WORKING PAPERS

Price Effects of a Merger: Evidence from a Physicians' Market

Thomas Koch
Shawn W. Ulrick

WORKING PAPER NO. 333

August 2017

BUREAU OF ECONOMICS
FEDERAL TRADE COMMISSION
WASHINGTON, DC 20580

Price Effects of a Merger: Evidence from a Physicians' Market

By Thomas Koch and Shawn W. Ulrick[1]

U.S. Federal Trade Commission

Abstract

Physicians' practices vary widely, as do their effectiveness and reimbursement. Using a merger of six orthopaedic groups in southeastern Pennsylvania, we find that such groups can generate large, anti-competitive price increases without any demonstrated increases in quality (indirectly measured by way of revealed preference) or efficiency. Further, we find that these price increases were targeted at certain beneficiaries, payors and codes, so any research design that omits care and billing along any of these dimensions is likely to be biased.

I. Introduction

The relationship between physician concentration and prices has been of interest to policy makers; of particular concern is the impact that physician consolidation has on price. Plausibly, larger groups in a concentrated market have more leverage with payors and thus extract higher reimbursement from them because it would be costly for an insurance plan to exclude the provider from the network. This phenomenon has been seen in research on concentration in hospital markets.[2]

In this paper, we examine a particular instance of an increased concentration in a physician specialty: We measure the impact on reimbursement of a merger between orthopaedists in Berks County, PA. A primary advantage to this paper is that we exploit a plausibly exogenous merger (i.e., increase in concentration), thereby better

---

[1] 600 Pennsylvania Avenue, NW, Washington, DC 20580. E-mail: Koch, tkoch@ftc.gov; Ulrick, sulrick@ftc.gov. The views expressed in this article are those of the author and do not necessarily reflect those of the Federal Trade Commission. We have benefitted greatly by the comments and suggestions from Seth Sacher, Louis Silvia, Dan Hosken, Mike Vita and Nathan Wilson.

[2] See, e.g., Noether (1988); Dranove, Shanley, and White (1993); Lynk (1995); Keeler, Melnick, and Zwanziger (1999); Town and Vistnes (2001); and Capps, Dranove, and Satterthwaite (2003).

isolating the impact of concentration on price from other endogenous factors. We use detailed, claims-level data from three private payors in the region to evaluate this relationship (to preserve confidentiality, we name them simply "Payor A," "Payor B," and "Payor C"). We use difference-in-differences to compare the change in the reimbursements of the merged physicians' practices to others in nearby areas plausibly not affected by a change in concentration. Depending upon the payor and plan type, the consolidation increased physician payments by twelve to twenty percentage points more than the predicted but-for world in which the merger had not occurred. Other payors experienced no price effect.

Our research is the first to look at a merger of physicians groups with detailed claims data and actual market prices from a sample of payors in the market. Most previous research with healthcare provider data with this level of detail has been for hospitals. The recent research in physician markets (Dunn and Shapiro 2014) makes cross-sectional comparisons and finds that doctors in more concentrated markets are paid more for their services. Carlin et al. (2017) looks at the effects of an acquisition of providers by a health system, which mixes horizontal effects (the system already employed competing providers) and vertical effects (the acquired providers were now newly part of a health system, i.e., hospitals). Clemens and Gottlieb (2014; 2017) have considered the physicians' incentives induced by health care financing, and how the structure of Medicare's physician reimbursement can affect commercial payor reimbursement.

Our findings have lessons for both academic researchers and antitrust practitioners. We broadly confirm the conclusion of recent research into physician markets: Physicians acquiring a dominant position in a region as small as a county may extract higher reimbursement rates. However, we also find a great deal of heterogeneity in the effects, as has been seen before for hospital mergers (Garmon and Haas-Wilson, 2011; Tenn, 2011; Thompson, 2011). The increases in physician reimbursement associated with the orthopaedist merger are limited to certain payors and medical services. That is, the effects of the merger were isolated to two of the payors for whom we have data but not the third. Moreover, the reimbursement rates for some CPT codes remained unchanged or decreased, even if overall the payor experienced a price increase. Our estimated merger effect on prices is similar in magnitude to the effects found in Dunn and Shapiro (2014).

Given this heterogeneity across payors, who provides the data can have consequences for measuring the consequences of a merger. Efforts to collect claims from private insurers, such as the HCCI or MarketScan, may miss the increased prices that resulted from a merger if the contributing insurers or included CPT codes were more or less likely to be the target of consolidation. Alternatively, omitting payors who did not negotiate higher prices because of a merger may overstate the average effect of a merger.

We further note that this merger provides insight into the relationship between group size and quality. One source of physician variation is group size: How many physicians work together as a single legal entity, and how does this integration affect their practice, performance, and quality. The consequences of this variation are poorly understood: Larger groups may be paid more because they perform better or more efficiently (Welch, et al. 2013; Casalino, et al. 2014). The ability of researchers to understand the interactions of these three forces (practice size, effectiveness, and reimbursement) has been hindered by their co-determination and a lack of reliable data and research designs. In this merger, physician practices increased reimbursement rates without materially changing quality, as measured indirectly by revealed preference: patients were not more or less likely to leave the Reading area for orthopaedic care after the merger.

For anti-trust practitioners, we demonstrate the importance of detailed data and how various ways of constructing a representative basket of composite goods can influence estimates of merger effects in this setting. Further, with our control group of close-but-separate markets, we show that there is coincidental, though unrelated, variation in prices. Not all price increases are due to a loss of competition, and we measure price increases contemporaneous with but unrelated to the Keystone merger. This provides further evidence that but-for the merger, prices would not have increased as much as they did. This demonstrates that the consequences of the merger were unusually large relative to the price increases other physicians' groups received. Due to the sample sizes available to researchers and practitioners, statistical significance does not necessarily demonstrate economic significance.

Our paper is constructed as follows: We describe the relevant market, our data sources, and background information in Section 2; we present our analysis of the merger effects in Section 3; we attempt to evaluate potential merger efficiencies in Section 4; and we provide a conclusion in Section 5.

II. Background and Data

Keystone Orthopaedic Specialists was formed out of six existing orthopaedic physician groups that served the greater Reading, PA area and Berks County generally.[3] Prior to the merger, which was agreed to on March 19, 2010, the six constituent groups were independent of each other , and each had separate contracts with each of the insurers (or payors) in the market. As described in the Complaint, the six groups competed with each other in two stages. First, they competed with one another in their negotiations with payors, in setting the prices at which they provided care in an in-network (i.e., covered) setting. Second, once in network, they competed against each other in terms of quality, to attract patients. The merger was consummated on January 1, 2011, at which time the previous constituent groups of Keystone ceased to exist as independent entities.

All but one group practiced and had providers that were board certified in general orthopaedics. Orthopods treat and perform surgery on the musculoskeletal system. As described below, this means that their care focuses on surgery on and rehabilitation of, as well as injections to manage, musculoskeletal problems. Reading Neck and Spine practiced in orthopaedic subspecialty covering neck and spine care. As a result of the merger, 19 of the 25 orthopods (including subspecialists) in the county worked in a single group practice that jointly negotiated with insurers, or payors, who provided coverage to beneficiaries who lived in the area

The merger was investigated after its consummation by the Federal Trade Commission as being potentially anti-competitive. The Commission filed a complaint against the Keystone group, which was ultimately settled prior to litigation. In 2014, in the midst of the investigation, one of the groups (Orthopaedic Associates of Reading or OAR) left the Keystone to become an independent practice. The settlement of the complaint maintained the separation between Keystone and OAR, among other things.

Table 1 shows how the merger changed concentration in Berks County. The practice composition of each group was measured using detailed review of documents

[3] The events described here are as described in the FTC's complaint and Aid to Public Comment available https://www.ftc.gov/news-events/press-releases/2015/10/two-pennsylvania-orthopedic-practices-settle-ftc-charges-merger.

from the merging parties and other local providers acquired via a subpoena. The aggregates are described in the complaint, while the individual group sizes were available from the parties' on their websites prior to the merger. The table shows how many orthopaedists were in Keystone and the other groups, just prior to the merger in late 2010 and just after the Keystone merger in early 2011. If we count each physician as having equal share, the merger increased HHI from 1,435 to 5,459 in Berks County.

Table 1

	Pre-merger	Post-merger
Keystone	--	19
Advanced Orthopedics	1	--
Arthritis Joint & Replacement	1	--
Berkshire	3	--
Commonwealth	5	--
Reading N&S	2	--
OAR	7	
Rest	6	6

In the course of the FTC's investigation, several payors provided billing and claims data pursuant to subpoenas in order to evaluate the consequences of the merger. To maintain the confidentiality of the data, we refer to them as Payors A, B, and C. The claims data capture the medical claims for beneficiaries in southeastern Pennsylvania and are similar in form to the broadly popular Truven or HCCI claims data. The claims data have several key characteristics suited to our purposes. Most importantly, the data are at the transaction level; that is, they give the actual allowed amounts (hereafter, "prices") separately for each service (characterized by its Current Procedural Terminology (CPT) code, including modifiers), billed by a doctor for a beneficiary on a given visit.[4] The claims we study here reflect the professional services fees, or those that

[4] See Herman and Ulrick (2012) for further discussion of the different pricing fields in a typical claims dataset, including "allowed amount," information on how payment may be split between insurer and patient, and examples of modifier codes. Throughout this paper, *Price* means the "allowed amount" of a particular CPT code, which is the total compensation contractually due to a physician, including both the insurer's and patient's shares. We make no distinction as to how the allowed amount is allocated

are paid to the provider for their services. These claims cover services provided in both the inpatient and outpatient settings, though are predominantly the latter. Each line item reports the professional service provided by the doctor, listing both the service (e.g., evaluation and management services, or the administration of an injection) and the amount paid. (Providers might also be paid facility fees, for use of the office or other equipment, though such fees are not part of the market under consideration here.)

The data also give diagnostic information (ICD-9 codes) and beneficiary information, such as ZIP Code and age. The data also reflect *all* claims to a given payor, rather than just a sample for a fraction of a plan's beneficiaries. The claims also identify the doctor by UPIN/NPI (a pair provider-specific identification numbers) and the Tax Identification Number ("TIN"), which indicates who receives the payment for the claim. Importantly, we also know whether the plan was employer-sponsored or provided under the Medicare Advantage (Part C) program.

The data cover the years 2008-2013 for beneficiaries and providers in counties in southeastern Pennsylvania. (We have redacted the exact counties for each payor in order to maintain their confidentiality.) This coverage allows us to measure the consequences of the merger on prices and some aspects of quality of care, since it covers a span of time both prior and after the merger, for beneficiaries who lived in the geographic area affected by the merger (Reading, PA) and other parts not affected by the merger (other parts of southeastern Pennsylvania). Due to these data limitations, we are not able to analyze the effect of the settlement, which maintained the separation of OAR from Keystone.

III. Merger Effects

In assessing the effects of the merger, we apply a difference-in-differences approach to identify merger price effects. This is because prices change "naturally" every year, *i.e.*, they change for reasons unrelated to mergers or anticompetitive activity. Therefore, we cannot identify harm from the merger by simply detecting whether Keystone raised prices post-merger. Rather, we must see if Keystone increased prices more than it would have absent the merger. We are not likely to know how much Keystone's constituent practices would have increased prices absent the merger. But we can use the average price change of nearby, non-Keystone orthopaedists as a

between patient and insurer. Because we seek to measure the price of a physician's services covered by a plan, we exclude those occurring out of network.

benchmark.[5] The estimated merger price effect is the difference between Keystone and non-Keystone price changes, *i.e.*, the difference between the differences.

We apply three methods. First, we plot average Keystone and non-Keystone prices for particular CPT codes over time. This allows us to establish the timing of the contracts. Second, we construct representative baskets for orthopaedic services, weighted by the frequency of each of the services provided by orthopods, and calculate the price change for such baskets associated with the merger. Finally, we use regression analysis to calculate the difference between Keystone's and non-Keystone's average price changes (averaged across several CPT codes, forming the bulk of revenue received by Keystone from these payors).

As we will show, the merger effect can vary across plan types for a particular payor. Commercial payors may offer three kinds of plans: private plans, such as traditional employer-sponsored insurance ("ESI"); Medicare Advantage plans; and Medicaid managed care plans. Medicare Advantage plans are an alternative to traditional Medicare and are only available to those beneficiaries eligible for Medicare (*e.g*., sixty-five years old or older). Eligibility for Medicaid managed care plans are restricted to children, with either express eligibility rules or subsidies determined by the child's family structure and income.[6] The per-beneficiary reimbursement to insurers for these government-sponsored plans is administratively determined. Due to data limitations, we focus our analysis on ESI and Medicare Advantage plans.

These plans also cover different beneficiaries and serve different classes of enrollees. Medicare Advantage faces competition from traditional Medicare; an able-bodied, thirty-five-year-old economist in Reading, Pennsylvania is not eligible for either. That economist's daughter would be eligible for Medicaid managed care or CHIP subsidies if the family were sufficiently poor. Because of the differences between Medicare or CHIP and private plans, payors face dissimilar constraints when negotiating over them with providers. For example, the way payors are paid by Medicare Advantage makes it very difficult for a payor to pass cost increases through to Medicare, since their payment follows a risk-adjusted cost index for the beneficiaries county.

---

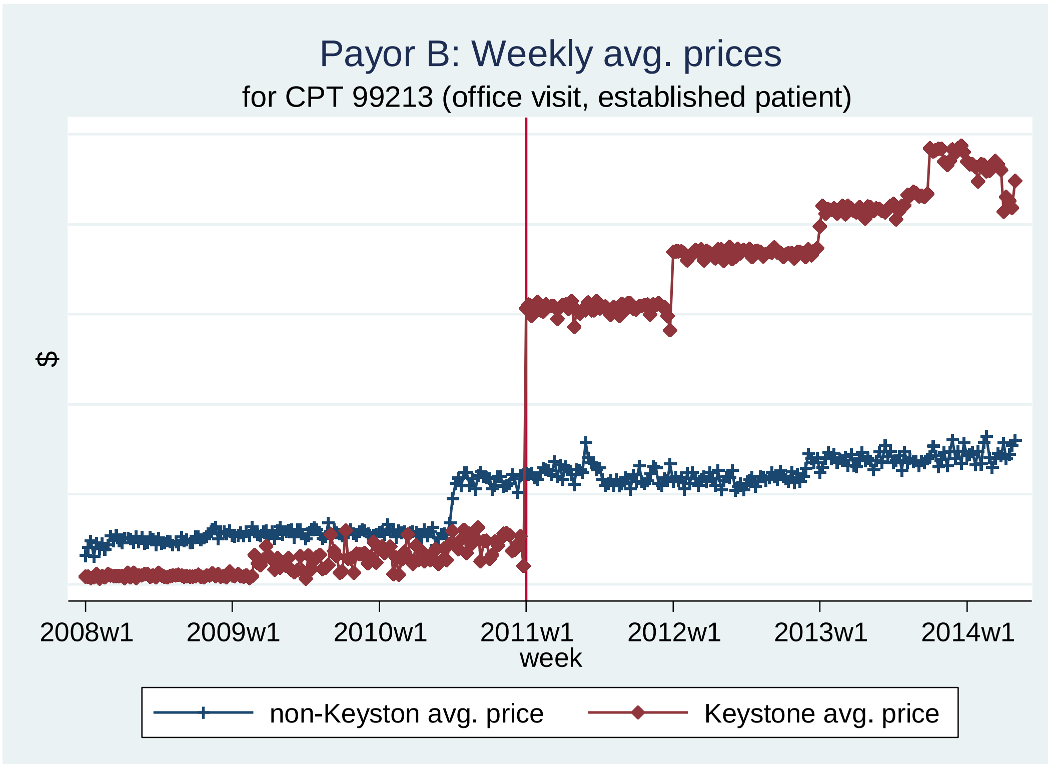
[5] The control group of physicians includes those in the same region as Keystone but in separate counties. These control counties vary by payor, as described in the Efficiencies section below.

[6] Eligibility is typically determined according to a cut-off threshold of the ratio of family income to the Federal poverty guideline. The Federal poverty guideline increases with family size and also varies by the child's age.

1. Keystone Prices of Individual CPT Codes

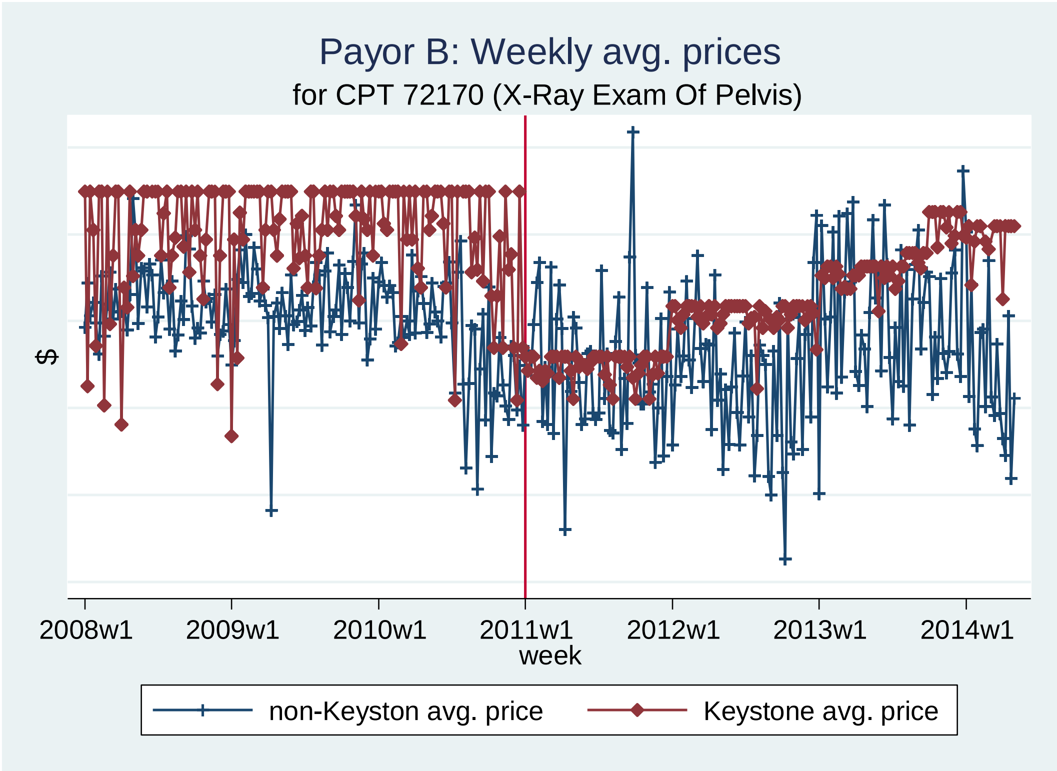
Payor B experienced a relative Keystone price increase for its ESI plans for some procedures, primarily in office visits, some injections, and radiological examinations. Figure 1 shows a typical pattern for a procedure with a relative price increase (specifically, CPT 99213, an established patient office visit). The figure plots weekly average prices of Keystone alongside those of the non-Keystone orthopaedists, for CPT code 99213, an established patient office visit. Keystone average prices include all component groups that eventually merged. The vertical line is the effective date of the merger. The pattern is striking: Keystone (on average) starts at about the same average price as non-Keystone orthopaedists but has substantially higher average price post-merger. There is some variation in weekly averages, which reflects weekly change in the composition of, patients' plans, for example, and variation in the composition of the control set of providers; but the figure clearly shows a price increase for this CPT code.

Figure 1



Keystone's relative price increases were limited to a subset of CPT codes. Moreover, some CPT codes decreased in price. Figure 2 provides CPT 72170 (an x-ray) as an example.

Figure 2



Time plots are illustrative but are not practical for more than just a few CPT codes. It is therefore important to employ methods that can convey information about a greater array of procedures. This must be done carefully because the mix of procedures can change over time, and some procedures are more expensive than others. If the composition changes differently across the various groups, it could bias our estimated merger effects. For example, an increase in average price for a particular group could simply reflect that group performing procedures that are more complicated.

2. Pre- and post-merger costs of a fixed basket of CPT codes

One way to measure the price change of a large number of products is to form a basket of commonly occurring items and calculate how the cost of purchasing the basket changes over time. We thus select a basket of frequently occurring CPT/modifier codes separately for each payor.[7] We then calculate what it would have cost to buy the pre-merger quantities in the baskets at the post-merger prices separately for Keystone and the controls. That is, we calculate the Laspeyres price index, given by:

$$I_L = \frac{\sum_{cpt} p_{post,cpt} \times q_{pre,cpt}}{\sum_{cpt} p_{pre,cpt} \times q_{pre,cpt}}$$

This approach has a strong intuitive appeal; it states how much the pre-merger basket of services would have cost post-merger. Because insurance limits a patient's exposure to the price of services, they are unlikely to substitute between services in the event that relative prices change, or the level of prices change. We find that the non-Keystone price change is near zero for both payors, whereas the price change for Keystone is substantial: 28.2% for Payor A and 14.5% for Payor B.

A limitation is that we must actively choose which services to include in the basket. Because we only include a limited number of services, the baskets account for 46% and 80% of the post-merger revenues (including beneficiary payments) earned by Keystone from Payors A and B, respectively. Several of the procedures have only a few (*e.g.*, three or fewer) observations in a bin. The prices and frequencies for such observations may be measured with considerable noise; expanding the baskets much further, to incorporate more of the revenue, would exacerbate this potential problem. For example, if an expensive but infrequently billed code were included in the bundle, a doubling of the rate of that code may reflect only an increase in the incidence of that code from twice a year to four times a year. Alternatively, there may be subsequent negotiation, bundling, rejection or other adjustments to these rare claims that could produce variability in the price. We have adjusted the bundles themselves, and marginal adjustments to the bundle composition do not materially alter our findings. Those results are available upon request.

3. Difference-in-Differences Regressions

The variation in prices changes across CPTs emphasizes the importance of controlling for the basket of codes. The basket approach of the previous subsection

[7] Each payor's basket includes all CPT/modifier code combinations with at least 100 observations, with at least one observation in each of the relevant bins. (The bins are pre-merger Keystone, post-merger Keystone, pre-merger non-Keystone, and post-merger non-Keystone.) The quantities for each provider-payor pair are set to the amounts for each payor. E.g., the Keystone basket for Payor A is Payor A's orthopedic basket across all providers, as it is for each of the providers when constructing their baskets for Payor A.

controls for basket but is limited in the breadth of the services that it can assess. Difference-in-differences regressions provide a convenient model that allows for including more codes in the estimation sample. The regressions permit us to include less frequently occurring modifier/CPT code combinations by imposing assumptions such as that a particular modifier code changes the price of all CPT codes by the same percentage.[8] We measure prices in natural logs because doing so makes imposing these restrictions easier. The log price model estimates give the *percentage change* in the average price of Keystone and the percentage change in the average price of the control groups. The difference between these percentage changes is the merger effect.

The model is

$$\ln(p_{gct}) = \beta_{merger}Keystone_g \times Post_t + \beta_{key}Keystone_g + \beta_{post}Post_t + CPT_c\gamma + \epsilon_{pcg}$$

where $\ln(p_{gct})$ is the natural log of the price paid for a claim to group g for code c on day t; $Keystone_g$ is an indicator of if the doctor was part of the Keystone group (or would be); $Post_t$ indicates whether or not the claim occurred after the 2011 merger of the group; CPT_c is a column vector of dummy variables for each CPT and modifier combination that appears in the estimation sample; and ϵ_{pcg} is the usual residual. The coefficient of interest is β_{merger}, which is the difference-in-differences merger effect.[9] We use the model to run four separate regressions: one for each of Payor A's ESI plans, Payor B's ESI plans, one payor's Medicare Advantage plan, and Payor C's ESI plans.[10]

---

[8] Because the regression controls for CPT code with dummies and prices are measured in logs, it imposes the restriction that, *ceteris paribus*, a particular modifier code changes the price of a procedure by the same percentage, for every procedure, pre-merger and post-merger. Similarly, the price of a particular CPT code differs from another by the same percentage post-merger as pre-merger, regardless of the price level. The model also restricts the percentage change in average price to be identical across all CPT codes. As demonstrated in Figures 1 and 2, the latter assumption is clearly violated. Nevertheless, the measured effect will be the average price change across all CPT codes. (The heterogeneous price changes fit the framework of a random coefficients model, which is consistent for the mean slope parameter. The resultant regression suffers from heteroscedasticity, which we combat with appropriate robust standard errors, clustering on physicians' group. See Greene, 1997, § 15.3.) Each observation is for a particular modifier code/CPT combination consumed by a particular patient on a particular day. The merger's effective dates are the points at which the new payor contracts became effective.

[9] The coefficient β_{merger} actually gives only an approximation to the merger effect. The larger the merger effect, the worse is the approximation. We therefore implement a simple transformation of the coefficients to obtain the exact percentage changes. (See Appendix 1 for details.) The transformation must be done appropriately to avoid what is called the "retransformation bias." We also implement the analogous transformations when reporting the other percentage changes stemming from the regressions. We use the delta method to calculate standard errors.

[10] Because not all payors might offer a Medicare Advantage, we redact the identity of the payor who offered the Medicare Advantage plan.

The claims provided to us to Payor C were for a limited set of CPT codes specific to orthopaedics, so its estimation sample is different from the other estimation samples. Because the physicians' group is the source of variation that identifies the merger effect, we cluster standard errors by group. The regression model expands the estimation sample for Payor A to 89 of the most frequent CPT codes. The broadened estimation sample for Payor B includes 350 of the most frequent codes.[11]

The results appear in Table 3. Some of Keystone's price increases are large relative to the non-Keystone groups'.[12] The difference between Keystone's and non-Keystone's percentage changes for Payors A and B is statistically significant for both payors at any reasonable significance level. The table also reflects the heterogeneity in the merger effect. The Medicaid Advantage plan and Payor C did not experience a price increase.

---

[11] The Payor A and Payor B estimation sample include every code that satisfies these three criteria: appears for Keystone, both pre-and post-merger; appears for non-Keystone, both pre- and post-merger; and has at least 25 observations.

[12] As a robustness test, we used data from only 2008 and 2013 through 2014. If the other groups tended to receive "natural" price increases (e.g., contract renegotiations reflecting inflation or changes in supply and demand conditions unrelated to the merger) early in the range of the data but not later (*e.g.*, closer to 2008 than to the Keystone merger), comparing the pre-merger to post-merger period would tend to overstate the relative price change of Keystone. Even if the other groups had the same sized price increase, they would have had the increased price for much of the pre-merger period, thereby lifting the average of their pre-merger price and reducing the difference in average prices from pre- to post-merger. This robustness test presented no material difference.

Table 3

	Payor A	Payor B ESI	Med Adv	Payor C
Effect	0.246***	0.109***	0.019**	-0.018
	(0.0030)	(0.017)	(0.009)	(0.0137)
Post	0.00873	0.021	0.06***	0.0139
	(0.0179)	(0.012)	(0.009)	(0.013)
Keystone	-0.0352	-0.009	-0.013*	0.041***
	(0.059)	(0.019)	(-0.007)	(0.011)
N=	redacted	redacted	redacted	redacted
CPT, Mod FE	x	x	x	x
Limited Bundle				x

This table reports the coefficient estimates and standard errors for the difference-in-difference merger effects. The reported coefficients and standard errors are corrected for the log-linearization. Standard errors clustered at the physicians' group level.

It is possible to expand the estimation sample further. The very long tail in the distribution of CPT codes implies that a large number of infrequently occurring procedures would have to be included in the estimation sample to more than trivially increase the revenue represented by it.[13] The tradeoff to expanding the estimation sample is a possible reduction in precision of the estimates. That is, a larger estimation sample will include procedures with fewer observations, perhaps reflecting unusual circumstances or increasing the variance in the estimated price changes.[14] We have

[13] For example, nearly 600 separate CPT codes appear four or fewer times.

[14] *Unusual circumstances*: A CPT code that appears just a few times over several years seems to reflect special circumstances by nature of its infrequency. It is also possible that special circumstances will not follow very standardized pricing formulae. *Variance*: Our model has a dummy variable to control for each CPT code/modifier combination. Increasing the number of CPT codes in the estimation sample also increases the number of dummy variables in the model that serve as controls for the various CPT codes. This has the tendency to increase variance. On the other hand, incorporating more CPT codes increases the number of observations in estimating the estimated average price change; this has the tendency to

experimented extensively with different estimation sample sizes. The large merger effects are robust to every estimation sample size we tried.

The magnitudes of the effects measured here are consistent with the previous findings of Dunn and Shapiro (2014). The baseline effect of concentration they calculated implies a ten percent price increase with the merger.[15] The pre-merger Berks orthopaedics market was unusually concentrated, placing it in the 60th percentile for their county-level measures. Post-merger, it was well above the 80th-percentile threshold. They found that more concentrated markets had a steeper relationship between concentration and price. Using the larger coefficients for the more concentrated markets implies a twenty percent increase with the merger. In either case, their magnitudes are consistent with the payor-specific effects we found.

The difference-in-differences results tell us that Keystone's prices increased by more than the average of the controls. The analysis is not complete until we put the *magnitude* of Keystone's price changes in context. If physicians' groups in the control set (unaffected by the merger) experienced price increases akin in size to Keystone's, it would be difficult to conclude that Keystone's price increase is because of the merger rather than a function of the factors caused the other groups' prices to increase. To elaborate, physician contracts specify pricing by a myriad of mechanisms, so we can expect heterogeneity in price changes even when groups have similar supply and demand shocks.[16] Therefore, there is a distribution of price changes, even absent an anticompetitive merger. Consequently, half of price changes are above average even if all price changes are unrelated to changes in the competitive environment faced by

---

reduce variance. The net effect depends in part on the number of additional observations included to offset the extra dummies. With codes that seldom appear, few extra observations would be included.

[15] By way of their footnote 52, the calculation is 0.077 (the IV coefficient for the 80-minute boundary in Table 4) times the difference in the log-HHIs associated with the Keystone merger.

[16] We offer the following as differences in how contracts might specify prices: (1) Physician contracts can specify prices in terms of percentage markup over some base fee schedule, such as Medicare, Medicaid, and proprietary fee schedules. These bases all change differently. Moreover, there is not a single, *e.g.*, "Medicare fee schedule"; the Medicare schedules (and hence changes to them) differ from county to county. (2) Pricing formulae may differ in how they use the base fee schedules. They might specify a fixed year of a schedule as the base and adjust prices annually with a COLA. They might use the current-year schedule as a base, so price changes are dictated by the year-to-year changes in the underlying schedule. (3) The base for a particular group may change from time to time. For example, a group may change its base from Medicare to Medicaid. Similarly, the method of calculating the markup over a base may change from time to time. (4) Carve-outs can be priced separately from the base fee schedules. Carve-outs may vary across groups. Then, the various groups would have different codes that are treated specially, which will result in different price changes for different groups. If a group's carve-outs change over time, it would exacerbate the noise. (5) Physicians may have different negotiating abilities and strategies. See Motheral and Fairman (1997) for more sources of noise in claims data.

providers and payors.[17] An analogy is the growth in infants from six to twelve months of age: Half of the babies will experience above average growth. Finding an above average price change cannot by itself signify an anticompetitive price increase (see Ulrick and Sacher, 2015). The prime issue is judging how atypical Keystone's price increases are among the plethora of changes in the control groups.

To this end, we compare Keystone's pre- to post-merger price changes to those obtained by other large groups over the same period. We use a simple model to calculate each group's price change, to be as inclusive of a broad range of CPT codes and modifiers as possible.[18] Figures 3 (Payor A) and 4 (B) show the results for all groups with at least 250 observations pre- and post-merger. (We omit this figure for Payor C, since it did not experience a relative post-merger price increase and thus there is no reason to establish its price change as being due to the merger.) Keystone's price changes stand out as exceptionally large for both payors. Most price changes are nearly zero or very small (*e.g.*, about 5% or less), compared to Keystone increases of nearly 25% for Payor A data and 15% for Payor B.

There are groups other than Keystone with somewhat large price changes. There are two roughly 15% price changes in the Payor A data and a few groups in the Payor B data have changes in the range of about 8%. Keystone's price increases are about twice as large as the next-highest amongst the controls.

It is also important to note that the estimation samples vary slightly across physician groups because they tend to bill different codes. Therefore, the figure does not purely reflect changes in physician contracts; some of the variation across groups reflects the heterogeneity in changes to the base prices of CPT codes.[19]

[17] This fact assumes a symmetrical distribution of price changes, but the assumption is only for expositional convenience. Even for asymmetrical distributions, the point remains: Some portion of groups will have price changes that are above average (the exact portion depends on the degree of asymmetry) for reasons unrelated to a merger or anticompetitive market powered gained thereby. The conclusion is the same: Absent other evidence regarding a properly constructed but-for world, an above average price change is not in itself sufficient evidence of an anticompetitive merger-effect.
[18] The model is log price regressed on modifier and CPT dummies separately for each physician group.
[19] For example, Medicare does not increase the price of all CPT codes by the same percentage each year. If two groups each receive 125% of Medicare, but the two groups bill different CPT codes, the groups will have divergent average price changes. In fact, Medicare occasionally decreases some prices. A group that substantially bills these codes may have an average price decrease, even if it remains at 125% of Medicare.

Figure 3

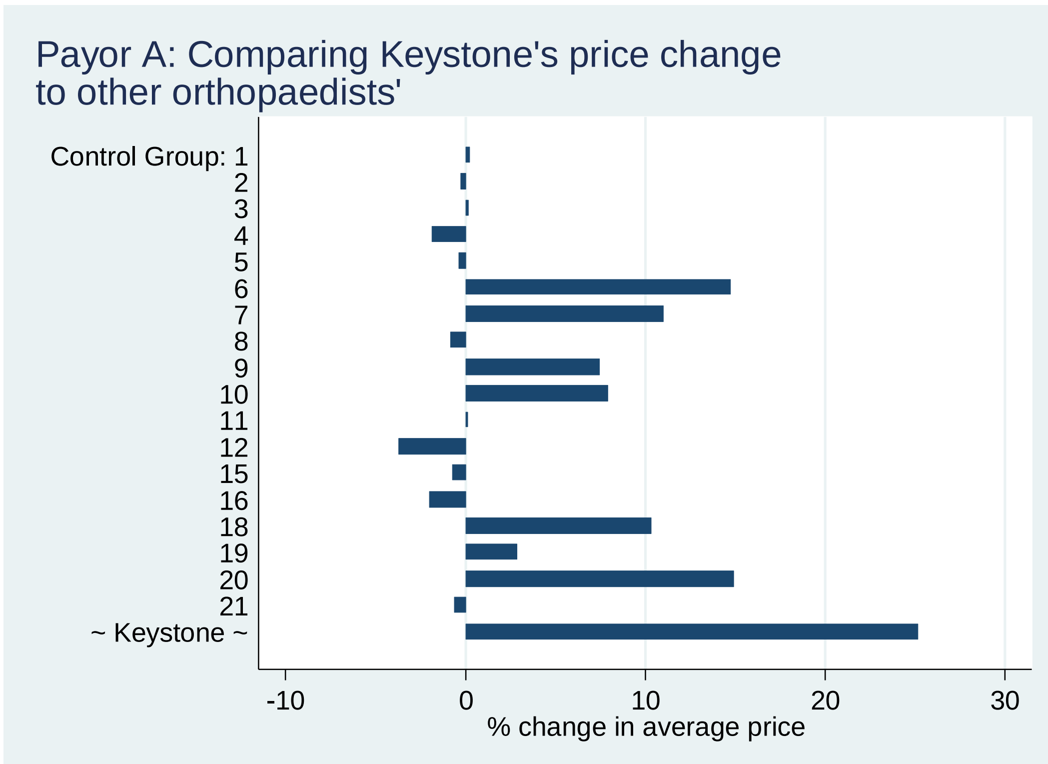
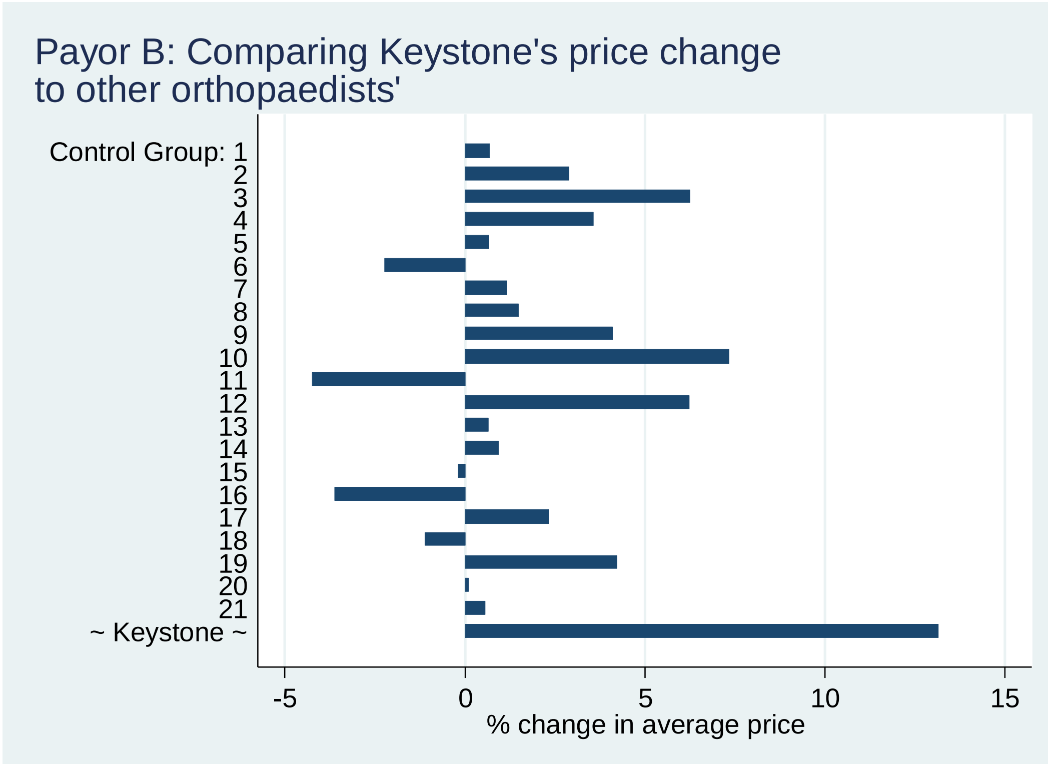


Figure 4



IV. Efficiencies

As mentioned before, an improvement in quality of care could offset an increase in prices. The FTC's complaint states that, to the extent efficiencies exist, they fail to offset the price increases. The complaint is based upon a lengthy investigation, including confidential document and interview evidence. As one way to confirm this claim with masked data, we consider patient flows.[20] We might expect that if the merged Keystone group improved care, it would draw more patients than before. This is, roughly speaking, a revealed preference approach to efficiency analysis. It has also been used before in Vita and Sacher (2001), to evaluate potential efficiencies due to a hospital merger.

Visits to Keystone largely comprise those to orthopaedists in in Berks County, so we use Berks County as a proxy for Keystone. That is, we measure the fraction of beneficiaries that lived in Berks County who stayed in the county to receive orthopaedic

[20] More direct, masked approaches are not available to us because the claims data do not report measures of quality that might be relevant to orthopaedic care, e.g., whether a patient limped after surgery or the extent he regained range-of-motion.

care.   A Keystone sufficiently higher in quality to more than offset its price increase to patients should persuade marginal Berks County beneficiaries to remain in the county (with Keystone) for care. That is, shares of county residents remaining in the county should increase. Because patients generally pay a small fraction of a price increase, it should not take much of an improvement in quality to overcome any potential price increase.[21]

We calculated the portion of Berks County residents who stay in the county for orthopaedic care separately for "new patient office visits" (CPT codes 9920_), "established patient office visits" (CPT codes 9921_), and all other services.[22] Payor B gave us service provider counties, so we focus on those. Calculations for Payor A provide similar qualitative results but are omitted from our discussion.

Our data include providers only for Berks and several local counties.[23] Some residents leave the region entirely; we do not know how many. One might argue that patients traveling so far face special circumstances not shared by those who stay, such as needing specialized care, an injury that occurred while out of town, or a desire to stay with family during treatment. If so, traveling patients face anomalous choices in providers, and their omission is of little consequence. On the other hand, if traveling patients resemble the typical Berks County inhabitant, our analysis understates the relevant outflows.

For Payor B, we calculated the patient flows separately for the pre- and post-merger periods. This is shown in Table 5. The share of Payor B's patients staying in Berks County is around 70% and does not change much pre- and post-merger. The

[21] Patients usually pay a copay or coinsurance. Therefore, the patients' share of a price increase is minimal.

[22] The data include the patient's home ZIP code rather than the patient's home county. ZIP codes do not lie strictly within counties. In some instances, the overlap is so small, a ZIP code does not deserve to be counted as Berks County. For example, only 0.4% of zip 18031 is in Berks County. (Information on ZIP code compositions can be obtained here: http://www.city-data.com/zips/xxxxx.html, where the xxxxx is the desired five-digit ZIP code.) In other cases an overlap, if not complete, is so great that it would be nonsensical to treat the ZIP as not being part of Berks County; *e.g.*, ZIP 19512 is 98.51% in Berks County. Other categorizations are not as obvious. For the flow analysis above, we classify a ZIP code as Berks if at least 25% of its population is in the county. (As a robustness check, we counted any ZIP having any overlap with Berks County as part of Berks County; the results were nearly identical.) The classification scheme means that several observations counted as Berks County residents seeking care outside the county *actually reflect flows from residents residing outside the county*. The implication is that this exercise is conservative and understates the portion staying within Berks County.

[23] Due to confidentiality concerns, we do not disclose the counties under consideration for each payor.

differences for the E/M visit codes were not statistically significant at the 95% level. The difference for the other codes was statistically significant (t-value of 4) but show an increase of the outflow from Berks County.

Table 5

Payor B: Percent of Berks County patients that stay in Berks County, 2010 and 2011

	Time relative to new, post-merger contract	
	Before	After
New patient visits [¤]	75%	75%
Established patient visits [¤]	69%	70%
All other codes [?]	72%	69%

Notes: All providors, in-network claims.
[¤]percentages based on number of visits.
[?]percentage based on revenue.
"Before" refers to time before 1st week of 2010.
"After" refers to 1st week of 2011 and after.
,

6. Alternative Data Sources

Much recent attention has focused on the use of claims data to generate prices in health care markets. Two important sources, the Health Care Cost Institute (HCCI) and Truven MarketScan data, both collect individual claims data from insurers. These claims data reflect the payments made to health care providers and are commercially distributed to researchers. Cooper and Gaynor (2016) use the HCCI data to compare hospital prices by hospital market concentration, among other things; Dunn and Shapiro (2014) use Truven data similarly for physician markets.

Since these are not all-payor databases, they may only have claims for some of the payors in a particular region, and the fraction of payors represented in their claims may vary by region. As we have demonstrated, the consequences of a merger may not be felt symmetrically across payors. Changes in market concentration may generate anti-competitive price effects, but these effects may not be measured by either the HCCI or Truven data if the affected payors do not submit their claims to either of these databases. Given there is heterogeneity in merger effects, it is possible a merger that adversely affects some but not all payors may (or may not) affect the payors in a limited data set. This can lead to discrepancies when measuring the overall effects of a merger. To evaluate the possibility of such a problem, we estimate the same merger effects model and price plots in the Truven MarketScan data as we did for the payor data above. The MarketScan data reflect claims for medical services paid for by employers and insurance carriers throughout the United States. The MarketScan sample is a convenience sample of patients whose insurer (employer or carrier) provided the data to Truven. Our data come from the outpatient portion of the Commercial Claims database. We focus on the years 2009-13, since they provide a two year window before and after the merger. Because the Truven MarketScan only provide the MSA of the provider (and no further information about the provider), we compare the average price paid for a claim in the Reading MSA before and after the merger. Data restrictions due to confidentiality issues do not allow us to compare Reading prices to other comparable MSAs in southeast Pennsylvania, such as York and Lancaster. This means that the estimates we report below are "difference" estimates, since we have no other differences to separate out. These are directly comparable to the difference estimates in Figures 3 and 4.

Figures 8 and 9 present the average price paid for the two evaluation and management CPT codes (99213 and 99212), plotting the prices for the Reading MSA. A few things are apparent for the figures. For the 99213 in particular, there may be a modest increase in price around January 2011, but that is small compared to jump up in price near the end of 2009. This pattern is not reflected in the average price for the 99212, which is flat.

There may be changes in insurer or provider composition in the MarketScan data. That is, some of the price variation for 99213 shown in the Figure 8 could be due to changes in who provides the data, which patients they see, or which doctors are covered. Further, there may be mechanical reasons unrelated to market position, such as changes to Medicare's reimbursement schedule, which often serve as a reference point in physician contracts, that drive discrete shifts in physician reimbursement. This highlights the importance of appropriate controls and appropriate control groups.

Figure 8

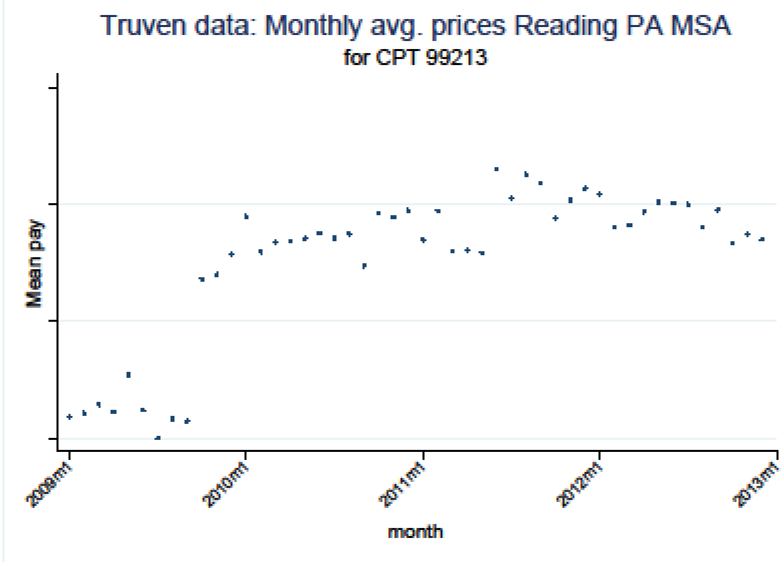


Figure 9

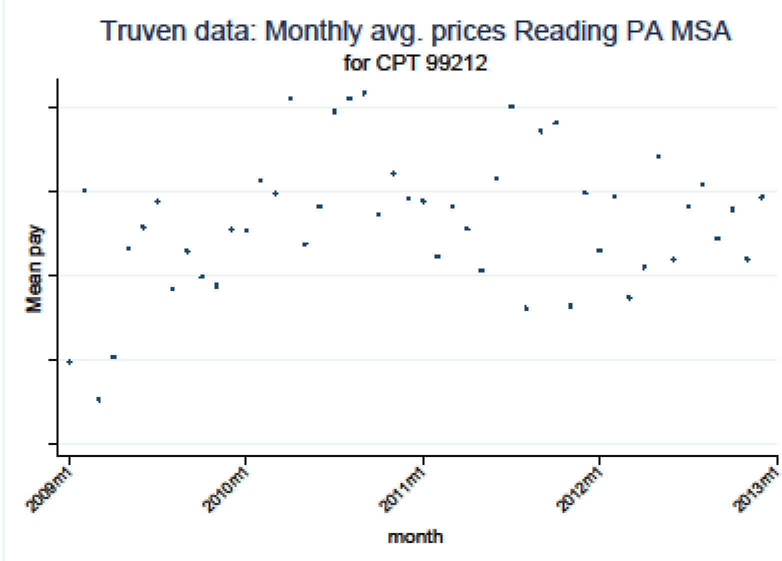


Table 6 reports the difference estimate for the merger effect as measured in the Truven claims data. The most naïve difference specification, using all of the data and with no controls (effectively a raw difference), does generate a merger effect of 13 percent. However, as we narrow the window and control for CPT codes with fixed effects, the change in prices associated with the merger falls to just over one percent, with a 95 percent confidence interval that includes zero but excludes prices changes much larger than four percent. As we have shown, physician prices do vary over time and by service (i.e., CPT codes) in ways that may not reflect changes in market position. Thus, it is important to control for what we can when a control group is not readily available.

In sum, the figures from the Truven data suggest that there are no changes in price when Truven claims data are used for our preferred specification, even though our estimates from other payors indicated an anti-competitive effect.

Table 6

Effect	0.13***	0.067**	0.042***	0.011
	(0.015)	(0.0022)	(0.019)	(0.015)
Constant	3.084***	3.87***	3.82	3.84
	(0.010)	(0.013)	(0.016)	(0.023()
N=	24,646	12,244	24,646	12,244
CPT, Mod FE			X	X
Years	2009-13	2010-11	2009-13	2010-11

This table reports the coefficient estimates and standard errors for the difference merger effects using Truven MarketScan data. Four specifications are included: whether or not CPT fixed effects are included, and width of the window around the merger date. Standard errors are robust to heteroskedasticity.

These findings underscore the limitations of claims databases that rely on a fraction of the payors in a market. Because firms may choose to exploit their market power against some, but not all, of the payors in a market, the use of un-representative claims data may lead to false negatives when assessing the consequences of market concentration.

7. Conclusion

We study the relationship between changes in price and concentration in the instance of a six-to-one merger among orthopaedists in Reading, PA. We measure prices with transaction-level claims data that are available to us for three commercial payors. Via difference-in-differences methods, we find the merged practice raised prices for two of the three payors. Depending on the payor and method of calculating the relative price changes, we find that these prices rose by up to twelve to twenty percentage points more than the expected change but-for the merger. Moreover, these price changes were substantively larger than those by other, nearby orthopaedists. The price increases were not uniform, however, because one payor's data includes two plan types, only one of which incurred the price increase. Further, the third payor did not experience any merger effect on prices.

Our findings thus present several lessons. First, merger effects may not be uniform across payors or plans. This implies that efforts to identify effects using only a subset of the payors may provide an incomplete picture of the consequences of the merger. Further, to the extent that we have measured it here, physicians groups do not necessarily induce efficiencies.

Appendix-- Calculating percentage changes from log-price regressions

Our basic difference-in-differences model uses the log of price to estimate the merger's impact:

$$\ln(Y) = a + b[postmerger] + c[postmerger][keystone] + d[keystone] + \epsilon$$

where Y is price of a CPT and $[postmerger]$ and $[keystone]$ are dummy variables denoting whether an observation belongs to Keystone or is for the post-merger period. Note that we omit other controls (including dummy variables for modifier codes), since they will vanish. The coefficient c is an approximation to the percentage point difference in Keystone's and non-Keystone's percent changes in price. For large percentage changes, this approximation can be quite bad. We therefore report the exact percentage point difference in this memo, which we obtain with a transformation on the model's parameters.

Obtaining the exact percentage changes requires some care, to avoid "retransformation bias."[24] In this appendix, we detail the correct transformation and supporting algebra.

To begin, rewrite the log-price model with price in levels:

$$\Rightarrow Y = e^{a} e^{b[postmerger]} e^{c[postmerger][keystone]} e^{d[keystone]} e^{\epsilon}$$
$$\Rightarrow E(Y|keystone, postmerger) = e^{a} e^{b[postmerger]} e^{c[postmerger][keystone]} e^{d[keystone]} E(e^{\epsilon})$$

Therefore, the expected value in logs is

(1) $\mathrm{E}[\ln(Y)] = a + b[postmerger] + c[postmerger][keystone] + d[keystone]$

[24] Retransformation bias is the bias in the predicted values that can occur when inappropriately retrieving expected price by taking $e^{\widehat{\ln[Y]}}$, where $\widehat{\ln[Y]}$ is the predicted value in logs. For further discussion on this problem, see, *e.g.*, Michael C. Newman, (1993) "Regression Analysis of Log-Transformed Data: Statistical Bias and Its Correction," *Environmental Taxicology and Chemistry,* 12, pp. 1129 - 1133. Note that Newman offers a method of obtaining an unbiased prediction of Y based on assuming normal residuals. Because we seek to estimate percentage changes and not the price levels, we take a more direct approach that requires no such assumption.

To calculate the exact percentage changes, we will need to know the expected values of prices, for each of the four relevant bins (the bins are pre-merger Keystone, post-merger Keystone, pre-merger non-Keystone, post-merger non-Keystone). It is useful to use more compact notation when identifying the bins: Let n denote non-Keystone, k denote Keystone, r denote pre-merger, and s denote post-merger. Then expected values *in logs* for each relevant bin are

<u>Non-Keystone, premerger:</u> $E(\ln Y\,|n,r) = a$

<u>Non-Keystone, post-merger:</u> $E(\ln Y\,|n,s) = a + b$

<u>Keystone, premerger:</u> $E(\ln Y\ \ |k,r) = a + d$

<u>Keystone, post-merger:</u> $E(\ln Y\,|k,s) = a + b + d + c$

In levels they are

<u>Non-Keystone, premerger:</u> $E(Y|n,r) = e^{a}E(e^{\epsilon})$

<u>Non-Keystone, post-merger:</u> $E(Y|n,s) = e^{a+b}E(e^{\epsilon})$

<u>Keystone, premerger:</u> $E(Y|k,r) = e^{a+d}E(e^{\epsilon})$

<u>Keystone, post-merger:</u> $E(Y|k,s) = e^{a+b+d+c}E(e^{\epsilon})$

Pre-merger, the percentage difference between Keystone and non-Keystone is

$$\frac{E(Y|k,r) - E(Y|n,r)}{E(Y|n,r)} = \frac{e^{a+d}E(e^{\epsilon}) - e^{a}E(e^{\epsilon})}{e^{a}E(e^{\epsilon})} = e^{d} - 1$$

The percentage change ($\%\Delta/100$) in non-Keystone expected price is

$$\frac{E(Y|n,s) - E(Y|n,r)}{E(Y|n,r)} = \frac{e^{a+b}E(e^{\epsilon}) - e^{a}E(e^{\epsilon})}{e^{a}E(e^{\epsilon})} = e^{a+b-a} - 1 = e^{b} - 1$$

Similarly, the percentage change ($\%\Delta/100$) in Keystone expected price is

$$e^{b+c} - 1$$

So, the percentage-point-difference in non-Keystone and Keystone changes is

$$(e^{b+c} - e^{b})100$$

Or, approximately, the percentage change ($\%\Delta/100$) in non-Keystone expected price is b. Approximately, the percentage change ($\%\Delta/100$) in non-Keystone expected price is $b + c$. Therefore, approximately, the percentage point difference in non-Keystone and Keystone is $100c$.

Standard errors are calculated with the delta method. So, for example, to calculate the standard error for the percentage-point-difference in non-Keystone, note that the function of the estimated regression parameters is

$$g(\hat{b}, \hat{c}) = (e^{\hat{b}+\hat{c}} - e^{\hat{b}})$$

where $\hat{b}$ and $\hat{c}$ are estimated regression parameters in the difference-in-differences regression (1). By a Taylor series expansion around b and c,

$$\mathrm{Var}[g(\hat{b}, \hat{c})] \approx [g_b(b)]^2 \,\mathrm{Var}(\hat{b}) + [g_c(c)]^2 \,\mathrm{Var}(\hat{c}) + 2 g_b(b) g_c(c) \,\mathrm{Cov}(\hat{b}, \hat{c})$$

where $g_b(b)$ and $g_c(c)$ are the following derivatives:

$$g_b(b) = \left.\frac{\partial g(\hat{b}, \hat{c})}{\partial \hat{b}}\right|_{\hat{b}=b, \hat{c}=c} = (e^{b+c} - e^{b})$$

$$g_c(c) = \left.\frac{\partial g(\hat{b}, \hat{c})}{\partial \hat{c}}\right|_{\hat{b}=b\ , \hat{c}=c} = e^{b+c}$$

Therefore,

$$\text{(2)}\quad \mathrm{Var}[g(\hat{b}, \hat{c})] \approx (e^{b+c} - e^{b})^2 \,\mathrm{Var}(\hat{b}) + e^{2(b+c)} \,\mathrm{Var}(\hat{c}) + 2(e^{b+c} - e^{b})(e^{b+c})\mathrm{Cov}(\hat{b}, \hat{c})$$

To feasibly calculate (2), obtain $\hat{b}$, $\hat{c}$, $\mathrm{Var}(\hat{b})$, $\mathrm{Var}(\hat{c})$, and $\mathrm{Cov}(\hat{b}, \hat{c})$ by ordinary least squares, and apply the Slutsky theorem to substitute $\hat{b}$ and $\hat{c}$ for b and c.

References

Capps, Cory, David Dranove, and Mark Satterthwaite. 2003. Competition and Market Power in Option Demand Markets. RAND Journal of Economics 34:737–63.

Carlin, Caroline S., Roger Feldman, and Bryan Dowd. "The impact of provider consolidation on physician prices." *Health Economics* (2017).

Casalino, Lawrence P., et al. "Small primary care physician practices have low rates of preventable hospital admissions." *Health Affairs* (2014): 10-1377.

Clemens, Jeffrey P., and Joshua D. Gottlieb. "Bargaining in the Shadow of a Giant: Medicare's Influence on Private Payment Systems." *Journal of Political Economy* (2017).

Clemens, Jeffrey, and Joshua D. Gottlieb. "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?." *The American economic review* 104.4 (2014): 1320.

Cooper, Z., Craig, S. V., Gaynor, M., & Van Reenen, J. (2015). *The price ain't right? Hospital prices and health spending on the privately insured* (No. w21815). National Bureau of Economic Research.

Dunn, Abe, and Adam Hale Shapiro. "Do physicians possess market power?." *The Journal of Law and Economics* 57.1 (2014): 159-193.

Dranove, David, Mark Shanley, and William D. White. 1993. Price and Concentration in Hospital Markets: The Switch from Patient-Driven to Payer-Driven Competition. Journal of Law and Economics 36:179–204.

Greene, William H. (1997), *Econometric Analysis*, 3rd ed., Upper Saddle River, NJ: Prentice Hall.

Haas-Wilson, D., & Garmon, C. (2011). Hospital mergers and competitive effects: Two retrospective analyses. *International Journal of the Economics of Business*, *18*(1), 17-32.

Herman, Douglas A. and Ulrick, Shawn W. (2012), "Measuring Price Dynamics: A Guide to Understanding Payer-Physician Claims Data," *Journal of Social and Economic Measurement*.

Keeler, Emmett B., Glenn Melnick, and Jack Zwanziger. 1999. The Changing Effects of Competition on Non-profit and For-Profit Hospital Pricing Behavior. Journal of Health Economics 18:69–86.

Lynk, William J. 1995. Nonprofit Hospital Mergers and the Exercise of Market Power. Journal of Law and Economics 38:437–61.

Motheral, B.R. and K.A. Fairman (1997) "The use of claims databases for outcomes research: Rationale, challenges, and strategies," *Clinical Therapeutics* 19(2), pp. 346-366.

Noether, Monica. 1988. Competition among Hospitals. Journal of Health Economics 7: 259–84.

Tenn, S. (2011). The price effects of hospital mergers: a case study of the Sutter–Summit transaction. *International Journal of the Economics of Business*, *18*(1), 65-82.

Thompson, A. (2011). The effect of hospital mergers on inpatient prices: a case study of the New Hanover–Cape Fear transaction. *International Journal of the Economics of Business*, *18*(1), 91-101.

Town, Robert, and Gregory Vistnes. 2001. Hospital Competition in HMO Networks. Journal of Health Economics 20:733–53.

Ulrick, Shawn and Seth Sacher, (*forthcoming*) "Inferring Anticompetitive Price Effects from Difference-in-Difference Analysis: A Caveat," *Journal of Competition Law & Economics*

Vita, M. G., & Sacher, S. (2001). The competitive effects of not-for-profit hospital mergers: A case study. *Journal of Industrial Economics*, 63-84

Welch, W. Pete, et al. "Proportion of physicians in large group practices continued to grow in 2009–11." *Health Affairs* 32.9 (2013): 1659-1666.