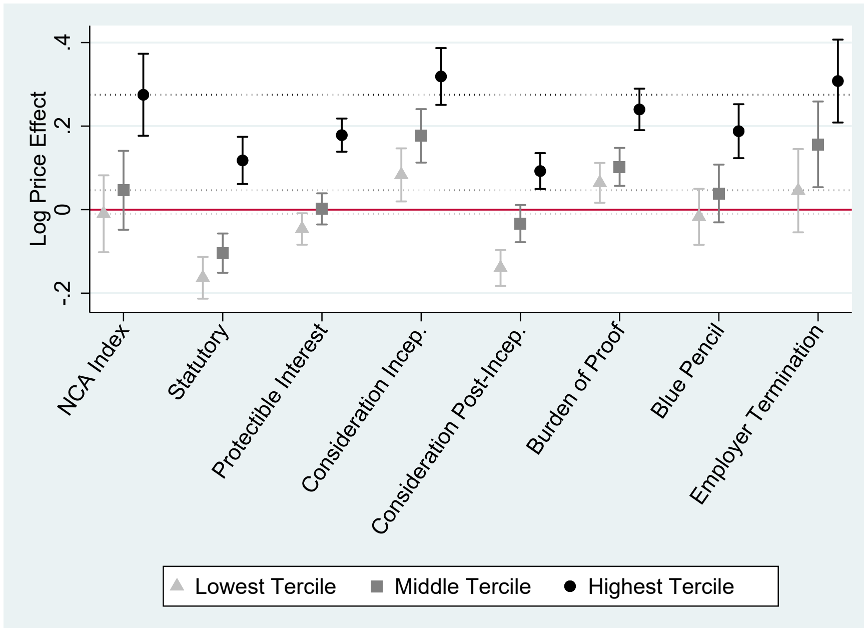
4.3 Heterogeneity in Price Effects and Potential Mechanisms

To help clarify what types of mechanisms might be driving these price effects, we investigate whether there is systematic heterogeneity across different types of medical procedures. Our test is motivated by the analyses of Clemens, Gottlieb, and Molnar (2017), who study the extent to which physician practices negotiate price schedules with private insurers that are benchmarked to Medicare prices. They find that in smaller physician groups, about 90% of procedure prices are negotiated relative to Medicare prices, while in larger practices (with at least $1 million in billings) only 40% of procedure prices are benchmarked to Medicare. In addition, they show that deviations from Medicare benchmarks are most likely to occur for procedures that use capital-intensive inputs, as opposed to labor-intensive procedures. Finally, they show that deviations from the Medicare price schedule tend to be negative for capital intensive procedures, potentially narrowing the gap between marginal costs and the average cost estimates used to set Medicare payments. This result suggests that insurers may extract a portion of the cost savings associated with larger physician practices to bring prices closer to marginal costs. In contrast, to the extent that prices of labor-intensive services deviate from the Medicare schedule, the deviations tend to be positive. This supports the notion that for at least some procedures physicians are not entirely price-takers, consistent with evidence from Kleiner et al. (2015) and Dunn and Shapiro (2014).

To get a sense of whether NCA laws may be affecting prices through practice size or through some other mechanism, we follow this intuition and test for differential effects of NCA laws on prices of procedures that have high versus low overhead costs and high versus low use of physician labor. We link the Marketscan price data to Medicare data containing the resource-based Relative Value Units (RVUs) for each procedure code. RVUs, which are used in calculating Medicare payments, are divided into three categories: 'Work RVUs' capture the amount of physician labor typically used in the procedure; 'Facility Practice Expense RVUs' capture the average use of equipment, office space, supplies, and non-physician labor expenses; and 'Malpractice RVUs' are designed to cover the costs associated with malpractice insurance for the procedure. For each procedure, we calculate the ratio of Facility Practice Expense RVUs to physician Work RVUs, and group procedures by tercile of this ratio. The top tercile, for example, contains procedures that use primarily capital and other practice costs, while using relatively less physician labor. We then interact these tercile indicators with the NCA variable in Equation 1, and re-estimate the model.

Figure 3 depicts the coefficient estimates and 95% confidence intervals from these regressions, which are estimated separately by NCA index, as in Table 2. The figure

Figure 3: Price Effects by Tercile of the Ratio of Medicare Facility Practice RVUs to Physician Labor RVUs



Notes: Estimates are similar Table 2, with NCA laws interacted with terciles of Medicare facility practice expense RVUs divided by physician work RVUs. Error bars are 95% confidence intervals, based on standard errors clustered by state-year.

shows that all seven indices have significant and positive price effects, ranging from 0.9% and 3.2% per 0.1 unit increase in the NCA index, for procedures with the highest ratio of practice RVUs to physician labor RVUs. Among procedures in the bottom tercile—those with a relatively intensive use of physician labor relative other inputs—there is no clear pattern of price effects: two of the indices have insignificant coefficients, and the remaining five range between -1.6% to +0.8% per 0.1 change in the NCA index. Using the weighted average NCA index, the coefficient in the bottom tercile is slightly negative, -0.1%, and statistically insignificant, compared to 2.7% in the top tercile of procedures. Moreover, for every dimension of the NCA indices, the effect on prices in the top tercile is significantly greater than the effect on those in the bottom tercile. These results help narrow the scope of potential mechanisms that could produce this pattern. Consistent with the evidence on negotiated prices from Clemens et al. (2017), the findings suggest that effects of NCAs on prices appear to occur through a mechanism related to practice organization and overhead costs, rather than through one related to physician labor costs.

We also test whether these price effects may be caused by changes in the aggregate relative supply of physicians. As NCA law changes may affect physicians' option sets within the local market, they may alter the overall attractiveness of the geographical location relative to others. If the laws drive flows of physicians across geographical markets, they may affect prices through total market size. To investigate this possibility, we take two measures of market size and regress them on the law changes. Table A16 shows no effect of the law changes, as measured by the NCA Index, on total physicians per capita, and column 3 of Table A17 shows no systematic effect of the law changes on population.

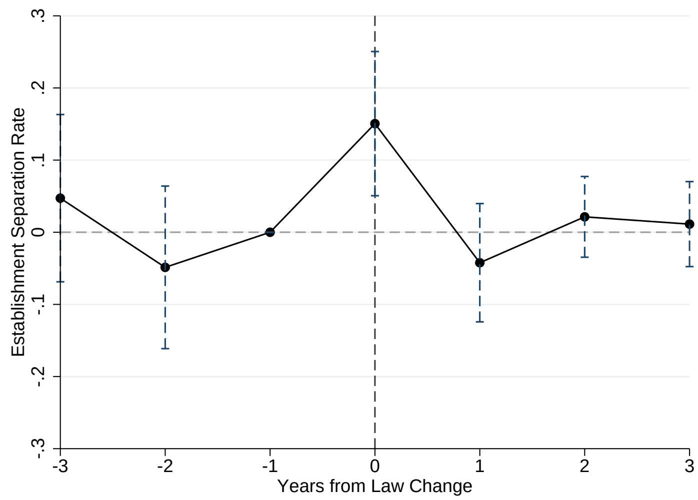
An additional means through which NCAs could affect prices due to physician sorting is through quality. Survey data from Lavetti et al. (2018), however, indicates no quality differences between physicians with versus without NCAs in a series of vignette-based questions designed by clinical experts to elicit knowledge of best practices, diagnostic skill, treatment patters, and clinical recommendations. Furthermore, to the extent that physician quality is correlated with prices, a quality difference would be reflected in a price difference between NCA users and non NCA users in the same market, but there is no such price difference conditional on practice sizes. Finally, physician experience—which is strongly correlated with measures of patient satisfaction and perceived quality (Choudhry et al., 2005)—does not differ across NCA-using and non-NCA using practices.

5 Effects of Law Changes on Practice Organization and Market Concentration

Figure 4 shows an event study plot with the same specification as in Figure 6, where the dependent variable is the average separation rate of physicians from practices, measured using the MPIER data. For ease of interpretation, the law changes are scaled by -1 so that the graph can be interpreted as an increase in the separation rate when NCA enforceability declines. The estimates suggest that an average decrease in NCA enforceability leads to a 15 percentage point jump in the rate of job separations in the year of the law change. There is again no clear anticipatory trend prior to the law change. We also test for pre-trends using a broader sample of all state-year observations prior to the second law change within each state. The p-value of an F-test that the three pre-period coefficients are equal is 0.14.

Figure 4 also shows that the average separation rate returns to the pre-event range. One potential explanation for this temporary increase is the presence of an accumulated stock of physicians who would like to switch practices but are prevented from doing so by an NCA. When the enforceability of the NCA restriction declines it becomes less costly to move, and a stock of physicians moves simultaneously. Once the moves are completed, there is less pent-up desire to switch practices and separation rates subsequently decline.

Figure 4: Event Study: Physician-Establishment Separation Rates Before and After Decrease in Enforceability



Notes: Sample includes treatment states with only one law change within the event window, and control states in the same Census division as the treatment state that had no law changes during the corresponding event window. Estimates are from fixed effects regressions including county effects, census division by year effects, and specialty effects. Specialties included in sample are primary care and non-surgical specialists. Dashed lines represent 95% confidence intervals based on standard errors clustered by state-year. Year 0 is the calendar year during which the law change occurred, and the dependent variable is normalized to zero in year -1.

The pattern in Figure 4 bolsters the evidence that NCA laws constrain physicians' choices over practices, suggesting that there are effects on practice organizational structures. Still, it is not obvious that even an exogenous event causing separations should change establishment sizes or market concentration. Separating physicians could start new small practices, reducing the average practice size, or join larger established practices, increasing establishment sizes. As shown in Appendix Table A4, the law changes also significantly affect the rates of new establishment births and incumbent establishment deaths. Alternatively, if separations are driven by idiosyncratic preferences, a spike in separations could simply lead to some physicians exiting a practice and other physicians entering it, with no net effect on concentration. The large spike in separations corresponding to the timing of law changes is only suggestive of an underlying mechanism that has the potential to cause the distribution of practice sizes to change.

Table 3, however, shows that these changes in separation rates and establishment counts also lead to changes in the average sizes of establishments. The dependent variable in this model is the log of the number of full-time equivalent physicians per establishment, where full-time equivalence is calculated by assigning equal fractions of each physician to

Table 3: Fixed Effects Models of Establishment Sizes

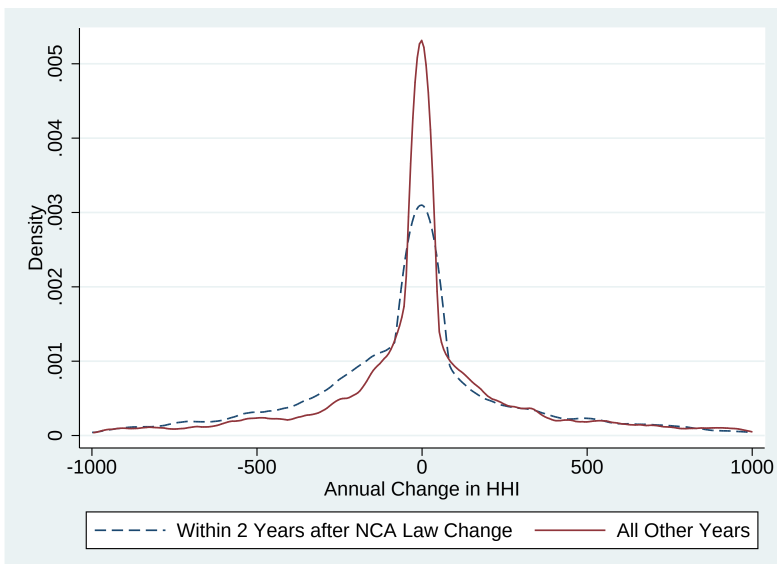
Dependent Variable: Log FTE Physicians per Establishment	By Component (1)	Combined (2)
Statistory Index$_{t-1}$	–0.169*	–0.140*
	(0.038)	(0.048)
Protectible Interest Index$_{t-1}$	–0.026	–0.178*
	(0.044)	(0.070)
Burden of Proof Index$_{t-1}$	–0.048	–0.262
	(0.042)	(0.146)
Consideration Index Inception$_{t-1}$	–0.121*	0.081
	(0.035)	(0.162)
Consideration Index Post-Inception$_{t-1}$	0.044	0.099*
	(0.031)	(0.032)
Blue Pencil Index$_{t-1}$	–0.151*	–0.163*
	(0.027)	(0.030)
Employer Termination Index$_{t-1}$	–0.159	–0.103
	(0.110)	(0.129)
N	379,370	379,370
R-Sq		0.23

Notes: Column 1 reports estimates from separate regressions on each law component, and column 2 reports estimates from a regression including all 7 components. Dependent variable is the log number of FTE physicians per establishment in a county-year. All specifications include controls for the aggregate supply of physicians in the county and fixed effects for county and census division by year. FTE establishment sizes are estimated by assigning equal partial shares (summing to one) to all establishments at which a physician is active. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

every establishment location at which they treat patients during the year. The independent variables include one-year lags of each legal dimension, as well as fixed county effects and census-division-by-year effects. Since many practices contain multiple physicians with different specialties, we do not condition on specialty in these specifications. Column 1 shows that 6 of the 7 indices are negatively correlated with establishment sizes when included separately, consistent with the patterns from the event studies in Figure 6, and that three of these 6 indices have significant coefficients. This result is potentially consistent with stricter NCA policies deterring physicians from joining established practices that require NCAs, leading more physicians to start new practices instead. The significant coefficients range from a reduction in establishment sizes of 12.1% to a reduction of 16.9% per unit change in each index, or about –3.6% to –5.1% per standard deviation change in each index. Column 2 presents estimates from a single regression on all 7 coefficients, in which the coefficients differ somewhat because each judicial decision can cause correlated changes in multiple indices at once. These estimates are again generally consistent with the negative relationship between NCA enforceability and practice sizes.

These establishment size changes seem to be reflected in changes in market concentration. Figure 5 depicts the unconditional kernel density functions of annual changes in establishment HHIs within markets. Each observation underlying these distributions is a market-year-specialty combination. The solid line shows the distribution of changes in HHIs from one year to the next when there have been no recent changes to NCA laws. This distribution is centered around zero and has a relatively small variance. The dashed line shows the same distribution in the two years following any change to NCA laws. In years just after a law change, the density function is visibly and statistically significantly altered (Kolmogorov-Smirnov p-value<0.001), with less mass near zero and more mass in the region of negative HHI changes.

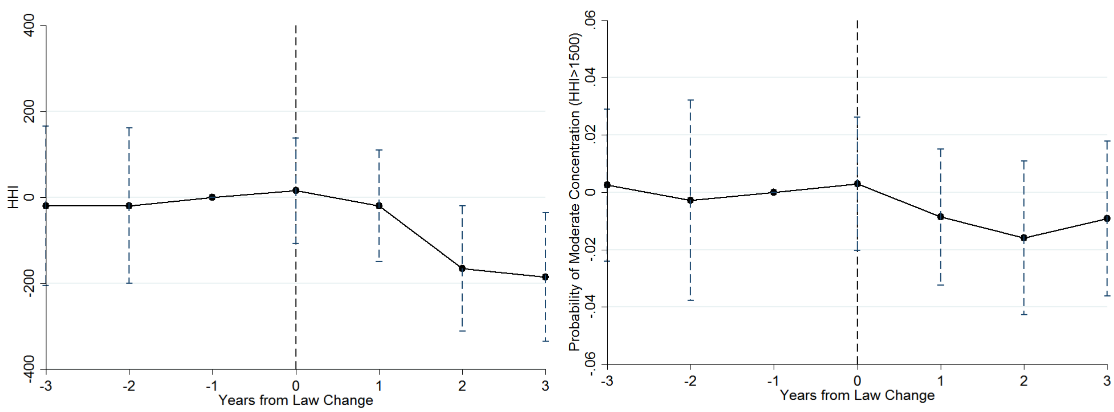
Figure 5: Distribution of Annual HHI Changes



Notes: Distributions are kernel density graphs of the change in annual HHI by CBSA-specialty for specialists. Distributions are truncated at $+/-$ 1000 for display. The p-value of Kolmogorov-Smirnov test of the equality of the full distributions is <0.001.

While the effects of law changes are clearly apparent in the unconditional comparison of HHI changes, our regression analyses control for geographic, intertemporal, specialty, and procedure variation. Figure 6 shows controlled event study plots using the same specification as in Figure 2, again restricting attention to law changes with uncontaminated event windows. The dependent variable in Figure 6a is the county-specialty HHI, and in Figure 6b it is a dummy variable that equals 1 if HHI is above 1500, the Department of Justice threshold for a 'moderately' concentrated market (DOJ Horizontal Merger Guidelines, August 2010). The figure shows that an average increase in NCA enforceability decreases HHIs by about 165 points within 2 years after the law change, with very little evidence of a differential pre-trend in treatment states. We again find no evidence contrary to the common trends assumption in the broader sample with uncontaminated pre-periods—the p-value of an F-test that all three pre-period coefficients are equal to each other is 0.87. Similarly, Figure 6b shows that the probability of exceeding the threshold for a moderately concentrated market decreases by about 1.6%. To be clear, our measure of establishment-level employment concentration is not directly comparable to the measure upon which the DOJ threshold is based, which is why we largely avoid making comparisons about HHI levels or using discrete thresholds in our analyses; this figure is only intended to be suggestive that changes in concentration occur both overall on average and at low to moderate concentration levels.

Figure 6: Event Study Plots: Concentration Before and After Law Changes



(a) Change in HHI

(b) Probability of 'Moderate' or Greater Concentration (HHI>1500)

Notes: Sample includes treatment states with only one law change within the event window, and control states in the same Census division as the treatment state that had no law changes during the corresponding event window. Estimates are from fixed effects regressions including county effects, census division by year effects, and specialty effects. Specialties included in sample are primary care and non-surgical specialists. Dashed lines represent 95% confidence intervals based on standard errors clustered by state-year. Year 0 is the calendar year during which the law change occurred, and the dependent variable is normalized to zero in year -1.

6 IV Model and Results

In this section, we evaluate the viability of the instruments and use them to estimate the effect of physician practice concentration on prices. We begin by describing the model specification and the identifying assumptions required to interpret the estimates as local average treatment effects (LATEs). We then present first and second stage IV estimates, followed by a brief overview of a wide range of sensitivity analyses.

6.1 Empirical Specification

We estimate a fixed effects two-stage least squares model, instrumenting for potentially endogenous variation in market concentration using the seven different dimensions of NCA laws as instruments. To differentiate between the effects of increases in concentration driven by larger firms as opposed to larger establishments, we allow both firm and establishment concentration to be endogenous regressors, the effects of which are overidentified by the seven instruments.[10] The first-stage equations for the two endogenous regressors are:

$$\begin{aligned} EC_{mc(t-1)} &= \alpha_1 + \beta_1 NCA'_{c(t-1,t-2)} + \beta_2 InsC_{s(c)(t-1)} + \eta_m + \pi_f + \theta_p + \gamma_c + \nu_{d(c)t} + \epsilon_{mc(t-1)} \\ FC_{c(t-1)} &= \alpha_2 + \beta_3 NCA'_{c(t-1,t-2)} + \beta_4 InsC_{s(c)(t-1)} + \eta_m + \pi_f + \theta_p + \gamma_c + \nu_{d(c)t} + \epsilon_{mc(t-1)} \end{aligned}$$

and the second-stage equation is:

$$\begin{aligned} P_{mfpct} &= \alpha_3 + \beta_5 \widehat{EC}_{mc(t-1)} + \beta_6 \widehat{FC}_{c(t-1)} + \beta_7 InsC_{s(c)(t-1)} + \eta_m + \pi_f \\ &+ \theta_p + \gamma_c + \nu_{d(c)t} + \varepsilon_{mfpct} \end{aligned} \tag{2}$$

where η_m, π_f, θ_p, γ_c, and $\nu_{d(c)t}$ are fixed effects for medical specialties, facility types, procedure codes, counties, and census division-by-years, respectively. NCA'_{ct} is a vector of the seven law instruments, measured at the state-year level. EC_{mct} is the establishment-based measure of market concentration, in contrast to FC_{ct}, the firm-based concentration measure. $InsC_{s(c)t}$ is the concentration of health insurance firms in the state, where $s(c)$ denotes the state in which county c is located. Our main specifications use HHIs as concentration measures, though we also present results using a range of alternative concentration measures including average practice size, the negative log HHI transformation, and the four and eight-firm concentration ratios.

The ability to distinctly observe both firms and establishments is a relatively unique feature of the data, and it allows us to estimate the marginal effect of each concentration measure on prices. The intuition behind this specification follows Dranove and Lindrooth

[10]See Malsberger 1991-2011 and Bishara 2011 for thorough discussions, and Appendix Table A2 for a brief overview, of the differences between these seven aspects of non-compete agreements in employment law.

(2003), who study the effects of hospital consolidation on operating costs. They find that when a hospital is acquired by another system there are no significant cost savings unless the acquisition leads to a physical consolidation of establishments, in which case median costs decline by about 14%. Building on this idea, we hypothesize that increases in establishment concentration, EC, conditional on firm concentration, FC, are likely to create cost efficiencies that may be partially extracted by insurers in negotiations, reducing prices. Conversely, evidence from the hospital setting suggests that increases in FC conditional on EC are less likely to result in cost savings, though they may improve bargaining position in negotiations with insurers, increasing prices. In Appendix Section 8, we more formally develop these hypotheses and derive the linkage between our estimands and more primitive underlying theoretical parameters in a model of simultaneous bilateral bargaining that builds on Ho and Lee (2014). Distinguishing between these two distinct forms of organizational consolidation can potentially provide useful insights into the factors that drive physician prices in the US.

Because we use county-level concentration measures and include county and census-division-by-year fixed effects, our estimates identify the effects of changes in bargaining position in local markets but do not incorporate potential bargaining power effects of multi-market physician systems, a distinction discussed in the context of large hospital systems by Lewis and Pflum (2015, 2017).

The model specifications use lagged concentration measures in the second stage, consistent with the literature as well as with the event studies. We also allow the instruments to affect concentration in either the contemporaneous or lagged year. Since the dependent variable in the first stage is lagged, the IVs include first and second lagged laws.

In Section 6.6 we evaluate a wide range of alternative specification assumptions, including alternative market definitions, assumptions about the treatment of multi-specialty practices in calculating HHIs, alternative measures of market concentration and firm sizes, omission of outlier law changes, and interaction effects between physician and insurer concentration.

6.2 Endogeneity Concerns

Our modeling approach follows the general structure-conduct-performance (SCP) framework for estimating effects of market structure on prices, which has several well-known limitations. One important class of concerns about SCP models described by Gaynor, Ho, and Town (2015) in their review of this literature is that measures of market structure are generally endogenous in pricing equations. A key difficulty in resolving this endogeneity is that there are many potential forms to consider. For example, latent variation in demand, costs, bargaining ability, or quality—all of which may affect prices—could be correlated with market structure, causing bias. Moreover, these bias components could oppose each other, creating ambiguity about the net direction of bias.

For example, consider the case of unobserved heterogeneity in practice cost functions. Since a high cost practice will negotiate higher prices in a standard bargaining model, ε_{ij} will contain some of this latent variation in practice costs. To the extent that insurers can steer patients towards low cost providers, the market share of high cost practices will be lower. The negative correlation between latent average cost and market share, which determines HHI, may cause downward bias in $\widehat{\beta}_5$.

On the other hand, a practice with high quality, unobserved to the researcher, is likely to have high market share. The error term contains the component of price variation caused by quality differences, and this error component is positively correlated with market share, possibly causing an upward bias in $\widehat{\beta}_5$.

In addition to being ambiguous, the sign of the net bias could depend on whether changes in practice size are motivated primarily by average costs or by bargaining position. Our empirical findings suggest that OLS estimates of β_5 and β_6 are attenuated towards zero. Our results generally support the conclusion that endogeneity of market structure in Equation 2 causes substantial bias. Previous empirical research on healthcare markets has also used instruments to address this endogeneity, as in the case of Dafny et al. (2012), which uses the merger of two large healthcare insurers as an instrument for concentration in local insurance markets. One contribution of our study is to develop new instrumental variables to overcome these biases in a variety of markets, including markets outside of healthcare in which NCAs are used frequently.

6.3 Model Assumptions

Structure, Conduct, and Performance: A second class of concerns described by Gaynor, Ho, and Town (2015) is that estimates can be sensitive to assumptions about market definition, conduct, and performance. We evaluate the sensitivity of our estimates across a range of potential market definitions and find the conclusions to be robust to this assumption. Perhaps more fundamentally, however, without estimating both conduct and performance, the choice of market structure measures can be arbitrary and potentially inconsistent with firm conduct. For example, choosing HHI as a market structure measure to estimate performance implies very specific implicit assumptions about conduct: homogeneous goods and Cournot competition. These assumptions are appropriately regarded with skepticism in many markets.

We make two points about firm conduct in our estimates. First, without a national panel of claims data covering our study period, we do not attempt to estimate firm conduct directly. Instead we take the approach that, using a variety of market structure measures, we identify patterns in negotiated prices under a broad conceptual framework. Each of these measures has underlying it a specific, and different, assumption about firm conduct. We show that the qualitative conclusions are identical regardless of our measure of market structure, suggesting that the assumptions of firm conduct do not substantially alter the findings once we correct for several other estimation challenges. We find the most important estimation challenge to be the endogeneity of these measures.

Second, there may be reasons to be less concerned about the implicit assumptions of homogeneous goods and Cournot competition in the case of physician practices, at least relative to hospitals. Hospitals often have observable (to the patient and econometrician) objective measures of quality, such as mortality rates, that vary substantially. In addition, consumers tend to have strong perceptions of quality differences. For example, research hospitals affiliated with prominent universities may be perceived to have sufficiently higher quality such that consumers are willing to pay higher premiums for insurer networks that include them (see Capps, Dranove, and Satterthwaite, 2003). Although some large physician groups have similar brand affiliations with prominent research hospitals, there is frequently no clear analogue among physicians to the dominant hospital phenomenon. There are few, if any, objective measures of physician-level quality outside of hospitals. Although consumers may have preferences for visiting a doctor that they personally know well, loyalty to a doctor is very different than a commonly-shared perception of quality, and it does not necessarily lead to correlation in willingness to pay across consumers.[11] We also condition on physician specialty, medical procedures, and geography, making the services closer to being conditionally homogeneous. Still, there is very little empirical evidence from the literature on measures of either objective heterogeneity in physician quality (outside of hospitals) or consumers' perceptions of differences in quality, and we have nothing concrete to add to the dearth of evidence on this question.

There is also some direct empirical evidence that the assumption of Cournot competition is reasonable in the case of physician practices. Gunning and Sickles (2013) estimate a structural model of conduct among physician practices that builds on the approach developed by Bresnahan (1989). Using data from the American Medical Association, they

[11] For example, if homogeneous consumers are uniformly distributed across doctors, even if each consumer is willing to pay more for an insurance network that includes their own doctor, the average willingness to pay for any particular doctor is the same, since willingness to pay is not correlated across consumers in the market.

estimate firm price elasticities and reject the null hypothesis of perfect competition, but they fail to reject the hypothesis of Cournot conduct, suggesting that using HHI as a market structure measure is consistent with firm conduct for physicians.

IV Assumptions: Interpretation of our estimates as local average treatment effects (LATE) requires several IV assumptions, including the existence of instruments with sufficient power in predicting the endogenous regressor. Adding to our discussion, above, of the evidence showing that NCA policies affect physician practice organization, we formally discuss instrument strength in Section 6.4 and show that our instruments exceed typical power thresholds.

The exclusion restriction necessary for the validity of the IVs requires that changes in NCA laws affect physician service prices only through physician market concentration. That is, changes in NCA laws must not be correlated with the error term in the second stage equation. In our structural equation, negotiated prices depend on market concentration and fixed specialty effects, county effects, medical facility type effects, procedure effects, and census-division-by-year effects. By conditioning on this set of covariates, law changes can mechanically only be correlated with the structural error if NCA laws affect negotiated prices across practices *within* a given market, defined by geography and medical specialty, and through some mechanism other than market concentration.

Although an exclusion restriction is not formally testable, we provide evidence supporting its validity in this setting. Using survey data from about 2,000 physicians with information on whether each physician has signed an NCA, linked to negotiated prices with private insurers for the most common office visit procedures, Lavetti et al. (2018) find that the use of NCAs has precisely no effect on negotiated prices conditional on the market and practice size. They find that, within a given geographic market, the standard deviation in negotiated prices across practices for a given procedure is about 39% of the mean price, but the average price difference associated with NCA use is only 2% of the mean and is not statistically significant. In addition, the price difference between NCA users and non-users is no different in higher versus lower NCA enforcement states. To the extent that NCAs affect prices, this evidence suggests that it occurs either across markets or through practice size and concentration measures, which is consistent with the requirements of the exclusion restriction.

The evidence presented in Section 4 on heterogeneity in price effects across procedures is also helpful for considering whether the exclusion restriction assumption is reasonable. One potential concern with this assumption is that changes in NCA laws could directly affect prices by causing labor market frictions that lead to a divergence between the earnings and marginal value products of physicians. However, Figure 3 shows that procedures using primarily physician labor as inputs have little systematic change in prices in response to changes in NCA laws. In contrast, the instruments cause large changes in the prices of procedures that use relatively high amounts of equipment, office space, and non-physician labor inputs. This evidence suggests that the instruments affect prices primarily through a mechanism outside of physicians' labor supply decisions, alleviating some concern about this potential form of violation of the exclusion restriction assumption.

A closer examination of the data also speaks to the assumption of random assignment, supporting the orthogonality of judicial changes on NCA laws to conditions in local physician markets. In addition to showing an absence of pre-trends in the event studies, which supports the notion that judicial decisions were not made in response to trends in physician concentration or prices, it is also informative to analyze the law changes directly. Since judicial decisions are accompanied by opinions written by judges that describe the rationales that led them to their decisions, we can identify the judicial decisions used in our data that were related to physicians and verify that our findings are not sensitive to excluding these events. In Section 6.6 we further examine law changes to verify that they were not systematically related to other state-level political and economic factors that could also affect prices, and find no such patterns.

The final IV assumption is monotonicity. The monotonicity condition in our case requires that any particular law dimension moves HHIs in all states in the same direction. To evaluate this condition we estimate the model using samples split on high versus low enforcement or positive versus negative law changes. The first stage results for these tests are generally consistent with the monotonicity assumption. Under these IV assumptions, each instrument identifies a separate LATE, and our second-stage estimand is an average of these LATEs. As we will show, the seven LATEs are all similar to each other, so this average is informative.

6.4 First-Stage Effects of NCA Laws on HHI

First-stage regression results corroborate the evidence from Section 5 that increases in NCA enforceability lead to reductions in physician market concentration. Table 4 presents estimates from the first-stage models based on employment. The first column shows results from seven separate regressions of establishment-level concentration on each of the instruments. Five of the seven legal indices are statistically significant, and six of the seven have negative coefficients. The dependent variable, HHI, is scaled to range from 0 to 100, so the coefficient on the Burden of Proof Index, for example, suggests that a one unit increase in the index decreases the HHI by 452 points on a 10,000 point scale.

Scaling by the standard deviation of the Burden of Proof Index (0.27) implies that a one standard deviation increase reduces the concentration by about 122 points.

Column 2 presents estimates from a similar specification that includes all 7 seven instruments. The Angrist-Pischke excluded instrument F-statistic is 87, and four of the instruments are statistically significant at the 0.01 level in this model. By comparison, the Stock and Yogo (1997) critical F-statistics thresholds range from about 9 to 12 for achieving 10% relative bias under 2SLS with one endogenous regressor and 3 to 14 instruments. The full table of first-stage results showing both first and second lags of all 7 instruments is shown in Appendix Table A5. When the second lags of each index are used as instruments, the F-statistic is similar (110.45), and when both first and second lags are used the F-statistic is 460, suggesting that any of these choices of lag specifications has sufficient power. In all three specifications, the fixed effects and excluded instruments explain about 75% of the variation in county-specialty-year concentration.

The main first stage IV results using Census data are presented in columns 3 and 4, which correspond to the two jointly-estimated first-stage equations from Section 6. Column 3 shows estimates from the establishment concentration first-stage equation, and column 4 from the firm concentration equation. There are three main points to note about these estimates. First, regarding instrument power, the main limitation of the models estimated using Census data relative to the MPIER estimates is that specialties are not observed, which substantially weakens the first-stage power. However, the F-statistics still suggest that the instruments are not overly weak. Since these models have two endogenous regressors, we report the jointly-estimated Kleinbergen-Paap F-statistic (12.81), which is comparable to the Cragg-Donald F-statistic (297.99) suggested by Stock and Yogo (1997), but is robust to non-independent errors. To corroborate that this reduction in power is caused by unobserved specialty, we re-estimate estimate analogous models in the MPIER data without conditioning on medical specialty and find the instrument strength declines substantially (Appendix Table A7), although the F-statistics are still somewhat higher in the MPIER model.

The first stage parameter estimates themselves are not affected much by the additional controls for firm and insurer concentration from the Census (comparing column 3 to column 2). The only clear exception is the coefficient on the Blue Pencil Index, which was the only positive coefficient in the just-identified first stage results but is insignificant and negative once the additional Census controls are introduced.

Finally, the table shows that the legal indices have different effects on the establishment and firm concentration measures, as can be seen by comparing column 3 to column 4. For example, an increase in the Statutory Index has a negative effect on firm concen-

Table 4: IV First Stage Estimates: Effect of NCA Laws on Employment-Based HHI

Dependent Variable:	Establishment HHI_{t-1} (1)	Establishment HHI_{t-1} (2)	Estab. HHI_{t-1} (3)	Firm HHI_{t-1} (4)
Statutory Index_{t-1}	–2.21	0.42	0.55	–5.72*
	(1.89)	(1.38)	(1.54)	(2.11)
Protectible Interest Index_{t-1}	–2.49	12.16*	14.72*	3.17
	(2.63)	(3.51)	(3.52)	(3.13)
Consideration Index Inception_{t-1}	–5.63*	21.71	17.58*	13.76*
	(0.83)	(38.57)	(6.89)	(6.62)
Consideration Index $\text{Post-Inception}_{t-1}$	–3.08*	–2.46*	–2.38*	1.65*
	(0.41)	(0.34)	(0.59)	(0.48)
Burden of Proof Index_{t-1}	–4.52*	–20.70	–16.47*	–11.50*
	(0.66)	(30.66)	(6.28)	(6.09)
Blue Pencil Index_{t-1}	13.03*	12.56*	–0.21	3.89
	(3.92)	(3.90)	(3.23)	(2.74)
Employer Termination Index_{t-1}	–10.81*	–19.15*	–24.80*	–8.73*
	(1.50)	(4.51)	(3.91)	(3.60)
Insurer HHI_{t-1}			0.00	0.01
			(0.01)	(0.01)
MPIER Data Used	Yes	Yes	Yes	Yes
Census Data Used	No	No	Yes	Yes
N	3,026,780	3,026,780	6,509,400	
N Clusters	121	121	319	
R-Sq		0.75	0.76	
F-Statistic		86.85	12.81	

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. Column 1 reports estimates from separate regressions on each law index, and columns 2-4 report estimates from a single regression with all 7 components. All independent variables are scaled to range between 0 and 1, where 1 is the strongest observed measure of the variable in any state and year in the data. HHIs are all based on employment levels, with establishment HHIs from the CMS MPIER file and firm HHIs from the Census LBD. HHIs are scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. Angrist-Pischke F-Statistic reported in column 2, Kleinbergen-Paap F-statistic reported in columns 3 and 4. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

tration of –5.7 but no effect on establishment concentration. In contrast, the Protectible Interest Index has a significant positive effect on establishment concentration of about 17.6 but very modest effects on firm concentration. This pattern suggests the presence of heterogeneity in the features of the legal indices that affect firm organizational incentives, with some laws having more impact on multi-establishment firm incentives and while others appear to impact the sizes of each establishment.

Table 5 presents first-stage estimates using concentration measures based on sales

Table 5: IV First Stage Estimates: Effect of NCA Laws on Sales-Based HHI

Dependent Variable:	Estab. HHI_{t-1} (1)	Firm HHI_{t-1} (2)
Statutory $Index_{t-1}$	–0.25	–3.09
	(1.48)	(2.33)
Protectible Interest $Index_{t-1}$	14.11*	7.23*
	(3.36)	(3.56)
Consideration Index $Inception_{t-1}$	17.26*	22.56
	(6.94)	(13.17)
Consideration Index $Post\text{-}Inception_{t-1}$	–2.18*	2.79*
	(0.59)	(1.01)
Burden of Proof $Index_{t-1}$	–16.15*	–19.49
	(6.33)	(12.31)
Blue Pencil $Index_{t-1}$	–0.15	0.53
	(3.13)	(3.49)
Employer Termination $Index_{t-1}$	–24.06*	–10.28*
	(3.77)	(4.31)
Insurer HHI_{t-1}	0.00	–0.04
	(0.01)	(0.02)
MPIER Data Used	Yes	Yes
Census Data Used	Yes	Yes
N	6,329,900	
N Clusters	319	
R-Sq	0.83	
F-Statistic	13.56	

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. Legal indices are scaled to range between 0 and 1, where 1 is the strongest observed measure of the variable in any state and year in the data. Firm HHIs are based on sales from the Census LBD and SSEL, and establishment HHIs are based on employment levels from MPIER. HHIs are scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors are clustered by state-year. Kleinbergen-Paap F-statistic reported. * indicates significance at the 0.05 level.

data from the Census LBD, Economic Census, and SSEL, instead of on employment. There are two main conclusions from these results. First, all seven of the coefficients in column 2 have the same sign as the coefficients in column 4 of Table 4, suggesting that the instruments have similar effects on both concentration measures. This consistency is reassuring, since there is little external evidence on the comparability of sales-based and employment-based concentration measures generally. Second, the instruments are slightly more powerful in the sales-based model, and the F-statistics remain above conventional weak-instruments thresholds.

6.5 The Effect of HHI on Negotiated Prices

Our main estimates are reported in Table 6. The top panel of the table presents results using the sales-based firm concentration measure. Column 1 uses first lags of each instrument, corresponding to the first stage estimates in Table 5. The coefficient on firm concentration of 0.02 implies that a 100 point increase in firm HHI, holding fixed both the establishment concentration and insurer concentration, causes a 2% increase in negotiated prices on average. This result is consistent with multi-establishment growth improving bargaining position relative to insurers. In contrast, the coefficient on the establishment concentration, –0.014, implies that holding firm concentration fixed but increasing the establishment HHI by 100 points leads to 1.4% lower prices.

These estimates suggest that the efficiency gains of larger group practices at a given location outweigh any effects of practice size on the bargaining position of physicians. However, consolidation of multi-site physician groups increases the insurance network value of the firm as a whole, and more than offsets any impacts of economies of scale. We elaborate on the intuition behind this result in a model of physician-insurer bargaining presented in Appendix 8.

The coefficient on insurer concentration is modest, 0.0007, although to be clear since insurers do not tend to use NCAs the law change events do not affect this variable, and the coefficient is identified only by the small intertemporal changes in insurer concentration that are not absorbed by county effects and census division by year effects. In contrast, previous studies that use more substantial sources of variation in insurer HHIs suggest that insurance market concentration plays an important role in affecting prices (Dafny et al. (2012)). We include this term only as a control variable, and we caution against the interpretation that insurance market concentration does not affect negotiated prices, since our identifying variation for this coefficient is potentially too small to be salient for bargaining, and since we do not have an instrument for insurance market concentration.

Column 2 of the table reports estimates using both first and second lags of each law index and yields qualitatively similar patterns. The coefficient on firm concentration declines to 0.01 but is still statistically significant, and the coefficient on establishment concentration remains similar, –0.013. To highlight the importance of addressing endogeneity in physician concentration, we also report the starkly different OLS estimates from the same sample: 0.0001 for both concentration measures. OLS estimates close to zero are consistent with evidence from previous studies using either cross-sectional variation or panel variation in an OLS specification (Dunn and Shapiro (2014) and Baker et al. (2014)).

The bottom panel of Table 6 presents corroborating evidence using employment-based

Table 6: Main Estimates: Effect of Market Concentration on Negotiated Prices

	Dependent Variable: ln($Price$)		
	IV First Lags	IV Both Lags	OLS
	(1)	(2)	(3)
Physician Firm HHI, Sales-Based	0.020*	0.010*	0.0001*
	(0.009)	(0.004)	(0.0000)
Physician Establishment HHI	–0.014*	–0.013*	0.0001*
	(0.006)	(0.005)	(0.0000)
Insurer HHI	0.0007	0.0003	–0.0001
	(0.0006)	(0.0004)	(0.0003)
N	6,329,900	6,329,900	6,329,900
N Clusters	319	319	319
F-Stat (Cragg-Donald)	270.27	143.79	
F-Stat (Kleinbergen-Paap)	13.56	10.32	
Physician Firm HHI, Employment-Based	0.016*	0.010*	0.0001*
	(0.007)	(0.005)	(0.0000)
Physician Establishment HHI	–0.014*	–0.017*	0.0001*
	(0.005)	(0.005)	(0.0000)
Insurer HHI	0.0000	0.0000	0.0000
	(0.0002)	(0.0002)	(0.0003)
N	6,509,400	6,509,400	6,509,400
N Clusters	319	319	319
F-Stat (Cragg-Donald)	297.99	174.35	
F-Stat (Kleinbergen-Paap)	12.81	10.44	

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. Data for physician firm and insurer HHIs in these regressions come from the Census' LBD (employment) and SSEL (sales). Physician establishment HHIs are from MPIER. HHIs are scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. Insurer HHIs are calculated from firm-level in-state sales. Medical specialties are observed in price data but not in Census data used to calculate physician HHIs. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

concentration measures. The estimates are again statistically significant and imply that a 100 point increase in firm HHI, conditional on establishment and insurer concentration, increases negotiated prices by about 1.0% to 1.6%, while the same size increase in establishment concentration decreases prices by 1.4% to 1.7%. The estimates across the two panels are consistent with the evidence from the first stage models that the sales-based and employment-based concentration measures yield similar results.

Since Census Bureau confidentiality restrictions impose cell size restrictions for every combination of subsamples across any analyses conducted, the disclosure requirements

Table 7: OLS and IV Second Stage: Effect of Establishment-Based Market Concentration on Prices

	Dependent Variable: $\ln(Price)_t$			
	IV	IV	IV	OLS
	(1)	(2)	(3)	(4)
HHI_{t-1}	–0.0283*	–0.0235*	–0.0251*	0.0002*
	(0.0056)	(0.0045)	(0.0047)	(0.0000)
Instruments	First Lags	Second Lags	Both Lags	
N	3,026,780	3,026,780	3,026,780	3,026,780
N Clusters	121	121	121	121
R-Sq	0.97	0.98	0.98	0.82
1st Stage F-Stat	86.85	110.45	460.22	

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors are clustered by state-year. Angrist-Pischke F-Statistic reported. * indicates significance at the 0.05 level.

grow exponentially with the number of samples used, making it impossible for us to use Census data to conduct a wide range of robustness analyses. Instead we rely on the MPIER data to conduct these analyses. The associated models therefore have one first-stage equation corresponding to EC, excluding FC and $InsC$ from Equation 2.

Table 7 presents IV estimates using only the MPIER data, where firm concentration is not observable. The results suggest that a 100 point increase in the establishment concentration leads to a 2.4% to 2.8% reduction in average negotiated prices, somewhat larger than in the estimates with Census data that control for firm and insurer concentration. The estimated effect is on the upper end of this range when first lags are used as instruments and on the lower end when second lags are used.

Since the majority of our robustness analyses can only be conducted using MPIER data, we first seek to understand why the MPIER estimates differ somewhat from the Census estimates. To that end, we collapse the MPIER HHI measures as though physician specialties were unobserved, and we re-estimate the IV models. The results, shown in Appendix Table A7, suggest that the effect of establishment concentration is between –0.011 and –0.015 in all specifications, very similar to Census estimates when the data structure is made comparable. We also estimate our main MPIER concentration specifications including the Census insurer concentration control but no firm concentration, and we find that it does not substantively alter those estimates either. These results provide some reassurance that the robustness analyses using MPIER data are relevant to our main estimates.

Returning to the discussion of the exclusion restriction from Section 6.3, one additional piece of evidence in support of this restriction comes from the consistency of estimates when we estimate the IV model using only one legal index at a time, shown in Table 8. Column 1 of the table presents second-stage estimates from 7 separate just-identified IV regressions using only the first-lags of each instrument, one at a time, and column 2 presents estimates using first and second lags of each index. All 14 models yield negative coefficients on establishment concentration, and 10 of the 14 estimates are statistically significant.

This result is reassuring because if the exclusion restriction were violated due to a direct effect of the instruments on practice cost functions conditional on practice size, the differences in the legal nature of the instruments would likely cause heterogeneity across instruments in the second-stage estimates. For example, whereas the Consideration Index changes the way employment contracts are written by affecting whether compensation for NCAs must be explicit, it is less clear that law dimensions such as the Burden of Proof Index or the Blue Pencil Index would impact practice cost functions. Both of these dimensions relate to the specific procedures used during the litigation of NCA contracts and become relevant when a job ends and litigation occurs. The Burden of Proof index could violate the exclusion restriction if, for example, the cost to the firm of producing evidence for the litigation affected prices negotiated with insurers. Similarly, the Blue Pencil Index could lead to a violation if the ability of a judge to ex post adjust the terms of a contract that was operable throughout an employment spell had a direct effect on prices. The consistency of estimates over a range of instruments, each of which has unique and distinct legal mechanisms for affecting organizational incentives, provides some reassurance that a potential violation of the exclusion restriction for any one legal measure is unlikely to drive the overall pattern of results.

Taken together, our results from MPIER and Census data suggest that the effects of consolidation on prices depend on a tradeoff between the efficiency gains of larger establishments and the increased negotiating power associated with bargaining as a larger organization. Larger establishments allow efficiency gains via economies of scale that appear to dominate any effects of bargaining position, leading negotiated prices to fall. These economies of scale can arise, for example, when physician practices share equipment, information systems, laboratory facilities, nurses, and technical and administrative staff over a larger number of physicians and patients. The contrasting estimates from

Table 8: IV Results Estimated Separately by Law Component

Dependent Variable:	ln(*Price*)	
Instruments:	First Lags	Both Lags
Statutory Index	–0.047	–0.029*
	(0.061)	(0.015)
	[1.37]	[3.58]
Protectible Interest Index	–0.003	–0.019
	(0.051)	(0.011)
	[0.90]	[3.65]
Consideration Index Inception	–0.043*	–0.032*
	(0.011)	(0.007)
	[45.68]	[246.67]
Consideration Index Post-Inception	–0.020*	–0.021*
	(0.009)	(0.008)
	[55.70]	[33.66]
Burden of Proof Index	–0.043*	–0.031*
	(0.011)	(0.006)
	[46.98]	[260.14]
Blue Pencil Index	–0.007	–0.008*
	(0.005)	(0.003)
	[11.05]	[24.31]
Employer Termination Index	–0.024*	–0.024*
	(0.006)	(0.006)
	[52.04]	[26.08]

Notes: Each cell shows the second stage IV estimate of the effect of lagged HHI on log prices using a single legal component as the instrument. The first column displays just-identified models using the first lag of each index. The second column includes both the first and second lags of the legal component as instruments. All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors, in parentheses, are clustered by state-year. First-stage Angrist-Pischke F-statistics are shown in brackets. * indicates significance at the 0.05 level.

the firm-level component of the variation, however, suggest that consolidation of multi-establishment firms increases the bargaining position of firms by more than any efficiency gains, leading to higher negotiated prices. This result is consistent with evidence from Dranove and Lindrooth (2003), and with the notion that most of the efficiency gains from larger physician firms come from increases in practice size *at a given location*; meanwhile, consolidation across locations has smaller efficiency gains but still affects bargaining position in negotiation, causing a net positive effect on prices.

6.6 Heterogeneity and Robustness

In this section we provide a concise overview of many supplemental analyses conducted to assess the robustness of our results to model assumptions and to potential data measurement concerns.

Heterogeneity by Procedure Type: Returning to the analysis in Section 4 showing

Table 9: IV Estimates by Tercile of Medicare Practice RVUs ÷ Physician Labor RVUs

	Dependent Variable: $\ln(Price_t)$		
Tercile	Lowest	Middle	Highest
$HHI_{(t-1)}$	-0.003	0.002	-0.016*
	(0.004)	(0.002)	(0.005)
N	909,540	759,617	793,161
1st Stage F-Stat	482.2	360.0	347.3

Notes: Specifications are the same as in Table 7, but estimated separately by tercile of Medicare facility practice expense RVUs divided by physician work RVUs. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors are clustered by state-year. Angrist-Pischke F-Statistic reported. * indicates significance at the 0.05 level.

that reduced-form price effects were largest for procedures with high facility practice RVUs relative to physician labor RVUs, we estimate the corresponding IV model. Consistent with the reduced-form evidence, Table 9 shows that the IV estimates are driven primarily by procedures in the highest tercile, with no significant effects of concentration on prices for procedures that are relatively more intensive in the use of physician labor.

Table 10: Heterogeneity by Procedure Type

	Dependent Variable: $\ln(Price_t)$					
					Imaging Subcategories	
	Evaluation	Tests	Procedures	Imaging	Technical Equipment	Professional Services
$HHI_{(t-1)}$	-0.0002	-0.0471*	-0.0159*	-0.0138	-0.0080	0.0177*
	(0.0028)	0.0088	(0.0063)	(0.0084)	(0.0084)	(0.0080)
N	1,776,856	1,044,303	218,887	318,383	226,708	48,473
F-Stat	440.2	335.5	30.0	118.1	55.6	77.9

Notes: Specifications are the same as in Table 7, but estimated separately by procedure type, as defined by Betos categories. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors are clustered by state-year. Angrist-Pischke F-Statistic reported. * indicates significance at the 0.05 level.

Table 10 presents similar evidence on heterogeneity in the effect of concentration on prices by procedure type. The table presents estimates for procedures by 'Betos' category, defined by Berenson and Holahan (1990) and used by Clemens, Gottlieb, and Molnar (2017) to study heterogeneity in private benchmarking to Medicare prices. We find that higher concentration leads to lower negotiated prices for tests and procedures—which tend to use more practice overhead and equipment resources—but not for evaluations, which primarily use physician labor. Greater concentration is also associated with lower prices for imaging procedures, although the effect is insignificant.

One convenient feature of imaging procedures is that procedure codes and prices are set separately for the equipment and physician time components of the cost. Separating these two subcategories reveals that there is a positive effect of concentration on prices for procedures that use only physician labor, where economies of scale are less likely to be present. This finding is also consistent with Clemens, Gottlieb, and Molnar (2017), who show that deviations from Medicare benchmark prices are more likely to be positive for labor-intensive procedures.

Market Structure Measure: Although our main estimates rely on HHIs, the most commonly used measure of market concentration in the literature (Gaynor et al. (2015)), interpreting these estimates as elasticities of demand requires the potentially undesirable assumptions that goods are homogeneous and firms engage in Cournot competition, as discussed in Section 6.3. Since we cannot estimate firm conduct directly without detailed claims data, we test the sensitivity of our estimates to these assumptions by re-estimating the model using the negative log HHI transformation, average establishment size, 4-firm market share, and 8-firm market share.

Table 11 shows that the qualitative conclusions are identical for all of these choices of market structure. In the negative log HHI specification, the sign is positive (which is consistent since the measure is negated), and the bias relative to the OLS specification goes in the same direction. When average establishment size is used we find that increasing the average number of physicians in a practice by one reduces negotiated prices by about 3.4% to 4.3%. Similarly, in markets that become more concentrated in terms of the market shares of the 4 largest or 8 largest establishments, average negotiated prices fall significantly. Across the variety of market structure measures and instrument specifications, we conclude that there is a statistically significant negative relationship between market concentration and negotiated prices in all fifteen models tested.

Geographic Market Definition: Although county is a commonly used market definition (See Baker et al. (2015), Schneider et al. (2008),) we also test whether the results are

Table 11: Alternative Measures of Market Concentration (Establishment-Based)

	Dependent Variable: $\ln(Price)_t$		
Instruments:	IV First Lags	IV Both Lags	OLS
Negative Log $\text{HHI}_{(t-1)}$	0.190*	0.283*	0.004*
	(0.084)	(0.092)	(0.001)
1st Stage F-Stat	[80.33]	[2043.69]	
Mean Establishment $\text{Size}_{(t-1)}$	–0.043*	–0.034*	0.0003*
	(0.014)	(0.013)	(0.0000)
1st Stage F-Stat	[245.13]	[652.49]	
4-Firm Market $\text{Share}_{(t-1)}$	–0.021*	–0.022*	–0.0001
	(0.007)	(0.006)	(0.0001)
1st Stage F-Stat	[21.96]	[11.81]	
8-Firm Market $\text{Share}_{(t-1)}$	–0.030*	–0.029*	–0.0001
	(0.011)	(0.010)	(0.0001)
1st Stage F-Stat	[13.96]	[4.95]	

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. All concentration measures are calculated from establishment sizes in MPIER data, provided by CMS. 4-Firm and 8-Firm Market Shares are measured from 0 to 100. Angrist-Pischke F-Statistics reported in brackets. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

sensitive to this choice. Market definition is often a crucial assumption in evaluating policies aimed at ensuring sufficient *levels* of competition, but since we rely on changes in concentration within markets, our estimates do not appear to be very sensitive to the assumption of market definition. The magnitudes of our estimates are very stable when using either smaller or larger market definition assumptions. Table 12 presents estimates of the main specification using counties, hospital service areas (HSAs), and primary care service areas (PCSAs) as potential market definitions. HSAs are defined by the Dartmouth Atlas of Healthcare using data on patient locations and their choices between hospitals. We chose HSAs as a plausible upper bound on the size of markets, since patients tend to travel farther on average to hospitals than they do for ambulatory physician visits. PCSAs are similarly defined by the Dartmouth Atlas but are based on choices of primary care physicians only. Since patients tend to travel farther to visit specialists than they do to visit primary care physicians, PCSAs are likely to be smaller on average than the appropriate overall market definition for physicians.

The estimates are very similar for all three market definitions, ranging from -2.3% in HSAs to -2.9% in PCSAs. This conclusion also holds within every physician specialty

Table 12: Sensitivity of MPIER IV Estimates to Market Definition

	Dependent Variable: ln($Price$)		
	County	HSA	PCSA
	Full Sample		
$\text{HHI}_{(t-1)}$	–0.025*	–0.023*	–0.029*
	(0.005)	(0.008)	(0.008)
1st Stage F-Stat	[460.2]	[2119.6]	[270.4]
	Primary Care		
$\text{HHI}_{(t-1)}$	–0.017*	–0.018*	–0.024*
	(0.005)	(0.005)	(0.004)
1st Stage F-Stat	[542.3]	[63.1]	[122.2]

Notes: All specifications include fixed effects for the corresponding geographic market, census division by year, procedure code (CPT), physician specialty, and facility type. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHIs are scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. In each specification the instruments include all lagged and twice lagged law components, as in column (3) of Table 7. Standard errors, in parentheses, are clustered by state-year. First-stage Angrist-Pischke F-statistics are reported in square parentheses. * indicates significance at the 0.05 level.

group, and there is a general pattern that effects are slightly larger in magnitude in smaller PCSA markets. Overall, we conclude that market definition assumptions do not substantively alter our conclusions.

Sensitivity to Large NCA Law Changes: Figure A1 shows that a small number of law changes are of much larger magnitude than the average change. Appendix Table A8 presents estimates in which we drop outlier states with very large NCA law changes. The estimates are very similar, remain statistically significantly different from zero, and are not significantly different from each other. The first-stage power increases slightly in all three specifications.

Treatment of Multi-Specialty Practices: Defining markets by specialty involves assumptions about how to treat physicians in multi-specialty practices. For example, when defining a market for orthopedists, how should one treat practices that contain orthopedists as well as radiologists? One approach is to ignore radiologists altogether, and only consider the market shares of orthopedists in the geographic market. However, an insurer concerned about the negative consequences of failing to reach an agreement with such a practice may care about the consequences of losing both the orthopedists and the radiologists. Our main specifications calculate HHIs using all physicians in any practice containing at least one physician in a given specialty. In Appendix Table A9 we consider 4 different possible sets of assumptions about the treatment of multispecialty practices in measuring concentration. The estimates are similar under every alternative assumption tested.

Heterogeneity by Specialty Type: As discussed in Section 2, there may be heterogeneity in the usage and value of NCA contracts for physicians with different specialties. Since the benefits of NCAs in protecting the value of patient stocks is likely to be largest for physicians with many repeated interactions with the same patients, hospital-based practices that employ specialists like surgeons may derive less value from NCAs. This

Table 13: Effect of Concentration on Prices by Medical Specialty

Dependent Variable:		$\ln(Price_t)$	
	Primary Care Physicians	Non-Surgical Specialists	Surgical Specialists
	(1)	(2))	(3)
HHI_{t-1}	–0.017*	–0.013*	–0.005
	(0.005)	(0.005)	(0.005)
N	473,033	300,990	272,913
1st Stage F-Stat	542.3	98.3	17.2

Notes: Specifications are the same as Table 7, using both instrument lags. 'Primary Care Physicians' includes Internal Medicine, Family Practice, Geriatric Medicine, and Pediatric specialists. 'Non-Surgical Specialists' includes Proctology, Urology, Dermatology, Cardiovascular Dis/Cardiology, Neurology, Gastroenterology, and Hematology. 'Surgical Specialists' includes General Surgery, Neurological Surgery, Orthopaedic Surgery, Thoracic Surgery, Anesthesiology, and Radiology. Angrist-Pischke F-statistics reported. Standard errors are clustered by state-year. * indicates significance at the 0.05 level.

notion is consistent with the findings of Lavetti, Simon, and White (2018) that hospital-based physicians are significantly less likely to have signed NCA contracts. In Table A13 we present results from subsamples of the data that are split by groups of medical specialties. The results suggest broad and consistent effects across both primary care and specialist physicians that are less likely to be hospital-based. We estimate that a 100 point increase in HHI reduces prices by 1.7% to 2.4% for primary care physicians and by 1.3% for non-surgical specialists. In contrast, for surgical specialists we find no significant price effects and the instruments are much weaker, consistent with these physicians being less likely to use NCAs.

Interactions between Physician and Insurer Concentration: Our main results in Table 6 include controls for insurer HHI, which have little effect on our estimates. This result is surprising given the previous literature, such as Dafny et al. (2012), which shows that insurer concentration is an important determinant of market outcomes. One important limitation to our findings is that the effect of insurer concentration on prices is identified only by year-to-year changes in insurer concentration, which may be both small in magnitude and endogenous, and for which we do not have an instrument. Taking another approach to understand how insurer concentration may affect our results, we re-estimate the MPIER model specification including interactions between physician establishment HHI and categories of insurer HHI using 2007 data on insurers from the American Medical Association. Appendix Table A12 corroborates our main results, showing that the effect of establishment concentration on prices has very little sensitivity to insurer HHI—the coefficient estimates remain between –0.023 and –0.025 in markets in which the insurer concentration level is below 2500, between 2500-5000, or above 5000. One potential explanation that could reconcile this finding with the literature is if the price effects of insurer concentration tend to be driven largely by the impact of disagreement on insurer profits (the first two terms in Appendix Equation 3 of the bargaining model), while the coefficient we estimate is driven primarily by the impact of practice costs on prices. If insurer concentration is not strongly correlated with the physician practice bargaining weight, it is conceptually possible for the two sets of findings to be consistent. This possibility is made more likely by the inclusion of county effects and census division by year effects, which absorb much of the impact of insurer concentration on price levels.

Heterogeneity in Urban and Rural Markets: One may expect NCA laws to affect practice organization and prices differently in urban versus rural areas, since urban markets tend to have lower levels of baseline concentration. Policy-makers, who have been increasingly concerned about inadequate supply of physicians in rural markets, are likely to also want to know the extent to which consolidation is occurring and affecting prices in these areas. On the other hand, the consequences of non-compete laws may be entirely an urban phenomenon.

Table 14 presents estimates splitting the sample by metro and non-metro counties. In metro counties the instruments are strongest, and we find that a 100 point increases in establishment HHI causes a 3.1% to 3.4% decline in negotiated prices. In non-metro counties the effect is somewhat smaller, 1.1% to 1.3%, but still statistically significant despite somewhat weaker instruments. This pattern is potentially consistent with greater economies of scale in metro markets, where input factor prices such an nursing and staff labor, rent, and equipment costs tend to be higher. The weaker first-stage in non-metro markets may also be potentially explained by the lower rate use of NCA contracts among physicians in rural markets documented by Lavetti, Simon, and White (2018). In most rural markets in the US the supply of physicians per capita is much lower, suggesting that physician groups may derive less value from using NCAs to protect patient stocks.

There is a large literature studying the distinction between urban and rural markets in

Table 14: Effect of Concentration on Prices in Urban and Rural Counties

	Dependent Variable: $\ln(Price)_t$					
	All Counties		Metro Counties		Non-Metro Counties	
Instruments	First Lags	Both Lags	First Lags	Both Lags	First Lags	Both Lags
	(1)	(2)	(3)	(4)	(5)	(6)
HHI_{t-1}	–0.028*	–0.025*	–0.034*	–0.031*	–0.013*	–0.011*
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)
N	3,026,780	3,026,780	2,077,627	2,077,627	949,153	949,153
1st Stage F-Stat	86.9	460.2	54.7	364.3	15.4	15.1

Notes: All specifications are identical to those in Table 7. Angrist-Pischke F-statistics reported. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

understanding the relationship between concentration and prices. Whereas a longstanding approach was to use patient flows to define relevant markets, Capps, Dranove, and Satterthwaite (2003) showed that this approach can be misleading particularly in metro areas, where the willingness of some patients to travel long distances to access lower prices may not indicate that all patients are similarly willing to travel, potentially understating local market power. This intuition may suggest that one would expect to see more positive price effects of a given change in concentration in metro areas. However, it is worth noting that the geographic distribution of willingness to pay affects prices in a bargaining framework through terms that are distinct from the cost function term, which we find dominates these other bargaining terms. Nonetheless, any attenuation in the positive impact of bargaining position on prices in metro areas could potentially cause the combined coefficient to be more negative, consistent with the pattern in Table 14, and we cannot distinguish this explanation from the possibility that economies of scale are larger in cities. Reassuringly, our estimates are similar when using market definitions based on patient flows (PCSA and HSA) or based on counties.

Balanced Panel: The sample size of the MarketScan price data increases over time. To test whether the imbalance in our panel caused by sample growth affects our baseline results, we re-estimate the model using only the subset of county-specialty pairs for which we have price data in all 12 years of our panel. The IV estimates, shown in Appendix Table A14, are similar but slightly larger in the balanced panel, –0.032 to –0.036, which is partially caused by the balanced panel containing a higher proportion of metro counties.

MPIER Fuzzy Matching Algorithm and Measurement Error: There are a few types of assumptions necessary to construct HHIs from the raw MPIER data. First, some addresses are missing, so we test the sensitivity of estimates to the treatment of missing addresses. To bound the effects, we estimate the main specification under the assumption that all missing addresses are separate solo practices, and again under the assumption that all missing addresses belong to a single practice. Appendix Table A9 presents estimates under each of these assumptions, interacted with the assumption about treatment of multi-specialty practices. In all eight specifications the estimates are statistically significant and negative and have large F-statistics of at least 460. The second stage estimates range from –0.014 to –0.025, and all 8 are statistically significant at the 0.05 level.

Second, the association of addresses to practices requires an assumption about the tolerance in the fuzzy matching algorithm. The algorithm allows characters in the addresses to be slightly different, to allow for typographic errors and abbreviations, while forcing numerical elements of the addresses to be exactly identical (that is, street numbers and office numbers must match exactly.) We use the normalized Levenshtein distance as a metric for the distance between all combinations of character subsets of addresses in the same zip code. Appendix Table A15 presents estimates from the main specification by re-calculating HHIs under alternative fuzzy matching thresholds that allow for stricter or more flexible matching of addresses. Smaller distance thresholds result in smaller average establishment sizes by forcing addresses to almost exactly match, while the opposite is true for larger thresholds. The results are not at all sensitive to this tolerance parameter, ranging from –0.025 to –0.026 (SE 0.005) in all nine specifications.

Exogeneity of NCA Law Changes: Using law changes as a source of identification generally raises the concern that the laws may not be exogenous to the outcome being investigated. The inclusion of county effects in our specifications removes average differences that may affect both NCA laws and outcomes, so our concern is limited to covariation within states over time. This could occur, for example, if political or economic environment that generated the law changes also affected the outcome of interest, potentially through other correlated laws, or through intermediate factors other than physician market concentration.

We test for evidence that NCA law changes are correlated with a variety of economic outcomes as well as state residents' subjective views from the Generalized Social Survey (GSS) on a variety of political, economic, and cultural topics and correlate them with the law changes. Appendix Table A17 shows that log payroll per worker, unemployment rates, and population are all uncorrelated with the law changes (columns 1-3). Politically, the share of votes to Republican candidates in presidential and congressional elections is also uncorrelated with the law changes (column 4).

Appendix Table A18 presents tests of correlations between law changes and GSS survey responses. The first five columns relate to the respondent's views on size of government and spending on social issues, such as cities, welfare, and medical care. The last two columns reflect the respondent's political identification and financial satisfaction, respectively. The law changes appear uncorrelated with views captured in the GSS; only one of 49 coefficients in the table is significant at the 5% level, suggesting that NCA laws are not systematically driven by or correlated with important changes in the local political or economic climate.

7 Discussion

This paper makes three main contributions towards understanding the relationships between physician practice organization and negotiated prices with private insurers in the US. First, we address several important data limitations that have impeded research on this topic. We build on existing work on physician markets by employing two comprehensive longitudinal data sets on physicians: one from CMS covering all physicians and practices in the US, and a second confidential database from the Census Bureau containing firm linkages for all multi-establishment practices using IRS tax IDs, and providing sales and payroll for every physician firm in the US. By linking these sources to a longitudinal database of negotiated prices between physicians and private insurers, we create a comprehensive new database with which to study physician markets, spanning virtually all markets in the country over 12 years. In addition to its breadth, this database has the advantage that it includes total sales of physician firms from all sources.

Second, we highlight the important role states play in affecting physician service prices through NCA policies. We show that even modest increases in NCA enforceability lead to meaningful increases in physician prices. As a rough back-of-the-envelope calculation, and abstracting from general equilibrium effects, our estimates suggest that if NCA enforceability decreased nationally by 0.1 units of the NCA Index, total physician spending would fall by about 3.7%—over $20 billion annually based on 2015 spending levels.[12] Yet 39 states have never legislatively chosen an NCA policy and instead leave the decisions to the judicial branch, in which common law traditions shape current policies. Our findings suggest that substantial value may arise from states conducting comprehensive assessments of NCA laws and actively legislating policies, drawing on the expanding research studying the impacts of NCAs.

Finally, we evaluate the validity of using judicial decisions that change NCA policies

[12]This calculation is based on our reduced-form estimate that a 0.1 unit increase in NCA enforceability led to a 3.7% increase in prices, and assumes an elasticity of demand for medical care of –0.2. Scaling a 3.7% price increase by quantity gives $\%\Delta Q = -0.2 * 3.7\% = -0.74\%$ and $\%\Delta PQ = 3.7\% * (1 - 0.74\%) = 3.67\%$, approximately the same as the percentage change in price.

as instruments for the potentially endogenous variation in physician practice organization and market concentration. After presenting evidence consistent with the IV assumptions, we use these instruments to estimate the effect of physician market concentration on negotiated prices. Our results highlight an important distinction between economies of scale in physician practices and the effect of larger practices on bargaining position. We find that when physician *establishments* grow larger, economies of scale dominate the effect of bargaining position on prices, leading to a net reduction in prices of about 1.3% to 1.7% per 100 unit increase in HHI. However, when physician *firms* grow larger conditional on establishment concentration the opposite is true—a 100 point increase in HHI increases prices by about 1% to 2%, suggesting that any associated economies of scale are outweighed by the effects of firm consolidation on bargaining position. These results have important implications for policies aimed at protecting competition in physician markets, suggesting that practice mergers that coincide with physical consolidation may be more likely to lead to lower prices. They also highlight the importance of measuring both establishment and firm sizes for understanding the impacts of practice organization on prices.

As a matter of interpretation, one question that we cannot fully address in our analyses is whether the estimated changes in concentration and prices are good or bad for consumers. Consolidation of multi-establishment practices may improve geographic access or other aspects of medical care that consumers value. Similarly, if multi-establishment consolidation causes price increases by affecting the bargaining weights of physicians relative to insurers, such price increases may be of less concern to antitrust regulators than if they were caused by changes in bargaining threat points. Interpretation of our estimates further depends on the margin of variation we use, which may be unique relative to patterns of consolidation in physician markets more generally. Our estimates are local average treatment effects driven by responses to changes in NCA enforceability, and the margin around which we identify effects on prices may differ from the margin that has prompted the recent trend of hospital acquisitions of physician groups, for example. More research is necessary to extend our findings before drawing conclusions about welfare effects.

References

[1] Anderson, G., Reinhardt, U., Hussey, P., and Petrosyan, V. (2003), "It's the Prices, Stupid: Why the United States is so Different from Other Countries," *Health Affairs*, Vol. 22, No. 3 (2003), pp. 89-105.

[2] Baker, Lawrence, M. Kate Bundorf, and Anne Royalty (2014), "Effects of Physician Practice Consolidation on Prices for Physician Services," Working Paper.

[3] Berenson, Robert A., and John Holahan (1990) "Using a New Type-of-Service Classification System to Examine the Growth in Medicare Physician Expenditures, 1985-1988". Urban Institute, Health Policy Center, 1990.

[4] Bishara, Norman (2011), "Fifty Ways to Leave Your Employer: Relative Enforcement of Noncompete Agreements, Trends, and Implications for Employee Mobility Policy," *University of Pennsylvania Journal of Business Law*, 13 (forthcoming).

[5] Bresnehan, Timothy F (1989), "Empirical Studies of Industries with Market Power. In Handbook of Industrial Organization, Volume 2, edited by Richard Schmalensee and Robert Willig, 101157. Amsterdam and San Diego: Elsevier, North-Holland.

[6] Capps, Cory, David Dranove, and Mark Satterthwaite (2003), "Competition and Market Power in Option Demand Markets," *RAND Journal of Economics*, Vol. 34, No. 4, (2003), pp. 737-63.

[7] Choudhry, Niteesh, R. Fletcher, and S. Soumerai (2005), "Systematic Review: The Relationship between Clinical Experience and Quality of Health Care," *Annals of Internal Medicine*, Vol 142. No. 4, (Feb. 2005), pp. 260-273.

[8] Clemens, Jeffrey and Joshua Gottlieb (2016), "In the Shadow of a Giant: Medicare's Influence on Private Payment Systems," forthcoming at *Journal of Political Economy*, 2016.

[9] Clemens, Jeffrey, Joshua Gottlieb, and Timea Molnar (2017), "Do Health Insurers Innovate? Evidence from the Anatomy of Physician Payments," *Journal of Health Economics*, Vol. 55 (2017), pp. 153-67.

[10] Cooper, Zack, Stephen Gibbons, Simon Jones, and Alistair McGuire (2011), "Does Hospital Competition Save Lives? Evidence From the English NHS Patient Choice Reforms" *The Economic Journal*, Vol. 121, No. 554, pp. F228-F260.

[11] Dafny, Leemore (2010), "Are Health Insurance Markets Competitive?" *The American Economic Review*, 100: 1399-1431.

[12] Dafny, Leemore, Mark Duggan, and S. Ramanarayanan (2012), "Paying a premium on your premium? Consolidation in the US Health Insurance Industry" *The American Economic Review*, 102(2): 1161-1185.

[13] Dranove, David and Richard Lindrooth (2003), "Hospital Consolidation and Costs: Another Look at the Evidence," *Journal of Health Economics*, Vol. 22 (2003), pp. 983-97.

[14] Dunn, Abe and Adam Hale Shapiro (2014), "Physician Market Power and Medical-Care Expenditures," *Journal of Law and Economics*, 57(1).

[15] Ericson, Keith and Amanda Starc (2012), "Pricing Regulation and Imperfect Competition on the Massachusetts Health Insurance Exchange," NBER Working Paper No. 18089.

[16] Fallick, Bruce, Charles A. Fleischman, and James B. Rebitzer (2006), "Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster," *The Review of Economics and Statistics*, 88(3): 472-481.

[17] Garmaise, Mark (2009), "Ties that Truly Bind: Non-Competition Agreements, Executive Compensation, and Firm Investment," Journal of Law, Economics, and Organization, Vol. 27, No. 2, (August) pp. 376-425.

[18] Gaynor, Martin, and W.B. Vogt (2003) "Competition among hospitals," *RAND Journal of Economics*, 34: 764-785.

[19] Gaynor, Martin, Kate Ho and Robert J. Town (2015), "The Industrial Organization of Health-Care Markets," *Journal of Economic Literature*, 53(2): 235-284.

[20] Gowrisankaran, G., Aviv Nevo, and Robert Town (2013) "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry," NBER Working Paper No. 18875.

[21] Gunning, Timothy and Robin Sickles (2013) "Competition and Market Power in Physician Private Practices," *Empirical Economics*, Vol. 44, No. 2, (Apr. 2013), pp. 1005-29.

[22] Ho, Kate and Robin Lee (2014) "Insurer Competition and Negotiated Hospital Prices," Working Paper.

[23] Kleiner, Samuel A., William D. White, and Sean Lyons (2015), "Market Power and Provider Consolidation in Physician Markets," *International Journal of Health Economics and Management*, Vol. 15, No. 1 (March, 2015), pp. 99-126.

[24] Lavetti, Kurt, Carol Simon, and William D. White (2018), "The Impacts of Restricting Mobility of Skilled Service Workers: Evidence from Physicians," *Working Paper*.

[25] Lewis, Matthew and Kevin Pflum, (2015) "Diagnosing Hospital System Bargaining Power in Managed Care Networks," *American Economic Journal: Economic Policy*, Vol. 7, No. 1 (2015), pp. 243-74.

[26] Lewis, Matthew and Kevin Pflum, (2017) "Hospital Systems and Bargaining Power: Evidence from Out of Market Acquisitions," *RAND Journal of Economics*, Vol. 48, No. 3 (Fall 2017), pp. 579-610.

[27] Malsberger BM, Ed. "Covenants Not To Compete: A State-by-State Survey" (Arlington, VA : BNA Books), 1st Edition 1991, 2nd Edition 1996, Cum. Supp. 1997, Cum. Supp. 2000, Cum. Supp. 2001, 3rd Edition 2003, 4th Edition 2004, 5th Edition 2006, 6th Edition 2008, Cum. Supp. 2009, 7th Edition 2011.

[28] Marx, M., D. Strumsky, and L. Fleming (2009). "Mobility, Skills, and the Michigan Noncompete Experiment," *Management Science* 55(6): 875-889.

[29] Pauly, Mark (1993), "US Health Care Costs—The Untold Story," *Health Affairs*, Vol. 12, No. 3, (1993), pp 152-159.

[30] Schneider, JE, P Li, D Klepser, NA Peterson, T Brown, and R Scheffler (2008), "The Effect of Physician and Health Plan Market Concentration on Prices in Commercial Health Insurance Markets," *International Journal of Health Care Finance and Economics*, Vol. 8, No. 1, (2008), pp. 13-26.

[31] Stock, James and Motohiro Yogo (1997), "Instrumental Variables Regression with Weak Instruments," *Econometrica*, Vol. 65 (1997), pp. 557-86.

Appendix: For Online Publication

Table A1: NCA Law Change Frequencies by Census Division

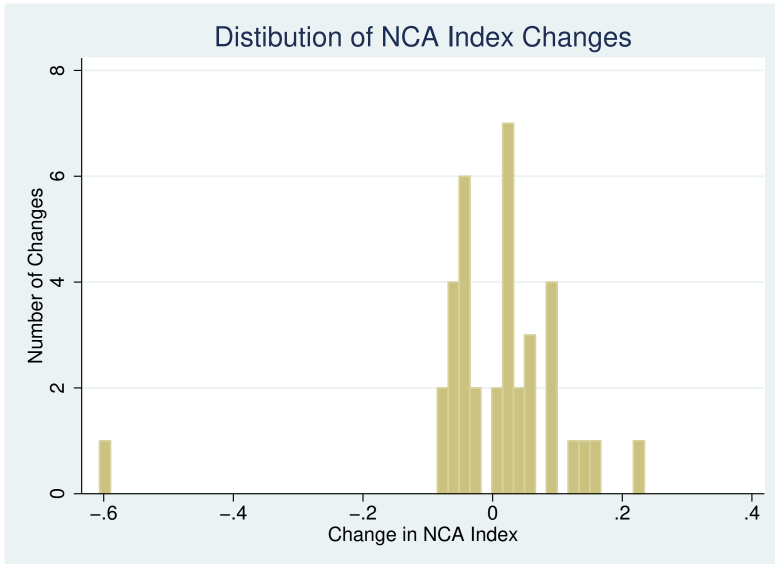
	East North Central	East South Central	Middle Atlantic	Mountain	New England	Pacific	South Atlantic	West North Central	West South Central	Total
	Positive Changes									
Statutory Index	0	1	0	0	0	0	0	1	0	2
Protectible Interest Index	0	1	1	2	1	1	1	2	2	11
Burden of Proof Index	1	1	1	2	1	0	1	0	0	7
Consideration Index Inception	0	0	0	0	0	0	0	1	0	1
Consideration Index Post-Inception	1	0	0	0	1	0	0	0	2	4
Blue Pencil Index	0	0	0	0	0	0	0	1	0	1
Employer Termination Index	0	0	2	0	0	0	0	0	0	2
	Negative Changes									
Statutory Index	1	1	0	2	0	0	1	0	2	7
Protectible Interest Index	1	1	0	0	0	1	0	0	0	3
Burden of Proof Index	0	1	1	0	0	0	1	1	0	4
Consideration Index Inception	0	1	1	0	0	0	0	0	0	2
Consideration Index Post-Inception	0	1	0	0	0	1	2	0	0	4
Blue Pencil Index	0	1	0	0	1	0	1	1	0	4
Employer Termination Index	0	0	0	0	0	0	0	0	0	0
Total All Dimensions	4	9	6	6	4	3	7	7	6	52

Table A2: Bishara (2011) Rating of the Restrictiveness of Non-Compete Agreements

Question #	Question	Criteria	Question Weight
Q1	Is there a state statute that governs the enforceability of covenants not to compete?	10 = Yes, favors strong enforcement 5 = Yes or no, in either case neutral on enforcement 0 = Yes, statute that disfavors enforcement	10
Q2	What is an employer's protectable interest and how is that defined?	10 = Broadly defined protectable interest 5 = Balanced approach to protectable interest 0 = Strictly defined, limiting the protectable interest of the employer	10
Q3	What must the plaintiff be able to show to prove the existence of an enforceable covenant not to compete?	10 = Weak burden of proof on plaintiff (employer) 5 = Balanced burden of proof on plaintiff 0 = Strong burden of proof on plaintiff	5
Q3a	Does the signing of a covenant not to compete at the inception of the employment relationship provide sufficient consideration to support the covenant?	10 = Yes, start of employment always sufficient to support any CNC 5 = Sometimes sufficient to support CNC 0 = Never sufficient as consideration to support CNC	5
Q3b	Will a change in the terms and conditions of employment provide sufficient consideration to support a covenant not to compete entered into after the employment relationship has begun?	10 = Continued employment always sufficient to support any CNC 5 = Only change in terms sufficient to support CNC 0 = Neither continued employment nor change in terms sufficient to support CNC	5
Q3c	Will continued employment provide sufficient consideration to support a covenant not to compete entered into after the employment relationship has begun?	10 = Continued employment always sufficient to support any CNC 5 = Only change in terms sufficient to support CNC 0 = Neither continued employment nor change in terms sufficient to support CNC	5
Q4	If the restrictions in the covenant not to compete are unenforceable because they are overbroad, are the courts permitted to modify the covenant to make the restrictions more narrow and to make the covenant enforceable? If so, under what circumstances will the courts allow reduction and what form of reduction will the courts permit?	10 = Judicial modification allowed, broad circumstances and restrictions to maximum enforcement allowed 5 = Blue pencil allowed, balanced circumstances and restrictions to middle ground of allowed enforcement 0 = Blue pencil or modification not allowed	10
Q8	If the employer terminates the employment relationship, is the covenant enforceable?	10 = Enforceable if employer terminates 5 = Enforceable in some circumstances 0 = Not enforceable if employer terminates	10

Source: Bishara (2011). Notes: The questions in the table correspond to the NCA law components used in the IV estimates throughout the paper. In the paper and tables, we refer to Q1 as the 'Statutory Index', to Q2 as the 'Protectible Interest Index', to Q3 as the 'Burden of Proof Index', to Q3a as 'Consideration Index Inception', to Q3b and Q3c together as 'Consideration Index Post-Inception', to Q4 as 'Blue Pencil Index', and to Q8 as 'Employer Termination Index'. In the raw data, the laws are scaled in each state-year from 0 to 10, as indicated by this table. In the estimations, each component is rescaled to range from 0 to 1, where 0 is the least restrictive observation in the data and 1 is the most.

Figure A1: Distribution of NCA Index Changes



Notes: Data points underlying the histogram are state-year observations of year-to-year changes in the NCA Index, which is a weighted sum of the 7 NCA law dimensions. The Index is scaled to range from 0 to 1, where 0 is the least restrictive state-year in the sample and 1 is the most restrictive. Changes in the Index can thus range from -1 to 1.

Table A3: NCA Law Components: Descriptive Statistics

	Mean	SD	N (State-Years)
Statutory Index	0.55	0.24	612
Protectible Interest Index	0.60	0.24	605
Burden of Proof Index	0.57	0.27	602
Consideration Index Inception	0.84	0.30	563
Consideration Index Post-Inception	0.70	0.33	526
Blue Pencil Index	0.53	0.34	538
Employer Termination Index	0.62	0.30	408

Notes: Statistics in the table represent data from 1996-2007 for each state-year in which a legal precedent exists. The minimum of each component is 0 and the maximum of each component is normalized to 1.

Table A4: Fixed Effects Models of Establishment Births and Deaths

Dependent Variable:	Births		Deaths	
	By Component	Combined	By Component	Combined
	(1)	(2)	(3)	(4)
Statutory Index$_{t-1}$	–1.281*	–0.612*	–1.348*	–0.734*
	(0.091)	(0.092)	(0.121)	(0.124)
Protectible Interest Index$_{t-1}$	0.583*	1.259*	0.609*	1.203*
	(0.066)	(0.158)	(0.089)	(0.177)
Burden of Proof Index$_{t-1}$	–0.633*	–3.684*	–0.532*	–3.659*
	(0.117)	(0.270)	(0.136)	(0.329)
Consideration Index Inception$_{t-1}$	0.024	3.389*	–0.354*	2.039*
	(0.088)	(0.299)	(0.090)	(0.265)
Consideration Index Post-Inception$_{t-1}$	–0.293*	–0.847*	0.081*	–0.458*
	(0.050)	(0.093)	(0.038)	(0.074)
Blue Pencil Index$_{t-1}$	0.235*	0.288*	–0.197*	–0.307*
	(0.041)	(0.060)	(0.048)	(0.065)
Employer Termination Index$_{t-1}$	–4.015*	–4.673*	–4.428*	–4.533*
	(0.513)	(0.630)	(0.682)	(0.780)
N		599,975		599,975
R-Sq		0.44		0.34

Notes: Columns 1 and 3 report estimates from separate regressions on each law component, and columns 2 and 4 report estimates from regressions including all 7 components. Dependent variables are the number of establishment births (columns 1 and 2) and deaths (columns 3 and 4) from the MPIER data. All specifications control for the aggregate supply of physicians and include fixed effects for county by medical specialty, and census division by year. Huber-White standard errors reported in parentheses. * indicates significance at the 0.05 level.

Table A5: IV First Stage: Effect of NCA Laws on Establishment-Based Market Concentration

	Dependent Variable: HHI_{t-1}		
	(1)	(2)	(3)
Statutory $Index_{t-1}$	0.42		1.80
	(1.38)		(1.99)
Protectible Interest $Index_{t-1}$	12.16*		8.05*
	(3.51)		(3.95)
Consideration Index $Inception_{t-1}$	21.71		15.22
	(38.57)		(41.59)
Consideration Index $Post\text{-}Inception_{t-1}$	–2.46*		–1.27
	(0.34)		(0.64)
Burden of Proof $Index_{t-1}$	–20.70		–13.90
	(30.66)		(33.39)
Blue Pencil $Index_{t-1}$	12.56*		0.99
	(3.90)		(2.29)
Employer Termination $Index_{t-1}$	–19.15*		–11.17*
	(4.51)		(4.41)
Statutory $Index_{t-2}$		–2.15	–1.85
		(1.14)	(1.77)
Protectible Interest $Index_{t-2}$		6.55*	2.92*
		(1.48)	(1.13)
Consideration Index $Inception_{t-2}$		–3.98*	0.36
		(1.39)	(1.15)
Consideration Index $Post\text{-}Inception_{t-2}$		–2.54*	–1.55*
		(0.36)	(0.61)
Burden of Proof $Index_{t-2}$		–1.99	–3.70*
		(1.34)	(0.82)
Blue Pencil $Index_{t-2}$		16.77*	16.18*
		(2.75)	(2.29)
Employer Termination $Index_{t-2}$		–6.71*	–1.62
		(1.77)	(1.25)
N	3,026,780	3,026,780	3,026,780
N Clusters	121	121	121
R-Sq	0.75	0.75	0.75
AP F-Stat	86.85	110.45	460.22

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. All independent variables are scaled to range between 0 and 1, where 1 is the strongest observed measure of the variable in any state and year in the data. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

Table A6: Sensitivity to IV Estimator

	Dependent Variable: $\ln(Price)_t$		
Instruments:	First Lags	Second Lags	Both Lags
	2SLS (Baseline)		
HHI_{t-1}	–0.028*	–0.024*	–0.025*
	(0.006)	(0.005)	(0.005)
	LIML		
HHI_{t-1}	–0.035*	–0.030*	–0.036*
	(0.007)	(0.006)	(0.007)
	2-Step GMM		
HHI_{t-1}	–0.023*	–0.016*	–0.020*
	(0.004)	(0.004)	(0.003)
N	3,026,780	3,026,780	3,026,780
N Clusters	121	121	121
1st Stage AP F-Stat	86.85	110.45	460.22

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

Table A7: IV Second Stage Estimates: MPIER HHIs, Markets defined by county only

	Dependent Variable: ln($Price$)			
	IV	IV	IV	OLS
Instruments:	First Lags	Second Lags	Both Lags	
$\text{HHI}_{(t-1)}$	–0.011*	–0.015*	–0.011*	0.000
	(0.050)	(0.005)	(0.005)	(0.000)
N	3,243,820	3,243,820	3,243,820	3,243,820
N Clusters	121	121	121	121
1st Stage AP F-Stat	29.03	53.70	168.33	

Notes:All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. Markets are defined by county only, and are not differentiated by physician specialty. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

Table A8: IV Estimates Dropping Largest NCA Law Changes

	Dependent Variable: $\ln(Price)_t$		
	(1)	(2)	(3)
HHI_{t-1}	–0.0263*	–0.0212*	–0.0218*
	(0.0052)	(0.0041)	(0.0039)
Instruments	First Lags	Second Lags	Both Lags
N	2,853,469	2,853,469	2,853,469
N Clusters	111	111	111
1st Stage AP F-Stat	92.47	119.33	470.95

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. The sample excludes state-years with the largest law changes, which account for 6.3% of the main sample. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

Table A9: IV Second Stage Estimates for Alternative MPIER HHI Measures

	Dependent Variable: ln($Price$)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$HHI_{(t-1)}$	–0.025*	–0.016*	–0.014*	–0.014*	–0.022*	–0.022*	–0.020*	–0.020*
	(0.005)	(0.003)	(0.004)	(0.003)	(0.007)	(0.007)	(0.008)	(0.008)
N	3,026,780	3,026,780	2,936,694	2,936,694	3,026,780	3,026,780	2,936,694	2,936,694
N Clusters	121	121	121	121	121	121	121	121
R-Sq	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
1st Stage AP F-Stat	460.22	704.58	1735.79	868.56	710.56	816.02	498.02	598.01

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. HHI is calculated from establishment sizes in MPIER data, provided by CMS. In column (1) the HHI is measured including all physicians in any group that has at least one member in a given specialty, and assumes physicians with missing addresses are solo establishments. The HHI in column (1) is the one used throughout the paper. The HHI in column (2) is similar to that in column (1), but assumes all physicians in a given market with missing addresses are in the same establishment. In column (3) the HHI is measured including all physicians in any group that has at least one member in a given specialty, drops observations with missing addresses if the same physician has another known address in the same zip code, and assumes all remaining missing addresses are solo establishments. The HHI in column (4) is similar to that in column (3), but assumes all remaining missing addresses in a given market are a single establishment. In column (5) the HHI is measured including only physicians in the given specialty within the market, and assumes physicians with missing addresses are solo establishments. The HHI in column (6) is similar to that in column (5), but assumes all physicians in a given market with missing addresses are in the same establishment. In column (7) the HHI is measured including only physicians in the given specialty within the market, drops observations with missing addresses if the same physician has another known address in the same zip code, and assumes all remaining missing addresses are solo establishments. The HHI in column (8) is similar to that in column (7), but assumes all remaining missing addresses in a given market are a single establishment. All HHIs are scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. In each specification, the instruments include all lagged and twice lagged law components (corresponding to column 3 in Table 7). All standard errors clustered by state-year. * indicates significance at the 0.05 level.

Table A10: IV Estimates Using Only Instruments with Negative First Stage

	Dependent Variable: $\ln(Price)_t$		
	(1)	(2)	(3)
HHI_{t-1}	–0.012*	–0.021*	–0.028*
	(0.005)	(0.005)	(0.005)
Instruments	First Lags	Second Lags	Both Lags
N	3,263,781	3,034,073	3,026,780
N Clusters	136	123	121
1st Stage AP F-Stat	109.18	126.30	515.27

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. Instruments do not include Blue Pencil Index, which is the only index with a positive coefficient in the univariate just-identified first-stage model, as shown in Table 8. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

Table A11: Positive and Negative Law Changes

	Dependent Variable: ln($Price$)		
	(1)	(2)	(3)
	Positive Law Changes		
HHI_{t-1}	–0.026*	–0.025*	–0.028*
	(0.008)	(0.006)	(0.005)
N	1,355,532	1,269,528	1,269,528
N Clusters	66	60	60
1st Stage AP F-Stat	88.07	108.36	700.24

	Dependent Variable: ln($Price$)		
	(1)	(2)	(3)
	Negative Law Changes		
HHI_{t-1}	–0.010	–0.027*	–0.026*
	(0.007)	(0.007)	(0.007)
N	2,798,107	2,623,004	2,623,004
N Clusters	99	90	90
1st Stage AP F-Stat	27.74	54.66	37.55

Notes: All specifications are the same as in Table 6. Top panel includes states with positive changes in NCA enforceability and control states, bottom panel includes states with negative changes in NCA enforceability and control states. All standard errors clustered by state-year. * indicates significance at the .05 level.

Table A12: Interactions between Physician and Insurer Concentration

	Dependent Variable: $\ln(Price)_t$		
	(1)	(2)	(3)
$Phys\ HHI_{t-1}$	–0.024*	–0.025*	–0.025*
	(0.004)	(0.005)	(0.006)
$Phys\ HHI_{t-1} \times I(Ins\ HHI > 2500)$	–0.001		
	(0.002)		
$Phys\ HHI_{t-1} \times I(Ins\ HHI > 4000)$		0.002	
		(0.002)	
$Phys\ HHI_{t-1} \times I(Ins\ HHI < 2500)$			0.002
			(0.002)
$Phys\ HHI_{t-1} \times I(Ins\ HHI > 5000)$			0.000
			(0.002)
N	3,026,780	3,026,780	3,026,780
N Clusters	121	121	121
1st Stage F-Stat	297.13	338.98	245.75

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. Physician HHIs are calculated from establishment sizes in MPIER data, insurer HHIs are state-level measures in 2007 from the AMA. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors are clustered by state-year. Kleinbergen-Paap F-Statistics reported. * indicates significance at the 0.05 level.

Table A13: Effect of Concentration on Prices, by Medical Specialty and Urban Status

Instruments	Dependent Variable: $\ln(Price)_t$					
	All Counties		Metro Counties		Non-Metro Counties	
	First Lags	First and Second Lags	First Lags	First and Second Lags	First Lags	First and Second Lags
	(1)	(2)	(3)	(4)	(5)	(6)
	All Physicians					
HHI_{t-1}	–0.028*	–0.025*	–0.034*	–0.031*	–0.013*	–0.011*
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.005)
N	3,026,780	3,026,780	2,077,627	2,077,627	949,153	949,153
1st Stage AP F-Stat	86.9	460.2	54.7	364.3	15.4	15.1
	Primary Care Physicians					
HHI_{t-1}	–0.024*	–0.017*	–0.036*	–0.026*	–0.007	–0.006
	(0.005)	(0.005)	(0.008)	(0.006)	(0.004)	(0.003)
N	473,033	473,033	306,449	306,449	166,584	166,584
1st Stage AP F-Stat	47.2	542.3	22.8	458.0	14.9	12.2
	Non-Surgical Specialists					
HHI_{t-1}	–0.013*	–0.013*	–0.026*	–0.023*	0.001	0.000
	(0.003)	(0.003)	(0.004)	(0.004)	(0.002)	(0.002)
N	300,990	300,990	234,402	234,402	66,588	66,588
1st Stage AP F-Stat	39.2	98.3	84.4	277.8	31.2	21.8
	Surgical Specialists					
HHI_{t-1}	–0.004	–0.005	–0.009	–0.009	–0.000	–0.003
	(0.007)	(0.005)	(0.006)	(0.005)	(0.008)	(0.003)
N	272,913	272,913	191,790	191,790	81,123	81,123
1st Stage AP F-Stat	8.3	17.2	29.21	55.0	1.1	2.4

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All estimates represent the second stage coefficient on HHI in 2SLS models corresponding to those in columns (1) and (3) of Table 7 for all counties, metro counties, and non-metro counties. The first two columns of the first panel reproduce the second stage results for all physicians in Table 7. The 'Primary Care Physicians' sample includes primary care MDs (excluding DOs), Internal Medicine, Family Practice, Geriatric Medicine, and Pediatric specialists. The 'Non-Surgical Specialist' sample includes specialists in Proctology, Urology, Dermatology, Cardiovascular Dis/Cardiology, Neurology, Gastroenterology, and Hematology. The 'Surgical Specialist' sample includes specialists in General Surgery, Neurological Surgery, Orthopaedic Surgery, Thoracic Surgery, Anesthesiology, and Radiology. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

Table A14: IV Estimates on Balanced Panel

	Dependent Variable: ln($Price$)		
	(1)	(2)	(3)
$\text{HHI}_{(t-1)}$	–0.036*	–0.034*	–0.032*
	(0.006)	(0.005)	(0.005)
N	2,032,976	2,032,976	2,032,976
N Clusters	121	121	121
1st Stage AP F-Stat	40.33	83.65	730.41

Notes: All specifications are the same as in Table 7, except the sample includes only observations corresponding to a county-specialty pair that is observed in all 12 years of the panel. All standard errors clustered by state-year. * indicates significance at the .05 level.

Table A15: Sensitivity of MPIER Second Stage IV Estimates to Fuzzy Matching Algorithm Parameter

Normalized Levenshtein Distance Threshold	IV Estimate	First Stage AP F-Stat.
0.01	–0.026* (0.005)	516.27
0.05	–0.025* (0.005)	535.96
0.10	–0.025* (0.005)	489.11
0.15	–0.026* (0.005)	478.49
0.20	–0.025* (0.005)	460.22
0.25	–0.027* (0.005)	530.28
0.30	–0.026* (0.005)	613.38
0.35	–0.026* (0.005)	607.27
0.40	–0.026* (0.005)	587.15

Notes: All specifications include fixed effects for county, census division by year, procedure code (CPT), physician specialty, and facility type. IVs are the full set of first and second lags of law components. The normalized Levenshtein Distance equals the minimum number of character insertions, deletions, or substitutions necessary to make two strings equal, divided by the length of the shorter string. The threshold value is the value of the normalized Levenshtein distance below which the character elements of two addresses in the MPIER are assumed to be equivalent. A larger threshold value results in over-estimating the size of establishments, while too low a value in the presence of typographical errors may lead to an underestimate of establishment sizes. The main estimates in the paper are based on a threshold value of 0.20. HHI is calculated from establishment sizes in MPIER data, provided by CMS. HHI is scaled to range from 0 to 100, so that a 1 unit change in HHI corresponds to a 100 point change in the typical 10,000 point scale. All standard errors are clustered by state-year. * indicates significance at the 0.05 level.

Table A16: Fixed Effects Models of Aggregate Physician Supply

Dependent Variable:	Log Physicians per 100,000 Population	
NCA Index$_t$	-0.027	
	(0.041)	
NCA Index$_{t-1}$	–0.022	–0.043
	(0.045)	(0.030)
Log Per Capita Income	0.156*	0.156*
	(0.030)	(0.030)
N	48,807	48,807
Adj. R Sq.	0.88	0.88

Notes: All specifications are fixed effects models and include county effects and census division by year effects. * indicates significance at the 0.05 level.

Table A17: Correlation of Law Changes with State Political and Economic Outcomes

Dependent Variable:	Log Payroll per Worker (1)	Unemployment Rate (2)	Population (3)	Republican Vote Share (4)
Statutory Index$_{t-1}$	–0.010	1.148*	–2183.276*	0.050
	(0.022)	(0.559)	(1073.492)	(0.033)
Protectible Interest Index$_{t-1}$	0.060	–0.636	724.974	–0.037
	(0.078)	(0.785)	(584.068)	(0.044)
Burden of Proof Index$_{t-1}$	0.051	0.762	–139.085	–0.034
	(0.040)	(0.859)	(520.628)	(0.061)
Consideration Index Inception$_{t-1}$	–0.056	0.328	678.706	–0.012
	(0.057)	(1.151)	(970.078)	(0.098)
Consideration Index Post-Inception$_{t-1}$	–0.038	–0.345	–367.454	0.035
	(0.023)	(0.599)	(252.488)	(0.035)
Blue Pencil Index$_{t-1}$	0.009	–0.702	–1485.250*	0.024
	(0.034)	(0.528)	(735.155)	(0.039)
Employer Termination Index$_{t-1}$	–0.119	–0.612	–567.853	–0.057
	(0.061)	(0.778)	(481.806)	(0.065)
N	969	969	969	510
N Clusters	51	51	51	51

Notes: An observation in these regressions is a state-year, and regressions are estimated by OLS with state and year fixed effects. All independent variables are scaled to range from 0 to 1, where 1 is the strongest observed measure of the variable in any state and year in the data. Standard errors are clustered by state. Data are from the Bureau of Labor Statistics (cols. 1 and 2), the Census Bureau (col. 3), and the Federal Election Commission (col. 4: presidential and congressional elections – every two years). Population is measured in thousands. Unemployment rate is measured in percentage points. * indicates significance at the 0.05 level.

Table A18: Correlation of Law Changes with Political and Economic Views in the GSS

Dependent Variable:	Respondent Thinks The Government Should Do Less:		Respondent Thinks We are Spending too Much On:			Respondent Considers Himself:	
	In General	To Help Pay for Medical Care	Urban Issues	Welfare	Nation's Health	A Republican	Satisfied With His Finacial Situation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Statutory Index$_{t-1}$	0.316	0.031	–0.166	–0.102	–0.121	–0.009	–0.297
	(0.177)	(0.120)	(0.257)	(0.322)	(0.155)	(0.169)	(0.216)
Protectible Interest Index$_{t-1}$	–0.026	0.074	–0.427	–0.513	0.021	–0.331	–0.074
	(0.376)	(0.196)	(0.372)	(0.462)	(0.210)	(0.365)	(0.363)
Burden of Proof Index$_{t-1}$	–0.103	–0.031	–0.685	0.394	0.215	–0.141	–0.454
	(0.383)	(0.360)	(0.515)	(0.745)	(0.343)	(0.502)	(0.317)
Consideration Index Inception$_{t-1}$	0.029	–0.092	0.819	–0.164	–0.453	0.463	0.527
	(0.438)	(0.340)	(0.558)	(0.758)	(0.347)	(0.502)	(0.422)
Consideration Index Post-Inception$_{t-1}$	0.144	–0.131	–0.034	0.151	0.546*	–0.062	0.001
	(0.123)	(0.086)	(0.407)	(0.244)	(0.208)	(0.271)	(0.237)
Blue Pencil Index$_{t-1}$	–0.297	0.365	–0.026	–0.121	0.268	–0.297	0.228
	(0.339)	(0.240)	(0.468)	(0.474)	(0.317)	(0.405)	(0.535)
Employer Termination Index$_{t-1}$	0.817	0.738	–0.974	–0.197	0.237	–0.325	0.631
	(0.532)	(0.568)	(0.484)	(0.791)	(0.590)	(0.538)	(1.006)
N	1,026	1,026	1,026	1,026	1,026	1,026	1,026
N Clusters	28	28	28	28	28	28	28

Notes: Regressions are linear probability models in which an observation is a survey respondent in a given year and a positive outcome represents the respondent's agreement with the statement presented in each column. All regressions include state, year, occupation, and industry fixed effects, as well as controls for age, education, marital status, and employment status. All independent variables are scaled to range from 0 to 1, where 1 is the strongest observed measure of the variable in any state and year in the data. Standard errors are clustered by state. Data on political and economic views are taken from the Generalized Social Survey for the years 1993-2010, where data exist (approximately every other year and in only 28 states). * indicates significance at the 0.05 level.

8 Bargaining Model

We model bargaining between physician groups and insurers following the setup of Ho and Lee (2016). The purpose of the model is to derive a relationship between negotiated prices and firm sizes under a set of plausible assumptions, and clarify how our empirical estimates can provide bounds on the underlying theoretical parameters. The market consists of a set of physician groups j and insurers i. Enrollees in insurance plan i can only visit a physician that is in the insurer's network, where the network is denoted by $\mathcal{G}_i \subseteq \{0,1\}^{i \times j}$. Similarly, $\mathcal{G}_j$ is the set of insurers with whom physician group j has contracted to be included in the network.

In each period of the model the following events take place. First, insurers and physician groups conduct simultaneous bilateral bargains over capitated prices p_{ij}, which are private knowledge of the negotiating parties.[13] Simultaneously with bargaining, insurers set profit-maximizing uniform premiums ϕ_i. Next, consumers form willingnesses to pay for insurance plans based on premiums and physician access in the network, measured by the amount of time a patient has to wait to get an appointment, $w_i(\phi_i, \mathcal{G})$, which depends on plan enrollment (and therefore plan premiums) and the size of the provider network. Finally, consumers probabilistically get sick and derive utility from being treated by a physician, and disutility from waiting for an appointment.

There are several simplifying assumptions about consumer choices. First, consumers are assumed to be incapable of discerning physician quality; they view physicians as homogeneous and value networks insofar as they differ in access. This assumption is made due to data limitations. In the hospital setting it is possible to obtain data on input choices for each hospital in a given market, which can allow researchers to estimate cost functions directly and model latent quality differences through fixed hospital effects (see Ho and Lee, 2016.) In physician markets there are no known similar data on the input choices of every physician office in a market, so the same estimation approach cannot be used. Second, we assume that insurers set uniform copayments. As a result, consumers are not directly affected by negotiated prices between physicians and insurers, although prices may have indirect effects on consumers through premiums or wait times. We abstract from specialties, but in the empirical estimates we consider each physician specialty to be a distinct market. The remaining model assumptions are similar to those made in models of hospital bargaining, such as Ho and Lee (2016) and Gowrisankaran et al. (2013).

The profit function of insurer i is:

$$\pi_i(\mathbf{p}, \mathcal{G}) = D_i(w_i, \phi)\phi_i - \sum_{r \in \mathcal{G}_i} D_{ir}(w_i, \phi) p_{ir}$$

where D_i represents the number of enrollees in insurance plan i, which depends on wait times $w_i(\phi_i, \mathcal{G})$ in network i, and D_{ij} is the number of enrollees in plan i who visit physician

[13] In reality many contracts are capitated, but for other contracts a capitated payment is conceptually similar to an average price for an expected bundle of services.

group j.[14] The profits of physician group j are similarly:

$$\pi_j(\mathbf{p}, \mathcal{G}) = \sum_{s \in \mathcal{G}_j} D_{sj}(w_i, \phi)(p_{sj} - c_j)$$

which equals the sum of enrollees D_{sj} over all insurers in the network of physician group j times the negotiated price p_{sj} minus c_j, the average per-patient cost for physician group j.

Prices are negotiated through simultaneous bilateral Nash bargains, where p_{ij} solves:

$$p_{ij} = \arg\max_{p_{ij}} \left[\pi_i(\mathbf{p}, \mathcal{G}) - \pi_i(\mathbf{p}_{-ij}, \mathcal{G}\backslash ij)\right]^{\tau_i} \times \left[\pi_j(\mathbf{p}, \mathcal{G}) - \pi_j(\mathbf{p}_{-ij}, \mathcal{G}\backslash ij)\right]^{\tau_j} \quad \forall\ ij \in \mathcal{G}$$

where $\pi_i(\mathbf{p}_{-ij}, \mathcal{G}\backslash ij)$ represents the disagreement profits of insurer i if they fail to reach an agreement over network inclusion with physician group j, and similarly $\pi_j(\mathbf{p}_{-ij}, \mathcal{G}\backslash ij)$ are the disagreement profits of physician group j. τ_i and τ_j are the bargaining power parameters of the insurer and physician group.

The first order condition of the bargaining problem simplifies to:

$$\begin{aligned}
\underbrace{p^\star_{ij} D_{ij}}_{\text{Physician Group Revenue}} &= \tau_j \left[\underbrace{\phi_i\left(D_i - D_{i-j}\right)}_{\Delta\text{Insurer Revenue}} - \underbrace{\left(\sum_{h \in \mathcal{G}_i \backslash ij} p^\star_{ih}\left(D_{ih} - D_{ih-j}\right)\right)}_{\Delta\text{Insurer } i \text{ Payments to Other Physicians}}\right] \\
&+ \tau_i \left[\underbrace{c_j D_{ij}}_{\text{Average Cost}} - \underbrace{\left(\sum_{n \in \mathcal{G}_j \backslash ij} \left(p^\star_{nj} - c_j\right)\left(D_{nj} - D_{nj-i}\right)\right)}_{\Delta\text{Physician Group } j \text{ Profits from Other Insurers}}\right] + \varepsilon_{ij}
\end{aligned} \quad (3)$$

where D_{i-j} is the number of enrollees in plan i if there is disagreement between i and j. The second term equals the additional payments that the insurer will have to make to other physician groups if group j is not included in the network, which is negative. $D_{ih} - D_{ih-j}$ is the effect of disagreement between insurer i and group j on the number of consumers in plan i who visit another group h, where $h \neq j$. The third term is the average cost to group j of treating an enrollee. The fourth term is the effect of disagreement between plan i and group j on the profits of group j from other insurers, which is negative. And ε_{ij} represents *iid* cost shocks.

Conditional on getting sick, consumer k derives utility from visiting a physician j in network i, which we assume takes the form:

$$u_{kij} = \eta_k + \frac{1}{w_{ij}}$$

[14] More precisely ϕ_i can be thought of as the premium for plan i net of any per-capita non-medical costs of running the plan.

where in equilibrium wait times will be equal within any network, so that $w_{ij} = w_i$. The average wait time for an enrollee who gets sick in network i is:

$$w_i = \beta \frac{\sum_{r \in \mathcal{G}_{i \times j}} \gamma N_i}{\sum_{r \in \mathcal{G}_{i \times j}} |P_j|}$$

where N_i is the number of enrollees in insurance plan i, γ is the probability of getting sick, $|P_j|$ is the size of physician group j, and $\mathcal{G}_{i \times j}$ denotes the connected subset of $\mathcal{G}$ that contains all insurers and physician groups that have any nodes in common with the networks $\mathcal{G}_i$ or $\mathcal{G}_j$. For an insurer i with an exclusive network of physicians that do not participate in other networks, this subset is simply $\mathcal{G}_i$.

As in Capps, Dranove, and Satterthwaite (2003) we consider willingness to pay (WTP) as a measure of the surplus that consumer k would lose if a given physician group were to leave the network. A consumer's change in utility caused by physician group j exiting the network is:

$$\Delta \text{WTP}_{kij} = u_{kij} \mid_{j \in \mathcal{G}_i} - u_{kij} \mid_{j \notin \mathcal{G}_i}$$

Each consumer's ex ante WTP is then $\gamma \Delta u_{kij}$. We express the WTP by the insurer for participation of group j in the network, which affects the premium charged by insurer i, as a constant proportion ξ of the average consumer surplus:

$$\Delta \text{WTP}_{ij} = \frac{\sum_k \Delta \text{WTP}_{kij}}{N_i} \xi = \frac{|P_j|}{\beta \gamma \sum_{r \in \mathcal{G}_{i \times j}} N_i} \xi$$

As a result $\frac{\partial \text{WTP}_{ij}}{\partial |P_j|} > 0$ since premiums reflect consumers' WTP. Also $\frac{\partial p^\star_{ih} \left(D_{ih} - D_{ih-j}\right)}{\partial |P_j|} < 0$, so the second term of Equation 3 gets increasingly negative as practice size increases, since the number of consumers who visit other physician groups increases when a larger group exits the network. The fourth term is also increasing with group size. If a plan fails to agree with a larger group, equalization of wait times implies the group will attract more consumers from other plans. Therefore the sum of the first, second, and fourth terms in Equation 3 cause prices to increase with group size. However, the cost function potentially opposes this effect. Without making assumptions, it is plausible that there are economies of scale, and that average costs (the third term) are declining in group size. In this case the sign of the aggregate effect of group size on negotiated prices is ambiguous.

To construct an empirical analogue of the FOC, suppose in disagreement the potential consumers of group j are distributed proportionally among the other physicians in the network. Then:

$$\begin{aligned} p^\star_{ij} \quad &= \quad a + |P_j| \, \tau_j \xi + \sum_{h \in \mathcal{G}_i \setminus ij} \tau_j p^\star_{ih} \frac{D_{ih}}{D_{ij}} \left(1 + \frac{|P_h|}{|\mathcal{G}_i| - |P_j|}\right) + \tau_i c_j(|P_j|) \\ &+ \quad \sum_{n \in \mathcal{G}_j \setminus ij} \tau_i \left(p^\star_{nj} - c_j\right) \frac{D_{nj}}{D_{ij}} \left(\frac{|P_j|}{|\mathcal{G}_i| - |P_j|} - \frac{|P_j|}{|\mathcal{G}_i|}\right) + \epsilon_{ij} \end{aligned} \tag{4}$$

This gives the equilibrium negotiated price, plugging the WTP values from the utility function into Equation 3. The negotiated price depends on the bargaining power parameters, physician group sizes, and the number of physicians in insurer i's network, $|\mathcal{G}_i|$, conditional on agreement with group j. Given the theoretical ambiguous effect of $|P_j|$ on $p^{\star}_{ij}$, it is an empirical exercise to determine this relationship.

8.1 Empirical Implementation

In our empirical setting we cannot estimate Equation 4 directly because we do not observe the bargaining parameters or practice-level demand. Instead we consider the combined impact of physician practice sizes on negotiated prices through two aggregated components: the value of including practice j in the network of insurer i, and the cost function of practice j:

$$p^{\star}_{ij} \equiv a + \beta_1 \times \text{Network Value}_j(|P_j|) + \tau_i \times \text{Average Cost}_j(|P_j|) + \epsilon_{ij} \tag{5}$$

where Network Value$_j(|P_j|)$ is defined by the sum of the first, second, third, and fifth terms in Equation 4, and β_1 captures the average effect of practice size on prices through network value. Average Cost$_j(|P_j|)$ is the fourth term, which has coefficient τ_i according to Equation 4.

There are several further adjustments to the model that must be made given our empirical setting and data. First, since we do not observe costs, what we can actually identify is an aggregate coefficient that combines β_1 and τ_i. Second, Equation 5 represents a specific market, where markets may be defined by a combination of geography, physician specialty, and time. In our analyses we use data from many markets, while controlling for latent market-specific variation. Finally, we do not observe the negotiated price for each practice; we only know the average price across all practices in a market.

The empirical analogue of the structural model we consider is thus:

$$\overline{p^{\star}_{mpct}} = \alpha + \beta_2 ES_{mct} + \beta_3 FS_{mct} + \eta_m + \pi_p + \gamma_c + \nu_{d(c)t} + \varepsilon_{mpct} \tag{6}$$

where ES_{mct} measures establishment sizes in specialty market m, county c, and year t; FS_{mct} measures firm sizes; and β_2 and β_3 represent effects of changes in each of the practice size measures on average negotiated prices. This specification allows the derivative of costs with respect to firm size to differentially affect prices depending on whether firm growth occurs within or across establishments. The equation includes controls for latent heterogeneity across services through medical specialty effects, η_m, and procedure code effects, π_p; across space through geographic effects, γ_c, for which we consider a variety of potential market definitions; and over time through census-division-by-year effects, $\nu_{d(c)t}$, which nest year effects while allowing prices to change arbitrarily over time across census divisions.

Given the limitations of the empirical model relative to the structural analogue, it is worth questioning whether the parameters are nevertheless useful for understanding the extent to which larger practice sizes may lead to higher prices by increasing the network value of the practice. In general they may not be very informative, since both β_2 and β_3 identify combinations of the effects of changes in average costs and network value, without separately identifying either parameter of interest. However, the estimates turn out to be informative in our setting because we find an important sign difference: $\beta_2 < 0$ while $\beta_3 > 0$. This combination of results implies lower bounds on both the network value parameter β_1 and the cost function parameter τ_i.

To understand why this result is informative, consider a hypothetical merger between two nearby physician practices that remain physically distinct after the merger but minimize costs jointly and negotiate with insurers jointly. The network value of the combined firm cannot decline, because otherwise the firm would prefer to negotiate separately by establishment, an option still within the choice set. Similarly, average costs cannot increase, since minimizing costs separately by establishment is still within the choice set. After the merger, there is no change in ES since the establishments remain distinct, but FS increases. If the merger were to increase negotiated prices, $\beta_3 > 0$, this would imply that the true effect of the merger on network value is at least as large as β_3, since τ_i is non-positive in this case.

Conversely, suppose the same two nearby firms merge and physically consolidate into a single establishment. In this case the change in FS is the same as in the case above, but ES now also increases. In our theoretical model, the network value of the post-merger firm depends on the total number of doctors (not on physical consolidation) and is thus the same as in the case above. A finding of $\beta_3 > 0$, then, suggests the effect of the merger on prices due to network value will also be positive in this case. However, a cost-reducing physical consolidation could put downward pressure on negotiated prices. If this merger were to generate a decrease in prices the implication would be that the average cost effect of τ_i dominates any change in network value, implying that β_2 is a lower bound estimate of τ_i.

In our empirical analyses we estimate an aggregated version of this model using establishment sizes from the MPIER data and firm sizes calculated by linking multi-establishment practices together using IRS tax IDs. Our finding that $\widehat{\beta_2} < 0$ and $\widehat{\beta_3} > 0$ suggests insurers extract the efficiency gains from larger establishments in the form of lower prices, but multi-establishment consolidation yields efficiency gains that are smaller than the effects on network value, causing negotiated prices to increase. This model aims to facilitate the interpretation of these empirical parameters as lower bound estimates of τ_i and β_1, the parameters of interest.