The Impact of Preferred Provider Incentives on Demand and Negotiated Prices *

Calvin A. Ackley†

February 28, 2024

Abstract

This paper studies the impact of a preferred network design on procedure-level spending for lab services. This plan structure, termed the "Site of Service" design, employs a two-tiered cost-sharing schedule for lab tests: patients incur no out-of-pocket costs at preferred providers but face a deductible at non-preferred providers. Using event-study methods and administrative data on two large carriers, I find that these tiered incentives lead to a considerable reduction in the price paid per lab, with effect sizes ranging from 14% to 36% across groups and time. I find that the preferred provider program generates savings both by steering consumers toward less expensive providers and by putting downward pressure on negotiated prices. Notably, I present explicit causal evidence linking the preferred network to substantial negotiated price cuts. I find that these price dynamics account for about half of the overall program savings while the steering mechanism accounts for the remainder.

1 Introduction

To address rising health care costs and disperse prices, insurers continue to innovate plan designs in an effort to encourage price-shopping among consumers. The most salient trend in this direction has been the marked increase in deductible levels. The share of the private market enrolled in high-deductible health plans (HDHPs) has soared from 8% in

*I am grateful to Randy Ellis, Tal Gross, Marc Rysman, Jihye Jeon, Chelsea Carter, Elena Prager, Abe Dunn, and Lasanthi Fernando for helpful comments. I also thank seminar participants at the Bureau of Economics Analysis, the Congressional Budget Office, and the Census Bureau for additional feedback. I also thank Mary Fields, the New Hampshire Department of Health and Human Services, and the New Hampshire Department of Insurance for providing the data. All analysis, conclusions, and recommendations presented herein are solely my own, and do not reflect the views of the New Hampshire Department of Health and Human Services, the New Hampshire Department of Insurance, or the U.S. Bureau of Economic Analysis. A large portion of this work was done at Boston University, prior to joining the Bureau of Economic Analysis.

†U.S. Bureau of Economic Analysis, Suitland, MD. Calvin.Ackley@bea.gov

2008 to 29% in 2018 (Claxton et al., 2018).[1] While these plans appear to reduce spending overall, studies typically find that HDHPs are largely ineffective in steering patients toward low-cost providers (Brot-Goldberg et al., 2015). Network-based designs, such as narrow networks, tiered or preferred networks, and reference pricing, provide alternatives to HDHPs that may offer consumers greater navigational ease. Importantly, designs with a strong steering capacity have the potential to reduce expenditure through both a direct change in demand, and indirectly, by exerting downward pressure on equilibrium prices.

In this paper, I study the impact of a tiered network design on procedure-level spending for lab services. Lab tests are the highest-volume procedure type in the United States, and their prices vary substantially across providers (Song et al., 2021; Robinson et al., 2016).[2] In New Hampshire, for example, the price of a lipid panel varies from around $10 at some independent labs to well over $100 at some hospitals. Given this high volume and variability, insurers and patients have opportunities to achieve considerable savings through both demand-side and supply-side channels. The plan design I study, called the "Site of Service" feature, employs a tiered cost-sharing schedule to generate clear incentives for patients to use low-cost providers. Under the Site of Service design, patients pay $0 at preferred lab providers but face a deductible at non-preferred providers.

This tiered design offers a number of potential advantages over deductible- and coinsurance-based plans in terms of yielding savings. On the demand side, the tiered design may be more effective at steering consumers by using clear and significant financial incentives (Prager, 2020; Ackley, 2022).[3] On the supply side, the preferred network setup offers insurers an additional bargaining lever to use in price negotiations with providers. Using administrative data from New Hampshire, I provide new evidence on the effectiveness of the tiered design along these dimensions.

I first estimate the effect of Anthem New Hampshire's Site of Service preferred network on the price paid per lab test. The price paid per test reflects the aggregate dynamics of both demand-side and supply-side forces. To study this measure, I utilize an event-study framework with both procedure-based and carrier-based control groups that are not subject to the tiered design. I examine two different treatment groups, one based on small-group plans and the other based on large-group plans, as treatment penetration differed slightly across groups. Overall, I find that the preferred network led to a considerable reduction in per-lab payments for both of the treated groups in the sample. For the small-group population, estimates imply a reduction in spending per lab of 36% by 2015. For the large-group population, the estimated reduction is 32% by 2015. Altogether,

[1] More generally, the fraction of privately insured individuals whose primary deductible level is greater than $1,000 has grown from 38% in 2013 to 58% in 2018 (Kaiser Family Foundation, 2018).

[2] In addition, an increasing share of lab spending is occurring out-of-network, which is typically associated with both higher prices and out-of-pocket costs for patients (Song et al., 2021)

[3] Of course, a host of other factors may impact a plan's ability to steer, including the geographic distribution of providers and their perceived quality.

these estimates imply that Anthem saved around \$48 million on lab tests relative to the counterfactual in which cost sharing for labs followed the trajectory of x-rays over the period.

I next study the demand side directly by estimating the extent to which the tiered incentives steer patients toward preferred labs. The share of tests performed at independent labs rose considerably over the sample period for both Anthem groups. Difference-in-differences estimates indicate that the preferred network design is associated with a 6 to 12 percentage point increase in the likelihood of using independent, preferred labs, representing almost a doubling relative to baseline.

I find that, relative to the trend for x-rays, the preferred provider design is associated with a net reduction in procedure-level cost sharing and out-of-pocket costs for patients. Moreover, I do not find much evidence that the design substantially impacts the total quantity of services on the extensive margin, in contrast to what much of the evidence on HDHPs shows.

I next analyze the dynamics of negotiated prices over the sample period, presenting some of the first causal evidence linking tiered-type designs to negotiated price changes. Theoretically, the Site of Service design offers several mechanisms through which downward pricing pressure may be exerted, all of which are rooted in the design's efficacy in steering patients. While there is some prior evidence on the steering capacity of the tiered design, this paper provides some of the first empirical evidence on the interplay of steering and negotiated prices.

To quantify changes in negotiated prices over the sample period, I estimate both event-study-based models as well as traditional Laspereyes-type price indices. Overall, I find strong evidence that Anthem's preferred network design generated downward pressure on negotiated prices. Across all providers, I estimate that prices fell by 12%-15% after the rollout of the preferred provider program. The most dramatic price changes are associated with providers that explicitly renegotiate lower rates to acquire preferred status. Among these providers, prices collectively drop by around 50% by the end of the sample period, evidence of a direct causal link between the network and prices.

Finally, I decompose the aggregate changes in payments per procedure into a direct steering component and a negotiated price component. The steering component reflects changes in per-lab spending that are accounted for by changes in provider demand, holding negotiated prices fixed at a prior level. The negotiated price component captures changes in negotiated prices, holding demand fixed. Broadly, I find that the steering component and the price component each account for around half of the overall decline in payments per lab, although there is heterogeneity across procedures and the two Anthem groups. The steering effect is somewhat more important for the small-group population than in the large-group population.

Overall, these results indicate that the preferred network design can generate downward pressure on medical spending through both demand-side and supply-side channels.

While this paper focuses on lab services and the commercially insured population, the mechanisms examined here are applicable for a broad range of services and payers. Understanding how best to deploy these types of incentives is crucial for insurers and policymakers seeking to contain costs, especially in the face of increasing provider consolidation (Baker et al., 2014; Lin et el., 2021; Fulton, 2017; Cooper et al., 2019).

1.1 Related Literature

This paper fits most directly into the small but growing literature on preferred networks and tiered cost sharing. Preferred/tiered provider designs differ from traditional deductible and coinsurance-based designs in that cost sharing varies discretely across groups of providers in the network.[4] Typically, plans of this style have either two or three groups of providers, with preferred providers entailing lower cost sharing. Prior work on these plans generally indicates that the tiered-type design is effective in steering. Sinaiko and Rosenthal (2014) find that a tiered physician network induces a small demand change among new patients. Frank et al. (2015) present a similar finding for hospital choice under a tiered network. Prager (2020) estimates a small but significant effect of tiered copayments on hospital choice among patients in Massachusetts. Starc and Swanson (2021) study both the demand-side and supply-side effects of preferred pharmacy networks in Medicare Part D. The authors find that preferred network plans pay around 2% less in drug prices relative to other plans, and that this difference is mostly explained by lower negotiated prices, rather than steering within a plan.

Most closely related to this paper is the prior study of Ackley (2022) on the Site of Service Program, although this paper departs from the prior study in many important ways. Most broadly, while Ackley (2022) focuses purely on the demand-side effects of the tiered network, this paper examines both the supply-side and demand-side, and provides novel decomposition of aggregate effects into a steering component and negotiated price component. The principal finding of Ackley (2022) is that the tiered design produces stronger price-responsiveness among patients than other popular designs, and is thereby more effective at steering. The principal finding of this paper is that the tiered design can generate considerable savings at the insurer level, both through steering and by impacting negotiated prices. Consistent with these broad differences in scope, the two papers also differ in both sample construction and research design. This paper focuses on lab tests, and utilizes both procedure-based and carrier-based control groups to identify negotiated price effects. In contrast, Ackley (2022) examines a much smaller sample of solely gastrointestinal procedures and focuses on estimating demand parameters.

A similar design to the preferred network I study here is reference pricing. In a

[4] Another type of network plan is the limited or narrow network design. Gruber and McKnight (2016), studying a narrow network plan offered by the Massachusetts Group Insurance Commission, estimate a 40% reduction in expenditures among individuals who switched to the plan.

reference pricing scheme, patients pay the full difference between a provider's price and a reference price established by the insurer, if the difference is positive. For providers that fall under the reference price, patients have full coverage. So, like the Site of Service design, reference pricing establishes large differences in cost sharing across two groups of providers, ideally simplifying the choice environment for patients.[5]

The bulk of the evidence on reference pricing thus far indicates that it is very effective in steering. Using data from a large firm that implemented reference pricing for lab tests, Robinson et al. (2016) find that the program led to about a 32% decrease in the price paid per test. Other work on the reference pricing program implemented by the California Public Employees Retirement System CalPERS also finds non-negligible steering effects (Robinson et al., 2015; Whaley et al., 2017). Aouad et al. (2021) study the effects of reference pricing on the distribution of spending for a range of procedures, finding the most significant reduction in the right tail of the insurer's spending distribution. There is less evidence on reference pricing's impact on negotiated prices, and the findings are somewhat mixed. Robinson and Brown (2013) report that reference pricing in the CalPERS system led to lower prices for knee and hip replacement. Also examining the CalPERS program, Whaley and Brown (2018) find modest price cuts for outpatient surgeries at ambulatory surgery centers, but not much change in hospital outpatient prices.

Relative to the existing work on tiered network designs and reference pricing, this paper offers several important innovations, both in research design and key findings. While the majority of existing work focuses on demand-side effects, this paper provides novel evidence on both negotiated price effects and steering effects. In terms of identification, this paper is the first to exploit variation across both insurers and procedure groups in tiered incentives. This design accommodates dynamic changes in bargaining power and other unobserveables at the insurer level, and is therefore particularly attractive for identifying negotiated price effects. In contrast, Robinson and Brown (2013) and Whaley and Brown (2018) examine only treated procedures covered by Anthem, while Starc and Swanson (2021) study drug claims covered by Medicare Part D. Additionally, this paper documents strong causal evidence in the form of exact price cuts made by some providers to obtain preferred status. Finally, this paper presents a novel, comprehensive decomposition of aggregate changes into a steering component and a negotiated price component, shedding new light on the mechanisms underlying the success of these preferred-provider-type designs.

1.2 Institutional Background

In New Hampshire, both insurers and state regulators have been active in promoting price transparency and consumer-focused plan designs. Each year, the state insurance

[5] A related design is centers-of-excellence contracting, where insurers use cost-sharing incentives to steer patients toward a small set of high-quality, low-cost providers (Robinson and MacPherson, 2012)

department convenes a meeting with various stakeholders to discuss initiatives in this area. In 2007, the state launched a price transparency website on which prospective patients could obtain price estimates for various procedures at different providers.

In terms of benefit design, HDHPs have been the most popular innovation. Indeed, HDHP enrollment increased from 1.5% of the commercial market in 2006 to about 47% in 2016 (NHID, 2008; Samgula et al., 2017). On the State Insurance Exchange, narrow network plans have been commonly offered.[6] In this paper, I study a tiered pricing feature introduced by Anthem in 2009-2010, termed the "Site of Service" design. These plans feature two tiers of in-network providers, which may be referred to as "preferred" and "non-preferred" providers. This type of tiered cost sharing covers an array of common services such as lab tests and outpatient surgeries. For lab services, which are the focus of this paper, patients incur zero out-of-pocket costs at preferred providers, but they face the full negotiated price, through a deductible, at non-preferred providers.

Motivated by the growing costs and price dispersion associated with outpatient services, Anthem first experimented with the Site of Service tiered design at the end of 2009 (Highland, 2012). Since 2010, the Site of Service tiered design has become standardized in Anthem's small-group plans, and has become popular in the large-group market as well (Gorman et al., 2013). Representatives from Anthem report that the Site of Service program has had a favorable impact of unit costs, and that this has helped constrain premiums (NHID, 2013; Grenier et al., 2013; Tu and Gourevitch, 2014). This has reportedly been particularly impactful in the small-group market, where purchasers may be particularly sensitive to premiums (NHID, 2013; Grenier et al., 2013; Tu and Gourevitch, 2014). This premium effect, along with dominance of fully-insured products in the small-group market, likely explains the greater early penetration of the Site of Service design in this group.

In 2013, Harvard Pilgrim, another large insurer in the state, introduced its own version of the tiered design called "LP" (for low-cost provider) (Tu and Gourevitch, 2014). By 2016, about 72% of the small-group market and 42% of the large-group fully insured market were covered by either Anthem's or Harvard Pilgrim's tiered designs (Smagula et al., 2016) .[7]

The Site of Service structure is a tiered-network-type design where cost sharing varies discontinuously across provider groups. For lab services, patients pay $0 at preferred providers but face a deductible at non-preferred providers. The Site of Service feature is distinct from traditional preferred provider organizations (PPOs) in that it is implemented in both HMO- and POS-type plans by defining a tiered schedule over in-network providers. Moreover, like other tiered-network designs, the Site of Service program offers

[6]Notably, in 2014, Anthem's limited network HMO was the sole plan offered on the state exchange (Smagula el al., 2016).

[7]In 2014, the New Hampshire state government, the state's largest employer, moved all covered employees onto plans with Anthem's Site of Service tiered design (Tu and Gourevitch, 2014).

a broader network than traditional limited-network HMO plans. The reference pricing design implemented by CalPERS is similar to the Site of Service design in terms of network breadth and in having essentially two tiers of cost sharing. The main difference between the two designs is that out-of-pocket costs in a reference pricing plan depend explicitly on an established reference price, whereas, in the Site of Service design, they depend only on preferred status.[8]

Figure B1 displays an advertisement for Anthem's Site of Service tiered design. This figure highlights the simplicity of the incentive scheme for each of the services described, in particular for lab services. Figure B2 shows another advertisement that details the names and locations of preferred lab providers. Importantly, preferred providers are dispersed throughout the state, ensuring that enrollees residing in all regions have reasonable access to these facilities. At the outset of the Site of Service program, independent, non-hospital providers were designated as preferred. After the initial rollout, hospital-based providers began to be added to the preferred tier after renegotiated lower rates (NHID, 2013; Tu and Gourevitch, 2014).

A frequent concern raised about tiered-type designs is that they may steer patients toward providers of lower quality. As described in Figure B3, Anthem bases tiers on both cost as well as quality benchmarks. This is similar to how low-cost providers are selected in many reference-pricing designs (Robinson and MacPherson, 2012). For lab services in particular, Anthem reports that independent labs offer the same level of quality as other outpatient labs, implying that associated cost savings does not come at the expense of quality results.[9]

1.3 Data and Study Design

The primary data source for this research is the New Hampshire All Payer Claims Database (NH-APCD). These data contain rich detail on plans, diagnoses, procedures, payments, and providers for all medical services rendered to individuals with a New Hampshire-based insurance plan. Plan information includes details on the carrier, the type of plan, and whether the carrier is operating an administrative services only (ASO) contract or a fully-insured contract. On the patient side, the data includes information on age, sex, race, zip code, and relationship to the insurance policyholder. In addition, I use each patient's history of ICD-9-CM and ICD-10-CM diagnosis codes to compute the Charlson Comorbidity Index (CCI) for each patient (Charlson, 1987). Each claim includes a procedure code as well as the amount billed by the provider for this service, the amount paid by the insurer, and any deductible, coinsurance, and copayments paid by the patient. The sum of payments made by the insurer and the patient represents the negotiated rate

[8] Anthem does not explicitly specify a reference-type price which determines preferred status. Rather, this status appears to depend on both the type of facility as well as the willingness to negotiate low rates.

[9] See Figure B4, for instance.

between the provider and the insurer.

The goal of this paper is to study the impact of Anthem's Site of Service preferred network on prices paid for lab services. To do so, I exploit longitudinal variation in preferred network incentives across two large carriers and two large procedure groups. The primary treated group is composed of lab claims associated with Anthem's small-group and large- group plans from 2008-2015. Lab services are a particularly attractive treated procedure group in this context because prices are easy to measure and exhibit considerable variation across groups and time.[10] Moreover, the high volume of labs performed yields a very large sample to analyze.

As noted above, the Site of Service preferred network became standard in all of Anthem's small-group plans in late 2010. While the network was not included in all large-group plans at the outset, the design became increasingly popular each year, reaching around 70% of large-group members by the end of 2014 (Tu and Gourevitch, 2014).

In most analyses, I treat the small group and large groups separately to account for the difference in treatment penetration. I use the entire large-group market, as opposed to isolating employer groups that have adopted the Site of Service program for several reasons. First, it is not always possible to identify whether or not a particular group has the preferred network feature.[11] Second, penetration of the preferred network in the large-group market increases over the sample period, which may introduce sample-selection concerns. Third, and most importantly, the supply-side effects of the program impact all plans within a contract group, even those group-plans without the Site of Service feature in a given year. Given the significant overall penetration in the large group, it is sensible to treat Site of Service as an important feature of the broad-based payer-provider negotiation process, which has downstream impacts on individual group plans. So, broadly, estimated effects for the large-group market reflect both an intent-to-treat component and a direct supply-side component.

I construct two control arms to account for both insurer-level dynamics related to procedure prices as well as and within-lab dynamics affecting prices. I first construct a control group of x-ray procedures. X-rays constitute an attractive comparison group in this context for several reasons. Most importantly, x-rays and other imaging services were not included in the initial rollout of the Site of Service program, and instead were covered by traditional cost sharing over the sample period.[12] In addition, like the sample of labs, the sample of x-rays is large, it includes many distinct procedure codes, and exact prices are easy to infer from claims. Moreover, labs and x-rays are both fundamentally

[10] It is more challenging to measure time-series variation in surgery prices, for instance, which typically have multiple charges for different procedure codes and providers.

[11] In many cases, the observed distribution of providers and out-of-pocket prices does identify the likely plan status for a group. Ackley (2022) focuses on those particular groups.

[12] Interestingly, as Figure B1 shows, imaging services were later added to the Site of Service design, likely in response to the success of the program.

diagnostic procedures with similar prices. In the absence of the Site of Service network, it is likely that cost sharing would have been similar between the two groups of procedures.

The x-ray control group helps account for insurer-level dynamics affecting patient choices and negotiated prices. I construct a second comparison group of lab and x-ray claims from Cigna plans to account for unobserved market dynamics that impact prices across the two procedure groups. For example, if a new technology reduced lab prices in the market, the appropriate time-series comparison to make would be within-lab and across Anthem-Cigna, rather than within-Anthem and across lab-x-ray. Cigna is the second-largest insurer in the state over this time period, and the characteristics of its enrollees are broadly similar to those of Anthem members.

Table 1 presents summary statistics covering all procedures included in the main sample for each of the three main insured groups. Panel A reports summary statistics for all lab procedures and Panel B reports these statistics for all x-ray procedures. Across the three insured groups, demographic characteristics and annual utilization measures are broadly similar. For both labs and x-rays, the two Anthem groups exhibit lower spending per procedure. Across the two procedure groups, cost sharing and distance traveled are similar within insured groups.

Table A1 presents additional summary statistics on patient-level demographic information for each insurer. Demographic characteristics such as age, sex, and Charlson score are quite similar across the insurers, as are zip-code level measures such as income and education. The geographic distribution of enrollees is also similar, with both insurers having the greatest concentration of enrollees in the same five hospital service areas (HSAs).

2 Effect on Payments per Procedure

In this section, I study the effect of Anthem's preferred provider incentives on the price paid per procedure. As described in Section 2, the average payment per procedure reflects the equilibrium interaction between patient demand-side behavior and negotiations between insurers and providers. The goal here is to summarize the combined effect of the preferred network coming from these two channels. I first present a descriptive analysis of temporal changes in per-procedure payments for each group. I then present a series of event-study-type models.

2.1 Descriptive Analysis

To begin, I compute the weighted average lab price paid by each of the three insured groups from 2009-2015, where weights are computed as the share of total spending. The top panel of Figure 1 depicts the time series. From 2009-2010, the three series show a similar upward trajectory. From 2011 onward, the average price paid declines precipitously for both Anthem groups, but continues a general upward trend for Cigna group claims. As a comparison group, the bottom panel of Figure 1 presents the analogous time series for x-ray procedures. Here, average payments per procedure follow a smooth upward trend over the entire period for each of the three insured groups. As Section 1.3 details, for Anthem small-group plans, the preferred provider design became standard for lab tests at the end of 2010, while x-rays continued to be covered by traditional deductible and coinsurance parameters. For Anthem large-group plans, the rollout of this new tiered design was more gradual, but reached around 70% by 2015. Consistent with this greater treatment intensity for small-group members, the observed decline in payments per procedure is larger for this group than for the large-group plans.

The temporal patterns depicted in Figure 1 suggest that Anthem's preferred provider design led to a reduction in the price paid per lab. To summarize some of the changes for particular procedures, Table A2 reports the total change in the mean price paid between 2010 and 2015 for the five most common tests across both anthem groups. The percent change in price ranges from -6% for a complete blood count test to -28% for a hemoglobin test. I also compute the change in the weighted average price across all procedures, as in Figure 1, which comes out to -17%. The last column of Table A2 reports a simple measure of the total savings attributable to the change. This measure is based on the difference between total spending in 2015 and counterfactual spending in 2015 under 2010 prices.[13] Overall, the price changes from 2010 to 2015 accounted for around $6.3 million in savings in 2015. Importantly, given that counterfactual prices would likely have risen in the absence of the Site of Service program, these simple savings estimates are more of a lower-bound on the total program savings rather than a precise treatment effect estimate. I next use event-study-based methods to obtain a more accurate measurement of the savings.

2.2 Event-Study Analysis

The descriptive analysis in the previous section suggests that the Site of Service tiered design generated considerable savings for Anthem on a per-procedure basis. Here I present a formal analysis of the program's effects using difference-in-difference(s) and event-study-type methods. As detailed in Section 1.3 and illustrated in Figure 1, preferred provider incentives vary across time, carrier, and procedure groups. I consider both two-dimensional and three-dimensional setups to explore each of the various component sources of variation.

I first estimate a series of event-study models of the following form:

$$log(p_{ijkt}) = \sum_{t \neq t_{ref}} \theta_\tau m(\text{lab}_k, \text{anth}_j) \times \{t = \tau\} + h(\text{lab}_k, \text{anth}_j, t; \Gamma) + \beta X_{ijt} + \lambda_t + \lambda_k + \epsilon_{ijkt} \tag{1}$$

[13]Specifically, the simple savings measure is: $\text{save}^k_{t_0,t_1} = p_{kt_1} \times q_{kt_1} - p_{kt_0} \times q_{kt_1}$, where p_{kt} is the average price paid for service k in year t, and q_{kt} is the total quantity of k rendered in t.

Here, $log(p_{ijkt})$ denotes the price paid for procedure k, obtained by individual i, who is enrolled in plan j, at time t. The function $m(\text{lab}_k, \text{anth}_j)$ identifies the treatment group of interest for a particular specification. In the first specification, $m()$ is simply an indicator for lab test and the model is estimated on Anthem claims from 2008-2015. In the second specification, $m()$ is an indicator for Anthem, and the model is estimated on Anthem and Cigna lab claims from 2009-2015.[14] The third specification utilizes both control arms in a triple-differences setup. In this case, $m()$ is the interaction of the Anthem indicator and the lab test indicator. In each specification, $h()$ includes all of the natural one-dimensional indicators and two-way interactions of $\text{lab}_k, \text{anth}_j$, and t that are not captured by $m() \times \{t = \tau\}$. Also in these models, X_{ijt} includes patient characteristics such as age, sex, and Charlson score, as well as calendar-month fixed effects and HSA fixed effects. I also include procedure fixed effects, λ_k, and year fixed effects, λ_t.

The parameters of interest in this model are $\{\theta_\tau\}_{\tau \neq t_{ref}}$, the coefficients on the interaction of year and treated group. The estimated sequence of $\hat{\theta}_\tau$s captures the dynamics of treatment-control differences over the sample period. In the two-dimensional models, identification relies on a standard parallel trends assumption– between Anthem and Cigna lab claims in the first case and between lab and x-ray claims in the second case. Contemporaneous shocks that affect Anthem, relative to Cigna, in the post period, such as an exogenous increase in bargaining power would challenge the first model. Contemporaneous shocks that affect lab tests, relative to x-rays, in the post period, such as a technology change would challenge the second model. The three-dimensional specification accommodates both insurer-specific and lab-specific shifts to relevant unobservables in the post period.

I also consider standard difference-in-differences and triple-differences versions of the event study which replaces event-time interactions with a post-period indicator post_t:

$$log(p_{ijkt}) = \theta m(\text{lab}_k, \text{anth}_j) \times \text{post}_t + h(\text{lab}_k, \text{anth}_j, \text{post}_t; \Gamma) + \beta X_{ijt} + \lambda_t + \lambda_k + \epsilon_{ijkt} \tag{2}$$

In this case, post_t is equal to one for 2011 and later. Figure 2 depicts the event study estimates associated with equation (1) for each of the three baseline specifications. The top panel plots the coefficient estimates for the Anthem small group and the bottom panel plots the results for the large group. Table A3 presents the precise numerical results. The estimates are broadly similar across the three specifications, and are consistent with the descriptive trends in Figure 1. Both the small group and large group show a sizeable reduction in per-lab prices in the post period, with magnitudes that are increasing over time. For the small group, the triple-differences estimate implies roughly a 26% reduction in payments per lab by 2012 and a 36% decrease by 2015. For the large group, the

[14] Cigna claims are incomplete before 2009, so I use 2009 as the first period in all models that include Cigna.

analogous implied changes are 14% in 2012 and 32% in 2015. The implied effects are slightly larger in the other specifications.

There are several factors that likely account for the dynamic effects shown in Figure 2. On the demand side, there may be a learning or information-dissemination element where it takes time for patients to become aware of and internalize the cost-sharing incentives in the Site of Service design. Section 4 presents evidence pertaining to this possibility.

Table 2 presents the results from the standard difference-in-differences and triple-differences models. For the small group, the coefficient estimate in the triple-differences specification is -0.35, which implies a reduction in the price paid per lab of about 30% in the post period. For the large group, the coefficient estimate translates to a 23% decline. For both groups, estimates are relatively similar across specifications, and are consistent with the event-study estimates.

The estimates presented above indicate that the Site of Service network yielded considerable savings over this period. The triple-differences event-study estimates imply a total savings of about \$14 million in 2015 across both groups and all procedures. The estimated total savings from 2011-2015 amounts to about \$50 million, or about \$12 per test.[15] Unsurprisingly, these savings estimates are much larger than the simple-comparison estimates in Table A2.

2.3 Heterogeneity and Robustness

Here, I consider a number of alternative specifications of the models applied above to accommodate heterogeneity and other possible dynamics on unobservables.

To explore heterogeneous effects across demographics, I adopt the difference-in-differences framework (2) and add interactions of the primary treatment indicator with three demographic variables of interest: age, CCI, and sex. Here I use the specification which includes only Anthem claims so that the treatment indicator is post $\times$ lab.

Table A6 reports the estimates associated with this interacted specification. For both insured groups, the treatment interaction with CCI is positive, indicating that treatment effects are smaller for less-healthy patients. This may reflect a stronger preference for hospital-based facilities among less-healthy individuals, or perhaps that these individuals are more likely to be in a hospital when labs are indicated. Conversely, these estimates imply that older patients and female patients demonstrate relatively larger responses to the tiered design.

I next consider alternative specifications of the baseline model for the price paid per test. Table A5 shows the results for four different versions of the difference-in-differences model based on Anthem lab and x-ray claims. The first specification excludes controls, the second specification uses zip-code-level fixed effects, the third specification includes

[15] To calculate implied savings in year t, I use the coefficient estimate $\hat{\theta}_t$ to compute the implied counterfactual average price $\bar{p}_t^{cf} = \bar{p}_t / \exp(\hat{\theta}_t)$. Then total savings is $q_t(\bar{p}_t^{cf} - \bar{p}_t)$.

patient-level fixed effects, and the fourth specification is estimated on the sample of the 20 most common lab and x-ray procedures. For both the small group and large group, the estimates are broadly similar across all of the alternative specifications and the baseline specification. Figure A2 shows the event-study estimates associated with each of the general event-study equations and each of the four alternative specifications for controls. As in the difference-in-differences case, these estimates are relatively stable across specifications.

Additionally, following Finkelstein (2007), I estimate a generalized linear model (GLM) version of (2) with a log link function to explicitly estimate $log(E(y|x))$, as opposed to $E(log((y|x))$. As Table A4 presents the results from this specification. The implied effects from the GLM model are broadly similar to the baseline estimates in 2.

3 Steering Effects

The results in the prior section encompass the total impact of both steering and changes in negotiated prices. In this section, I look specifically at the direct steering effects of the Anthem's tiered cost-sharing design. As Section 2 details, the extent to which the tiered incentives actually steer patients toward preferred providers can strongly impact current spending and is a critical upstream determinant of negotiated prices. In prior work, Ackley (2022) presents evidence on the steering effects of the Site of Service program for endoscopic procedures. The key finding in that work is that the tiered cost-sharing schedule has a non-negligible impact on out-of-pocket price sensitivity relative to a traditional deductible-coinsurance-based design. While Ackley (2022) estimates marginal out-of-pocket price responsiveness directly in a discrete choice model, in this paper, I focus on a simpler choice model where the outcome is an indicator for independent lab use.

Independent labs are the basis of the preferred provider tier under Anthem's Site of Service lab network, so the most direct measure of steering in this context is the share of labs performed at these providers. Several hospital-based providers gain preferred status at different times after the program rollout, but I omit these providers from the independent group here to avoid a mechanical increase in preferred provider share. The two main independent lab providers are Quest Diagnostics and Labcorp, although there are several other smaller independent labs in the state which carry preferred status. In 2014, for example, of all claims at preferred providers, over 50% were associated with Quest and Labcorp.

Figure 3 plots the procedure-weighted share of lab tests performed at independent labs, for each insured group over time.[16] From 2009 to 2010 the independent share dropped slightly for each group. From 2010 to 2015, the independent share rose considerably for

[16] I also consider an unweighted version of these calculations and the results are materially similar.

the Anthem small group, roughly doubling from 18% to 36%. For the Anthem large group, the independent share increased from about 13% to around 24% over this period. By comparison, in the Cigna population, the independent share grew from 18% in 2010 to 20% in 2015.

To further examine these dynamics I adapt the prior difference-in-differences strategy to the independent lab outcome:

$$y_{ijkt} = \alpha \text{anth}_j + \gamma \text{post}_t + \theta \text{anth}_j \times \text{post}_t + \beta X_{ijt} + \lambda_t + \lambda_k + \lambda_i + \epsilon_{ijkt} \tag{3}$$

Here, y_{ijkt} is an indicator that is equal to one if the provider is an independent lab. I consider two specifications to account for detailed individual-level characteristics that likely mediate provider choice. In the first specification, I include demographic information such as age, sex, and Charlson score, as well as fixed effects for HSA. In the second specification, I include individual-level fixed effects, exploiting the fact that many patients have multiple lab tests over time.

Table 3 reports the results for both specifications of (3) and each Anthem group. In all cases, the coefficient estimates are similar across specifications and indicate a positive steering effect of the preferred provider incentives, consistent with Figure 3. In particular, the coefficient estimate associated with the small group and individual fixed-effects specification is 0.124, implying around a 12 percentage point increase in the probability of independent lab use. For the Anthem large group, the coefficient estimate is 0.063. Table A7 shows the analogous results for a logit GLM specification of (3).

These results strongly suggest that the preferred network design is effective in steering patients toward incentivized providers. This mechanism explains some of the total procedure-level savings documented in Section 3 and provides a basis for changes in negotiated prices, which I explore in the next section.

4 Changes in Negotiated Prices

In this section, I examine the relationship between the Site of Service design and negotiated prices. Understanding the extent to which tiered-type designs can put downward pressure on negotiated prices is critical for insurers and policymakers aiming to reduce spending. This has become increasingly important in the face of increasing vertical and horizontal integration of providers, which has been shown to raise prices (Baker et al., 2014; Lin et el., 2021; Fulton, 2017; Cooper et al., 2019). However, existing evidence on plan design and negotiated prices is limited.

Theoretically, there are several channels through which the tiered design may impact prices. First, by sharply raising the relative out-of-pocket price of non-preferred providers, the design can raise the effective price elasticity of this group. The extent to which the non-preferred elasticity rises depends on the degree of substitution toward preferred providers.

The results in Section 3 indicate that this type of substitution is achievable. The second channel, which is related to the first, is that the tiered schedule may affect underlying patient price sensitivity directly. Prior work by Prager (2020) and Ackley (2022) supports this hypothesis. Specifically, these studies estimate that price-responsiveness is essentially zero for deductible and coinsurance-based plans, but is significantly different from zero under a tiered design. This finding implies that the price elasticity of demand will rise for non-preferred providers in the tiered setting.

A third channel through which the preferred network may affect prices is by expanding the set of network levels that are bargained over. That is, insurers and providers bargain over not just inclusion in the general network, but also inclusion in the set of preferred providers. Providers have an incentive to bargain for inclusion in the preferred tier if expected profits are higher under that agreement than under the non-preferred agreement.

4.1 Empirical Evidence

To examine trends in negotiated prices, I first estimate a regression-based price index for each carrier and procedure group using the model:

$$log(p_{ijkt}) = \delta_{jt} + \beta X_{jt} + \lambda_{kr} + \epsilon_{ijkt} \tag{4}$$

Here, λ_{kr} denotes procedure-by-provider fixed effects and δ_{jt} denotes payer-by-year fixed effects. The estimated payer-year fixed effects $\hat{\delta}_{jt}$ capture the evolution of negotiated prices for each payer over the sample period. Given the significant overlap in procedure-provider negotiated prices across Anthem plans, I combine both the small-group and large-group plans into a single Anthem group. While this negotiated-price overlap is substantial, some across-plan variation remains. To account for this, I include plan-type fixed effects (i.e., HMO, PPO, etc.) and contract-type fixed effects (i.e., ASO or fully insured). In subsequent analyses, I focus on specific contracts, within-which prices are homogeneous for a given procedure-provider-time cell.

To summarize dynamic differences in negotiated price dynamics more directly, I also collapse the price index model above into the difference-in-differences framework from Section 2:

$$log(p_{ijkt}) = \theta m(\text{lab}_k, \text{anth}_j) \times \text{post}_t + h(\text{lab}_k, \text{post}_t, \text{anth}_j; \Gamma) + \beta X_{jt} + \lambda_t + \lambda_{kr} + \epsilon_{ijkt} \tag{5}$$

The most important difference between these negotiated price models and the previous models (1) and (2) is the addition of procedure-by-provider fixed effects. The objective of including these fixed effects is to isolate variation in negotiated prices across insurers and time. This variation is distinct from changes in the distribution of providers used, conditional on prices.[17] I estimate each of the three versions of this specification on the sample of providers that perform both lab tests and x-ray procedures. The triple-differences specification is particularly attractive for this sample of providers because it allows for dynamic changes in bargaining power and other relevant unobservables at the insurer level and procedure-group level. For example, an increase in Anthem's bargaining power in the post period is captured by the second-order term $\text{post}_t \times \text{anth}_j$. Then, as in the analogous model of Section 3, identification here relies on the absence of anthem-lab specific shocks in the post period that were not related to the Site of Service network.

Figure 4 plots the index coefficient estimates for each insurer and procedure group, where the Anthem-2010 coefficient serves as the reference point. For Anthem, lab prices fall considerably in the post period. The coefficient estimates imply that lab prices declined by around 13% from 2010 to 2015 for Anthem. In contrast, lab prices associated with Cigna and x-ray prices for both insurers are increasing over the period.

Table 4 reports the difference-in-differences estimates of the negotiated price changes. The coefficient estimates across the three specifications range from -0.127 to -0.161, implying a reduction in negotiated prices of 12%-15%. These estimates are broadly consistent with the price index estimates above.

Figure A4 depicts the analogous estimates from the event study versions of (5). Both these estimates and the index estimates in Figure 4 show that negotiated price changes first become evident in 2012, and then increase in magnitude after that. There are several explanations that are consistent with this temporal pattern. First, the share of Anthem large-group enrollees in the Site of Service program increased over the post period. This likely boosted the incentives of providers to negotiate preferred status. Similarly, patients may take time to adjust to the tiered incentives, leading to a demand-side effect that is increasing over time. The results in Figure 3 are consistent with this possibility. Finally, the complexity of price negotiations may produce a lagged response to changes in market conditions.

Overall, these estimates above suggest that changes in negotiated prices are a meaningful component of the overall decline in per-lab payments documented in Section 3. Roughly speaking, comparing the negotiated price results above with the baseline results in Section 3 suggests that around half of the total decline in per-lab payments may be due to negotiated price changes. Indeed, the formal decomposition results in Section 6 support this general allocation. This finding lends support to some of the theoretical mechanisms connecting preferred networks, steering, and prices discussed above.

[17] Of course, changes in negotiated priced may be related to changes in demand in an equilibrium sense. The goal here is to document the magnitude of price changes over the period.

4.2 Heterogeneity Across Providers

To better understand the driving forces behind the falling lab prices paid by Anthem, I break down price patterns separately for three distinct groups of lab providers based on their relation to the Site of Service network: (1) always preferred, (2) newly preferred, and (3) non-preferred. The always-preferred groups is made up of independent, non-hospital-based labs such as Quest Diagnostics and Labcorp. These providers formed the original group of preferred providers under the Site of Service program. The newly-preferred group consists of a set of hospitals that switch from being non-preferred to preferred after explicitly renegotiating lower prices with Anthem. Non-preferred providers are not preferred during the sample period.

The regression-based index (4) is useful for summarizing trends at the broad, payer level. For these these narrower provider groups, I compute a more explicit Laspeyres-type price index using exact prices and quantities. Specifically, the index $\tilde{P}_t$, takes the following form:

$$\tilde{P}_t = 100 \times \frac{\sum_{k \in K} \sum_{r \in R} \bar{q}_{kr\bar{t}} \times p_{krt}}{\sum_{k \in K} \sum_{r \in R} \bar{q}_{kr\bar{t}} \times \bar{p}_{krt}} \tag{6}$$

Here, $\bar{q}_{kr\bar{t}}$ denotes the total quantity of procedure k at provider r in the base period, $\bar{t}$, while q_{krt} reflects this quantity in year t. Also, p_{krt} denotes the price of k at provider r in year t. In this index, the basket is defined as the total quantity associated with each procedure-provider pair in the reference period. Therefore, the index tracks changes in negotiated prices over time, holding the distribution of providers and procedures fixed.

Importantly, as noted above, negotiated prices can vary across contracts for a given procedure-provider-insurer cell. To capture changes in negotiated prices that are most closely related to the Site of Service preferred network, I focus on the Anthem contract associated with the majority of fully-insured small-group and large-group claims. To construct this contract group, I follow a similar bunching-type procedure as in Cooper et al. (2019) and Craig et al. (2021).[18]

Figure 5 plots the Laspeyres-type indices for each provider group, taking 2010 as the reference year for all. The most salient result in this figure is the dramatic decline exhibited by newly-preferred providers. Relative to 2010, the price index for these providers falls by nearly 50% by 2014-2015. This group consists of five hospitals that renegotiated lower rates with Anthem during the middle of the sample period in order to obtain preferred status in the Site of Service network (Tu and Gourevitch, 2014). Figure A5 plots the price indices for several of these providers. As Figure A5 shows, different providers renegotiated

[18] I first construct a plan variable, which is defined as the concatenation of payer, plan type, group size, and ASO/fully-insured status. I then define a contract as a group of plans for which procedure-provider level prices match to the $0.01-level. For the main contract I focus on here, the match rate between the plan-level price and the contract-imputed price is about 90% across procedure-provider cells.

in different years. This contributes to the increasing price effects observed over the post period.

The always-preferred group also shows a considerable decline in prices in the post period. As Figure A5 shows, this is driven by a drop in prices associated with Quest from 2011 to 2012. The price decline among always-preferred group is somewhat counterintuitive, given that patients in the Site of Service program face no cost sharing at these providers. One possible explanation for this is that value of preferred status increased over the post period, as more patients became enrolled and responded to the incentives. This would, in turn, raise the bargaining leverage of Anthem in negotiations over price and preferred status.

In contrast to the preferred groups, the price index for never-preferred providers increases relatively smoothly over the entire sample period. To summarize the difference between the never-preferred and newly-preferred groups more precisely, Table A8 reports differences-in-differences estimates from (5) separately for each group.[19] For the newly-preferred group, these estimates imply a 29%-36% reduction in prices in the post period. The analogous estimates for the never preferred group are in the range of 6%-7%.

These results indicate that negotiated-price dynamics are an important lever through which preferred network designs may affect spending per procedure. Indeed, the estimates in this section show that changes in negotiated prices are an important component of the overall decline in per-lab spending associated with Anthem plans. In the next section, I quantify the relative contribution of these negotiated price changes to overall changes in spending per procedure using a formal decomposition.

5 Decomposition of Price and Steering Effects

The results presented thus far indicate that the Site of Service tiered design led to a considerable reduction in spending per lab test. Moreover, the overall change in per-lab payments is driven by both a steering effect and a negotiated price effect. In this section, I perform a precise decomposition of aggregate price changes into a steering component and a negotiated price component.

First, let $D_{krt} = \frac{1}{\sum_{r \in R_k} q_{krt}} \times q_{krt}$ denote the market share of provider r for service k at time t, which I also call demand. Then, the average price paid for this procedure in t is given by:

$$\bar{P}_{kt} = \sum_{r \in R_k} p_{krt} \times D_{krt} \tag{7}$$

Much of the analysis to this point has focused on the time series of $\bar{P}_{kt}$ across services and insured groups. For a particular service and insured group, the change in average price

[19] I omit the always-preferred group from this analysis because these providers do not perform x-rays.

paid between years t_0 and t_1 can be broken out into:

$$\bar{P}_{kt_1} - \bar{P}_{kt_0} = \underbrace{(\sum_{r \in R_k} p_{krt_0} \times D_{krt_1} - \sum_{r \in R_k} p_{krt_0} \times D_{krt_0})}_{\equiv SE_k(t_1,t_0) \text{ (steering effect)}} + \underbrace{(\sum_{r \in R_k} p_{krt_1} \times D_{krt_1} - \sum_{r \in R_k} p_{krt_0} \times D_{krt_1})}_{\equiv PE_k(t_1,t_0) \text{ (price effect)}} \tag{8}$$

The first term is the steering effect, $SE_k(t_1, t_0)$, which holds prices constant at the t_0 level and allows demand to change between t_1 and t_0. The second term is the price effect, $PE_k(t_1, t_0)$, which holds demand constant at the t_1 level and allows prices to change between t_1 and t_0. It is straightforward to compute an aggregate decomposition across multiple procedures by taking a weighted average:

$$\sum_k w_k \bar{P}_{kt_1} - \sum_k w_k \bar{P}_{kt_0} = \sum_k w_k (SE_k(t_1, t_0) + PE_k(t_1, t_0)) \tag{9}$$

for sensible weights w_k. It is worth noting that the decomposition defined above should be viewed more as an accounting exercise than as a breakdown of causal effects. In equilibrium, prices and demand are jointly determined via an intricate bargaining process. To conduct a more-detailed causal breakdown of the preferred network design, a general equilibrium model is necessary.[20]

Table 5 reports the decomposition results for the 10 most common lab services and the aggregate measure across all procedures. Here I use 2010 as the base year and 2015 as the comparison year. The first two columns report the 2010 and 2015 average prices, respectively, and column 3 reports the percentage change. Column 4 reports the steering component of the overall percentage change and column 5 shows the negotiated price component of the total percentage change.

The aggregate change in the mean price paid is 14.6%. Of this, the steering component accounts for about 6.7% and the price component accounts for about 7.9%. Across the 10 individual procedures, the aggregate change ranges from a little under 6% to around 25%. In most cases, the steering component and price component are relatively similar in magnitude.

To explore heterogeneity across the two main Anthem groups, I perform the decomposition separately for small-group and large-group claims. Figure 6 depicts the price and steering decomposition separately for these groups for the 5 most common tests and the aggregate measure. Consistent with prior results, average payments per lab fall by more across the board for the Anthem small group than for the large group. For the aggregate measure, the average price paid falls by almost 20% between 2010 and 2015 in the small group. For the large group, the analogous decline is about 12%. As in the combined

[20]Both Prager (2016) and Ackley (2020) develop extensions of the Gowrisankaran et al. (2015) framework which incorporate tiered cost-sharing into the equilibrium model. These models permit the researcher to simulate counterfactual prices under alternative plan designs.

results of Table 5, there is some heterogeneity across procedures, although the relative heterogeneity is similar across the small and large groups.

The results in Figure 6 suggest that the steering effect is somewhat more dominant in the small group than it is in the large group. For the aggregate measure in the small group, the steering effect accounts for slightly more than half of the overall change. In the large group, steering accounts for a little under half of the overall decrease in the aggregate measure. Overall, these results are consistent with the imperfect penetration of the Site of Service network in the large group. The fact that not all large-group enrollees have tiered incentives over the entire post period implies that steering effects would be smaller, while these groups may nonetheless realize spillover effects from negotiated price changes.

6 Effects on Cost Sharing and Total Utilization

6.1 Out-of-Pocket Costs

Here I examine the effect of the program on patient cost sharing, the direction of which is ex-ante ambiguous. While deductibles are generally rising over the sample period, the Site of Service plan imposes zero out-of-pocket costs on patients when they use preferred providers. Consequently, the net effect of the plan design on out-of-pocket costs depends on the degree to which patients substitute toward preferred providers. Understanding these cost sharing dynamics is important because out-of-pocket costs are a first-order determinant of consumer welfare in this context.

To probe this question, I use the first version of the difference-in-differences model (2), which uses the longer time period of Anthem lab and x-ray claims. Table 6 reports the difference-in-differences estimates for two different cost-sharing measures and both Anthem groups. The first outcome variable is the log of the total out-of-pocket price plus 1 and the second outcome variable is the total cost share–that is, the ratio of the out-of-pocket price to the total price paid. For the cost share outcome, I consider both a linear specification and a Poisson GLM specification. Figure A6 shows the analogous event-study estimates.

For both the small group and large group, these estimates indicate that cost sharing for labs declined, relative to x-rays, as a result of the tiered design. For the small group, the estimates imply that the tiered design is associated with a lower average cost share of almost 37 percentage points. For the large groups, the analogous estimate is 14 percentage points. Moreover, the log out-of-pocket price results indicate that the lower cost share for labs translates to lower out-of-pocket prices in absolute terms.

6.2 Extensive-Margin Effects

The primary outcome of interest in this paper is the price paid per procedure, which depends on both patient choices and insurer-provider negotiations. Of course, plan design may also impact the total quantity of services demanded, which may indirectly affect negotiated prices. Indeed, the literature on HDHPs indicates that these plans lead to a reduction in utilization on the extensive margin, but not to any changes on the intensive margin (Haviland et al., 2016; Brot Goldberg et al., 2017). However, because patients incur zero cost sharing at preferred lab providers, a rising deductible need not impact the extensive margin as much as a standard HDHP.

Figure 7 depicts the time series for three annual utilization measures for each group: total lab spending, total lab count, and an indicator for having any labs in the year. The total lab count and indicator for any labs reflect the extensive margin of demand entirely, while total spending reflects both the extensive margin and the price paid per test. The top panel shows that mean annual lab spending dropped by about 25% over the sample period for both Anthem groups, but increased for the Cigna group. The second and third panels of Figure 7 show that, for both Anthem groups, the average quantity of labs is mostly flat over the sample period. This implies that the decline in average annual lab spending is due almost entirely to a decrease in the price paid per test. Moreover, the event-study estimates above show that the decrease in spending per lab is not simply an artifact of a change in the composition of labs received, but reflects a genuine reduction in test-specific conditional prices.

7 Conclusion

In this paper, I study a preferred lab network that employs a tiered cost-sharing schedule to incentivize patients to use low-cost providers. I find that these plans are very effective at reducing spending per procedure, and that savings are achieved through both a demand-side-steering channel and a supply-side-price channel. On the demand side, I find that patients meaningfully shift toward preferred independent labs. On the supply side, I find that changes in negotiated prices account for about half of the overall decline in spending per procedure, with the largest changes coming from providers that renegotiate lower rates to obtain preferred status.

Collectively, these results shed new light on design-based mechanisms to reduce spending on health care. These findings suggest that tiered-type incentives offer insurers and policymakers an attractive method to counter increasing provider consolidation, which tends to drive prices up (Baker et al., 2014; Lin et el., 2021; Fulton, 2017; Cooper et al., 2019). In an environment with both increasing vertical and horizontal integration, the capacity to both steer and renegotiate is critical in cost containment efforts.

While this paper focuses on lab services, these results could be applied to other insurance settings and procedure groups. Importantly, there are several key factors that govern the effectiveness of the tiered design in reducing spending in a particular setting. The first is the extent to which steering is possible. This likely depends on the type of service, patient population, and distribution of providers. Diagnostic radiology procedures may be a good candidate, as they are particularly similar to lab tests in terms of being common, relatively homogeneous, and exhibiting price dispersion. Prior work has shown that the tiered/preferred-type design is effective in steering for surgical procedures and even, to some extent, in an inpatient setting (Robinson and Brown, 2013; Whaley et al., 2018; Ackey, 2022; Prager, 2020). The second key factor is the extent to which a particular insurer is able to leverage the steering component to extract lower prices from providers. This, in turn, depends on demand elasticities, insurer bargaining power, and market structure. The results in this paper suggest that savings can be achieved through plan design alone, without necessitating an increase in insurer market share.

More work is needed to assess how tiered-network-type plans like the Site of Service program can be designed optimally. On the consumer side, it is important to accurately measure the welfare differences across different plan designs. On the insurer-provider side, it is critical to understand the impact of this network design on the bargaining process and how this may evolve over time as these networks become more prevalent. Future work may compare different realized designs using observational or experimental-type frameworks. It would also be valuable to compare different possible parameterizations of the tiered-network-type setup using a fitted general equilibrium model. Overall, it is likely that improved plan designs will lead to meaningful welfare gains in this market.

References

Ackley, Calvin A. 2020. "Essays on insurance design and the demand for medical care." PhD diss. Boston University.

Ackley, Calvin A. 2022. "Tiered cost sharing and health care demand." *Journal of Health Economics*, 85: 102663.

Aouad, Marion, Timothy T Brown, and Christopher M Whaley. 2021. "Understanding the distributional impacts of health insurance reform: Evidence from a consumer cost-sharing program." *Health economics*, 30(11): 2780–2793.

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

Brot-Goldberg, Zarek C, Amitabh Chandra, Benjamin R Handel, and Jonathan T Kolstad. 2017. "What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics." *The Quarterly Journal of Economics*, 132(3): 1261–1318.

Brown, Zach Y. 2019. "An empirical model of price transparency and markups in health care." *Manuscript.*

Charlson, Mary E, Peter Pompei, Kathy L Ales, and C Ronald MacKenzie. 1987. "A new method of classifying prognostic comorbidity in longitudinal studies: development and validation." *Journal of Chronic Diseases*, 40(5): 373–383.

Claxton, Gary, Matthew Rae, Michelle Long, Anthony Damico, and Heidi Whitmore. 2018. "Health benefits in 2018: modest growth in premiums, higher worker contributions at firms with more low-wage workers." *Health Affairs*, 37(11): 1892–1900.

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

Frank, Matthew B, John Hsu, Mary Beth Landrum, and Michael E Chernew. 2015. "The impact of a tiered network on hospital choice." *Health Services Research*, 50(5): 1628–1648.

Gorman, Bela, Jennifer Smagula, and Jon Camire. 2013. "New Hampshire Insurance Department 2011 Medical Cost Drivers." Gorman Actuarial.

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

Gruber, Jonathan, and Robin McKnight. 2016. "Controlling health care costs through limited network insurance plans: Evidence from Massachusetts state employees." *American Economic Journal: Economic Policy*, 8(2): 219–50.

Haviland, Amelia M., Matthew D. Eisenberg, Ateev Mehrotra, Peter J. Huckfeldt, and Neeraj Sood. 2016. "Do "Consumer-Directed" health plans bend the cost curve over time?" *Journal of Health Economics*, 46: 33 – 51.

Highland, James. 2012. "Report to the State of New Hampshire Insurance Department 2010 Cost Drivers." Compass Health Analystics.

Lin, Haizhen, Ian M McCarthy, and Michael Richards. 2021. "Hospital pricing following integration with physician practices." *Journal of Health Economics*, 77: 102444.

NHID. 2008. "Supplemental Report of the 2007 Health Insurance Market in New Hampshire." New Hampshire Insurance Department.

NHID, ed. 2013. "Public Hearing Concerning Premium Rates in the Health Insurance Market (meeting transcript)." Concord, New Hampshire:New Hampshire Insurance Department.

Prager, Elena. 2016. "Tiered hospital networks, health care demand, and prices." *Job Market Paper, The Wharton School, University of Pennsylvania.*

Prager, Elena. 2020. "Healthcare Demand under Simple Prices: Evidence from Tiered Hospital Networks." *American Economic Journal: Applied Economics*, 12(4): 196–223.

Robinson, James C, and Kimberly MacPherson. 2012. "Payers test reference pricing and centers of excellence to steer patients to low-price and high-quality providers." *Health Affairs*, 31(9): 2028–2036.

Robinson, James C, and Timothy T Brown. 2013. "Increases in consumer cost sharing redirect patient volumes and reduce hospital prices for orthopedic surgery." *Health Affairs*, 32(8): 1392–1397.

Robinson, James C, Christopher Whaley, and Timothy T Brown. 2016. "Association of reference pricing for diagnostic laboratory testing with changes in patient choices, prices, and total spending for diagnostic tests." *JAMA Internal Medicine*, 176(9): 1353–1359.

Robinson, James C, Timothy T Brown, Christopher Whaley, and Emily Finlayson. 2015. "Association of reference payment for colonoscopy with consumer choices, insurer spending, and procedural complications." *JAMA Internal Medicine*, 175(11): 1783–1789.

Sinaiko, Anna D, and Meredith B Rosenthal. 2014. "The impact of tiered physician networks on patient choices." *Health Services Research*, 49(4): 1348–1363.

Smagula, Jennifer, Don Gorman, Gabriella Lockhart, Linda Kiene, and Bela Gorman. 2016. "New Hampshire Insurance Department Final Report on 2014 Medical Cost Drivers." Gorman Actuarial.

Song, Zirui, Timothy Lillehaugen, and Jacob Wallace. 2021. "Out-of-network laboratory test spending, utilization, and prices in the US." *JAMA*, 325(16): 1674–1676.

Starc, Amanda, and Ashley Swanson. 2021. "Preferred pharmacy networks and drug costs." *American Economic Journal: Economic Policy*, 13(3): 406–46.

Tu, Ha, Rebecca Gourevitch, et al. 2014. "Moving markets: lessons from New Hampshire's health care price transparency experiment." Mathematica Policy Research.

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

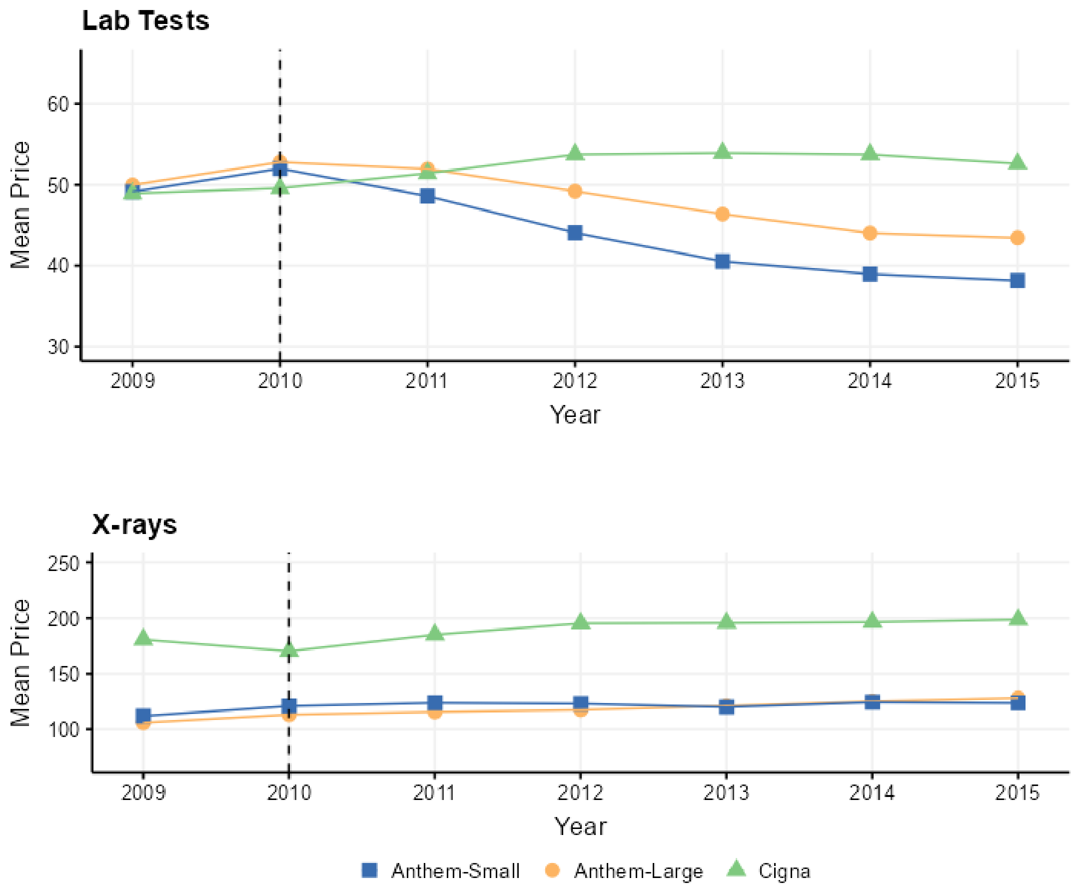
Whaley, Christopher M, Chaoran Guo, and Timothy T Brown. 2017. "The moral hazard effects of consumer responses to targeted cost-sharing." *Journal of Health Economics*, 56: 201–221.

Table 1: Summary Statistics

	Anthem-Small		Anthem-Large		Cigna	
	Mean	Sd	Mean	Sd	Mean	Sd
Panel A: Labs						
Total Payments	44.98	70.19	48.42	74.47	52.73	83.83
Out-of-Pocket	11.58	38.87	5.34	26.01	8.80	25.66
Miles to Provider	18.71	22.68	19.57	23.64	21.38	25.30
Age	43.91	15.55	45.82	15.68	44.89	15.87
Female	0.60	0.49	0.61	0.49	0.61	0.49
Charlson Score	0.58	1.39	0.67	1.44	0.76	1.52
Annual Spend	320.72	503.14	372.23	518.39	385.49	542.44
Annual Count	7.13	9.14	7.69	9.28	7.31	8.62
Count	520084		4746848		2241996	
Panel B: Xrays						
Total Payments	117.03	91.89	116.77	87.78	189.24	168.70
Out-of-Pocket	50.91	76.83	23.99	57.90	45.05	77.05
Miles to Provider	16.83	21.46	16.42	21.77	14.75	16.98
Age	40.43	17.76	43.59	18.00	41.56	18.53
Female	0.51	0.50	0.54	0.50	0.53	0.50
Charlson Score	0.43	1.16	0.55	1.28	0.66	1.39
Annual Spend	195.34	312.08	204.65	313.77	291.26	353.34
Annual Count	1.67	1.37	1.75	1.50	1.54	1.21
Count	54965		525182		179507	

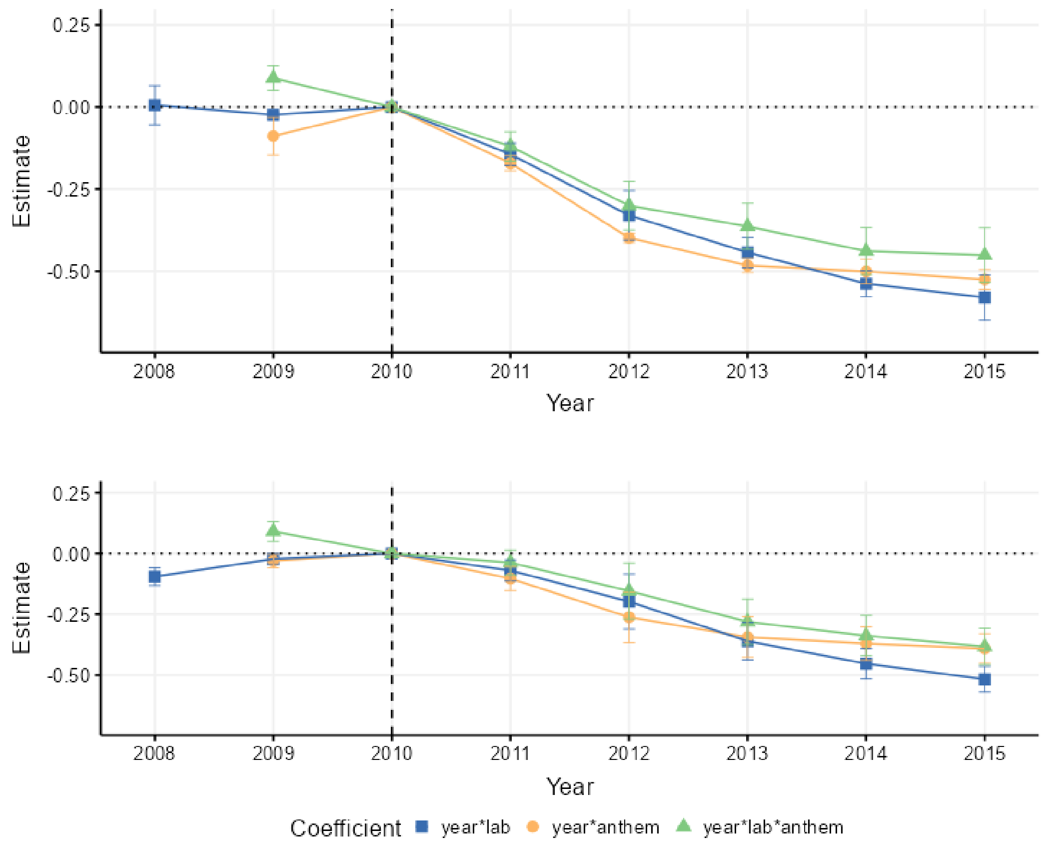
Summary statistics covering all lab and x-ray procedures in the primary analysis sample 2009-2015. The first three rows reflect characteristics of services received while the remaining rows reflect characteristics of the individuals that received the tests. Annual spending denotes the total spending on labs (top panel) or x-rays (bottom panel) in the year that a particular encounter-level procedure is received. Similarly, annual count is the total quantity of procedures received in the year.

Figure 1: Average Price Paid per Procedure by Group



Notes: Top panel shows the weighted average lab price paid for claims associated with each of the three main insured groups described in the text. Weights are computed as each procedure's share of total volume over the sample period. The bottom panel shows the analogous plot for x-rays.

Figure 2: Event Study Estimates: Price Paid Per Procedure



Notes: Top panel depicts the event-study estimates associated with equation (1) in the text for the Anthem small group. The bottom panel shows the analogous estimates for the Anthem large group. Each line reflects the estimates from a particular version of the general event-study equation, where the dependent variable is the log of the total price paid for a particular procedure. In the $year * lab$ case, the sample is Anthem claims and the treated group consists of lab claims in the post period. In the $year * anthem$ case, the sample is entirely lab claims from both Anthem and Cigna, and the treated group consists of Anthem claims in the post period. In the $year * lab * anthem$ case, the sample contains both lab and x-ray procedures from both insurers, and the treated group consists of Anthem lab claims in the post period. Standard errors are clustered at the plan level. Cigna claims begin in 2009, so models with Cigna are estimated on 2009-2015 data, while the Anthem-only model is estimated using 2008-2015 data.

Table 2: Difference-in-Differences Estimates: Price Paid Per Procedure

	Anthem-Small			Anthem-Large		
	(1)	(2)	(3)	(4)	(5)	(6)
Estimate	-0.376***	-0.345***	-0.354***	-0.261***	-0.267***	-0.266***
	(0.049)	(0.012)	(0.031)	(0.045)	(0.041)	(0.043)
Model	DD	DD	DDD	DD	DD	DDD
Coefficient	post*lab	post*anthem	post*lab*anthem	post*lab	post*anthem	post*lab*anthem
R^2	0.53	0.39	0.46	0.54	0.45	0.5
N	733,088	3,412,048	3,631,900	6,639,634	8,228,967	8,862,894
Mean Dep Var	50.31	51.79	59.29	53.78	50.21	56.46

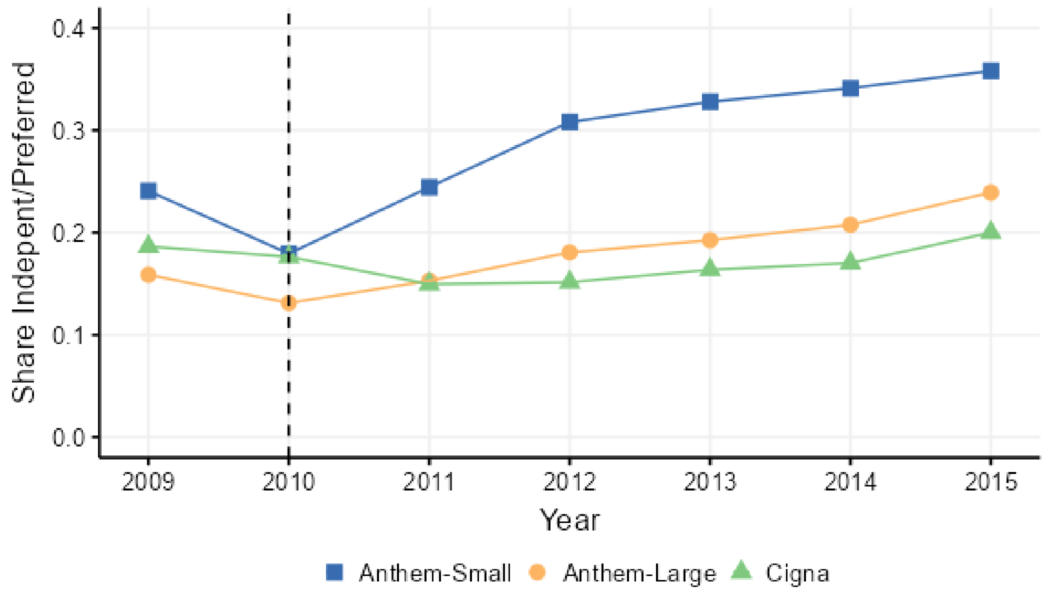
Notes: Table shows difference-in-differences (DD) and triple-differences (DDD) estimates associated with equation (2) in the text. he dependent variable is the log of the total price paid for a particular procedure. In the $year * lab$ case, the sample is Anthem claims and the treated group consists of lab claims in the post period. In the $year * anthem$ case, the sample is entirely lab claims from both Anthem and Cigna, and the treated group consists of Anthem claims in the post period. In the $year * lab * anthem$ case, the sample contains both lab and x-ray procedures from both insurers, and the treated group consists of Anthem lab claims in the post period. Standard errors, clustered at the plan level are shown in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 3: Difference-in-Differences Estimates: Independent Lab Use

Treated Group:	Anthem-Small	Anthem-Small	Anthem-Large	Anthem-Large
Estimate	0.118***	0.124***	0.062*	0.063***
	(0.012)	(0.007)	(0.03)	(0.004)
Individual FE	No	Yes	No	Yes
R^2	0.19	0.63	0.17	0.59
N	3,412,048	1,692,867	8,228,967	4,924,108
Mean Dep Var	0.19	0.18	0.17	0.16

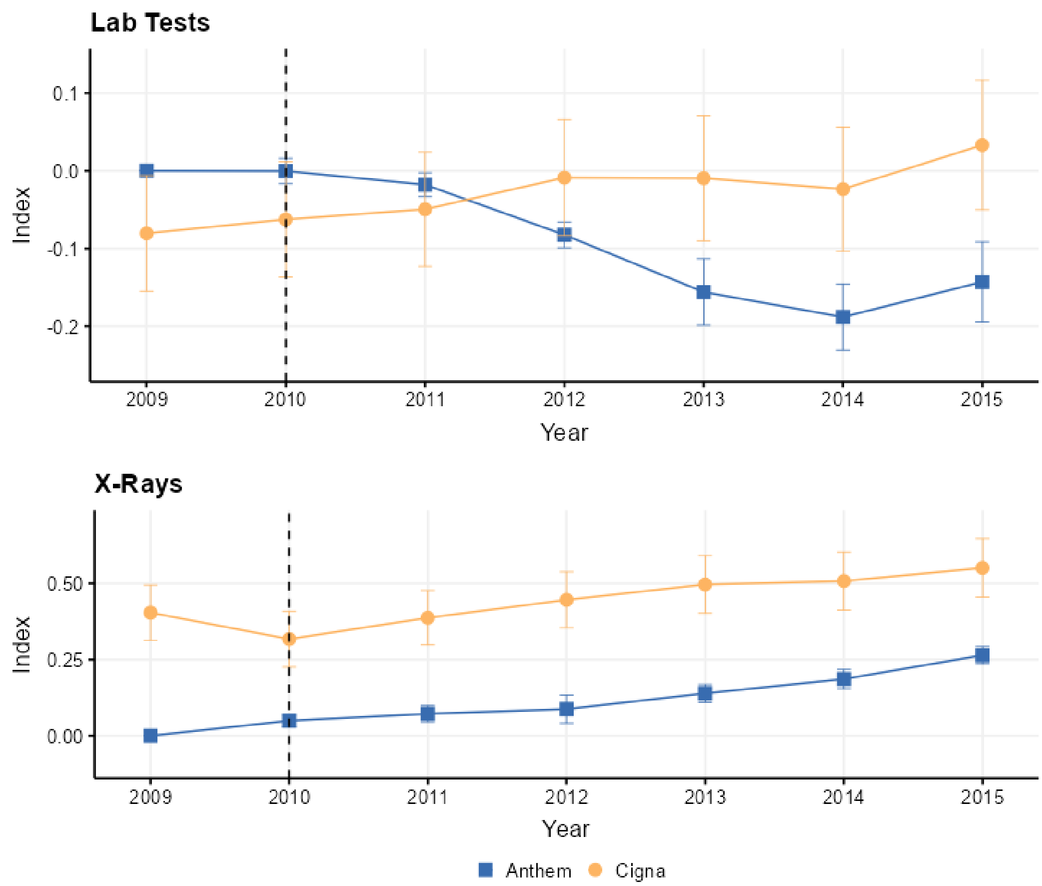
Notes: Table shows difference-in-differences estimates associated with equation (2) in the text. The dependent variable is an indicator that is equal to one if the servicing provider is an independent lab and zero otherwise. The models with individual fixed effects are estimated on the sample of patients with a lab claim in both the pre period and post period.

Figure 3: Independent Lab Share Over Time



Notes: Figure plots the share of lab tests performed at independent labs, preferred under the Site of Service program, for each insured group over time. Shares are computed with procedure-volume weights to account for changes in the mix of procedures.

Figure 4: Regression-Based Price Indices



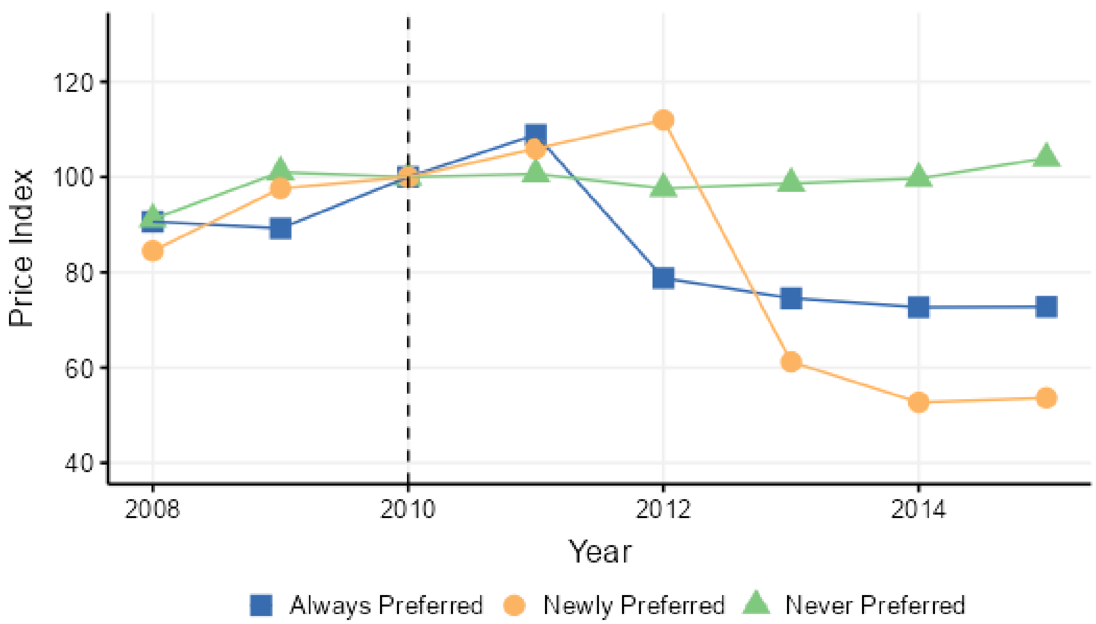
Notes: Top panel shows the regression-based price index for lab procedures associated with equation (4) in the text. Bottom panel shows the analogous estimates for x-rays. Points represent the estimated insurer-year fixed effects for a particular procedure group. Standard errors are clustered at the plan level.

Table 4: Difference-in-Differences Estimates: Negotiated Prices

	(1)	(2)	(3)
Estimate	-0.127***	-0.161***	-0.134***
	(0.019)	(0.018)	(0.026)
Model	DD	DD	DDD
Coefficient	post*lab	post*anthem	post*lab*anthem
R^2	0.72	0.78	0.74
N	6,684,650	5,461,588	7,163,850
Mean Dep Var	59.23	62.31	65.9

Notes: Table shows difference-in-differences (DD) and triple-differences (DDD) estimates associated with equation (5) in the text. The dependent variable is the log of the total procedure price. Sample is restricted to hospital-based providers that perform both lab tests and x-rays. Standard errors, clustered at the plan level are shown in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure 5: Price Index by Preferred Group



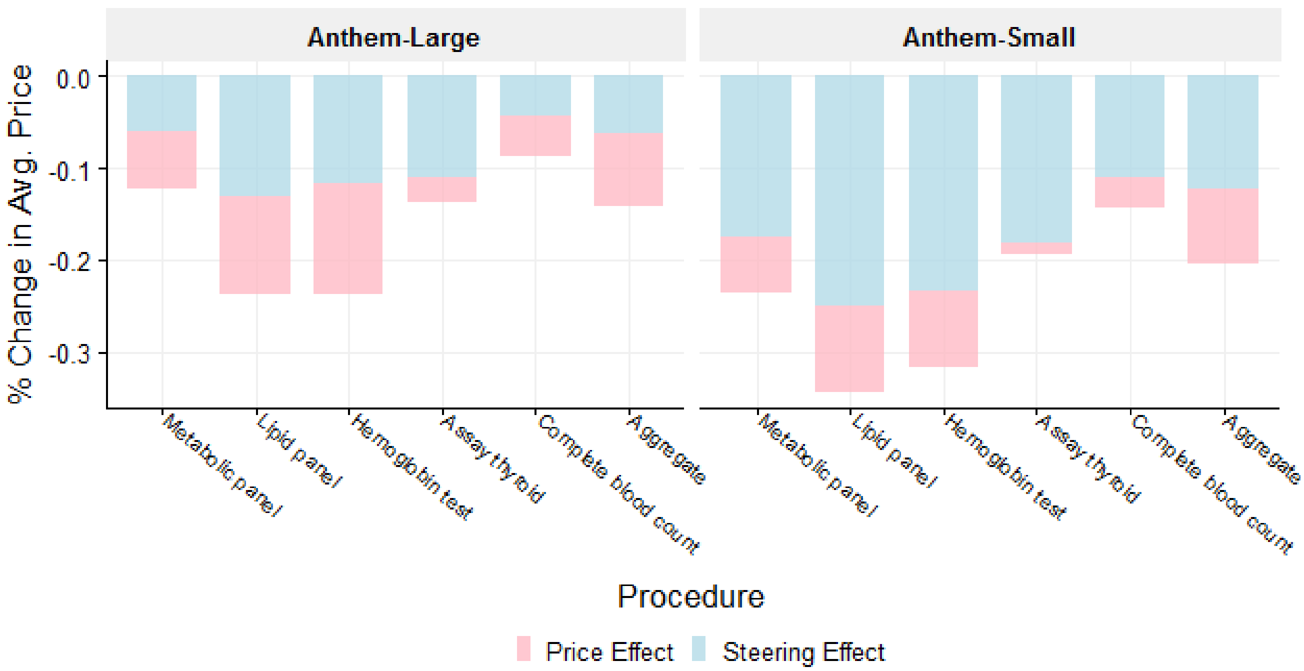
Notes: Figure shows Laspeyres-type price index estimates associated with equation (6) in the text. Always preferred providers include the independent labs that are preferred over the entire program period. Newly preferred providers switch from being non-preferred to preferred after explicitly renegotiating lower prices with Anthem. Never-preferred providers are not preferred during the sample period. 2010 is taken as the reference period for each group.

Table 5: Decomposition of Steering and Price Effects

	$\bar{P}_{2010}$	$\bar{P}_{2015}$	%Δ	%Δ Steer	%Δ Price	N
Aggregate	73.91	63.10	-14.62	-6.73	-7.89	1241872
80053 Comprehen metabolic panel	43.63	37.89	-13.14	-6.88	-6.25	119200
85025 Complete cbc w/auto diff wbc	31.85	28.92	-9.19	-4.82	-4.37	107791
80061 Lipid panel	51.32	38.74	-24.51	-13.86	-10.65	105647
84443 Assay thyroid stim hormone	66.44	56.99	-14.22	-11.47	-2.75	70206
83036 Glycosylated hemoglobin test	36.40	27.54	-24.36	-12.41	-11.94	38938
80048 Metabolic panel total ca	36.61	30.81	-15.83	-6.41	-9.42	35870
82306 Vitamin d 25 hydroxy	105.05	78.29	-25.48	-14.24	-11.23	34151
85610 Prothrombin time	17.02	14.33	-15.81	-7.86	-7.96	29595
87086 Urine culture/colony count	25.88	21.88	-15.45	-4.41	-11.04	27564
81001 Urinalysis auto w/scope	14.32	11.33	-20.88	-11.73	-9.14	25283
85027 Complete cbc automated	25.55	22.90	-10.37	-8.57	-1.80	23380
84153 Assay of psa total	52.52	38.16	-27.35	-15.54	-11.82	21160
88305 Tissue exam by pathologist	190.58	179.96	-5.57	5.34	-10.91	19963
81003 Urinalysis auto w/o scope	6.95	6.25	-10.09	4.04	-14.13	19222

Notes: Table shows price-steering decomposition results associates with equation (8) in the text. The decomposition is shown for each of the 10 most common tests. The first row shows the aggregate decomposition for all lab tests in the sample, which is computed according to (9). The first column shows the mean price paid per test in 2010 and the second column shows this statistic for 2015. The sample includes both Anthem small-group and large-group claims. The steering component reflects amount of the total change that is accounted for by a change in demand. The price component reflects the amount of the total change that is accounted for by changes in negotiated prices.

Figure 6: Decomposition of Steering and Price Effects by Group



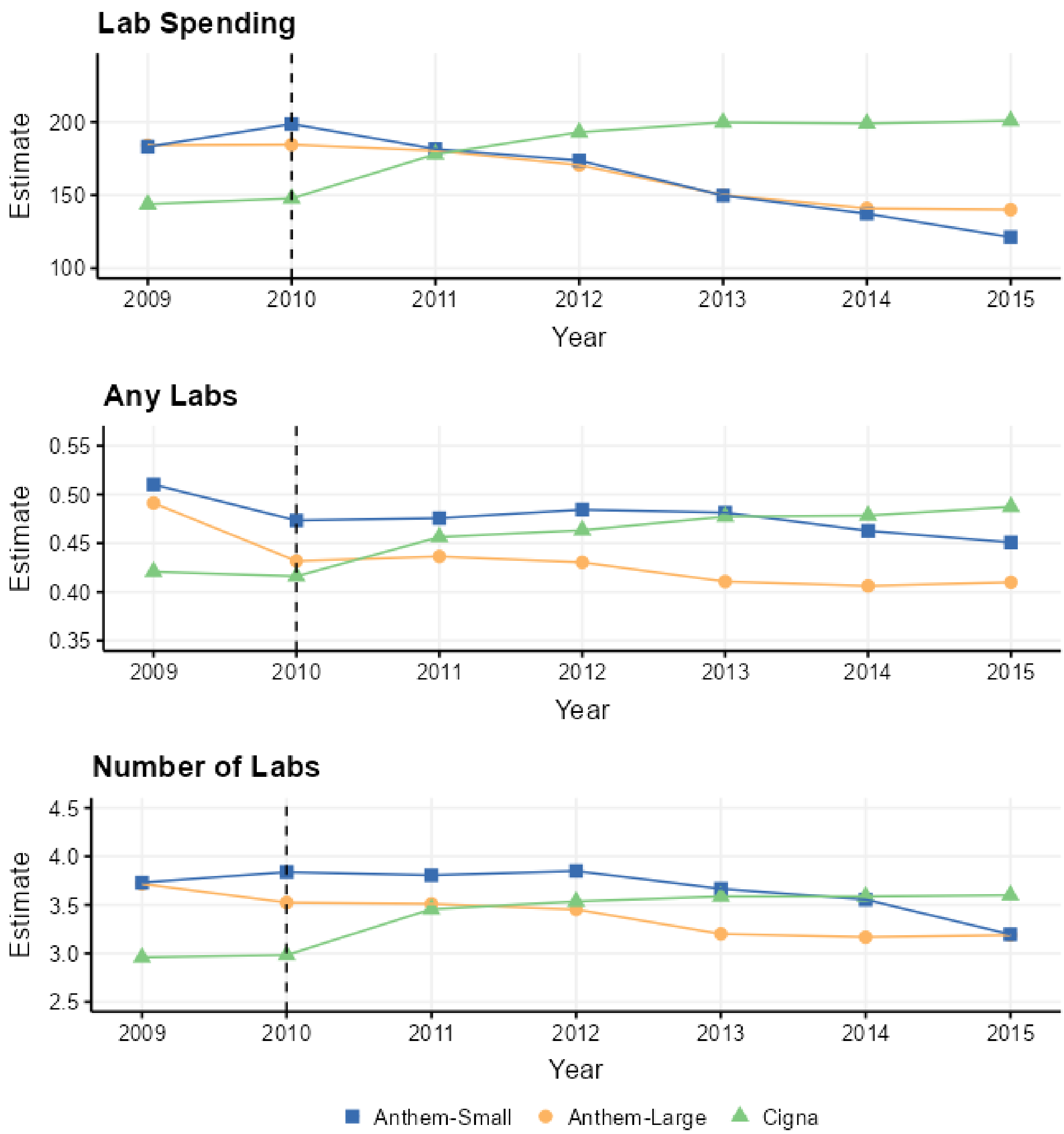
Notes: Graphical depiction of the price-steering decomposition results associates with equation (8) in the text. The total bar height reflects the change in the average price paid for procedure k between 2010 and 2015. The blue shade of the bar represents the portion of the total change that is accounted for by a shift in the distribution of providers used. The pink shade of the par represents the portion of the total change that is accounted for by changes in negotiated prices. The left panel shows the decomposition results for the 5 most common tests and the aggregate measure for the Anthem large group. The right panel shows the analogous results for the small group.

Table 6: Difference-in-Differences Estimates: Cost-Sharing

	Anthem-Small			Anthem-Large		
Dependent Variable:	log(OOP)	cost share	cost share	log(OOP)	cost share	cost share
Estimate	-1.728***	-0.367***	-0.817***	-0.752*	-0.144*	-0.584***
	(0.383)	(0.051)	(0.102)	(0.348)	(0.072)	(0.111)
Specification	OLS	OLS	Poisson	OLS	OLS	Poisson
R^2	0.12	0.08	0.05	0.07	0.05	0.06
N	733,088	733,088	733,088	6,639,634	6,639,634	6,639,634
Mean Dep Var	14.48	0.23	0.23	6.8	0.11	0.11

Notes: Table shows difference-in-differences estimates associated with equation (2) in the text. The two dependent variables are the log of out-of-pocket price plus one and the total cost share, which is defined as the ratio of out-of-pocket price to total price. The sample is Anthem lab and x-ray claims from 2008-2015 and the treated group consists of lab claims in the post period. Standard errors, clustered at the plan level are shown in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure 7: Annual Utilization Measures



Notes: The top panel plots the mean total annual lab spending for each of the three main insured groups. The middle planel plots the annual average of the any labs indicator, which is equal to one if the patient had any lab tests performed in the year and zero otherwise. The bottom panel plots the mean annual total number of labs.

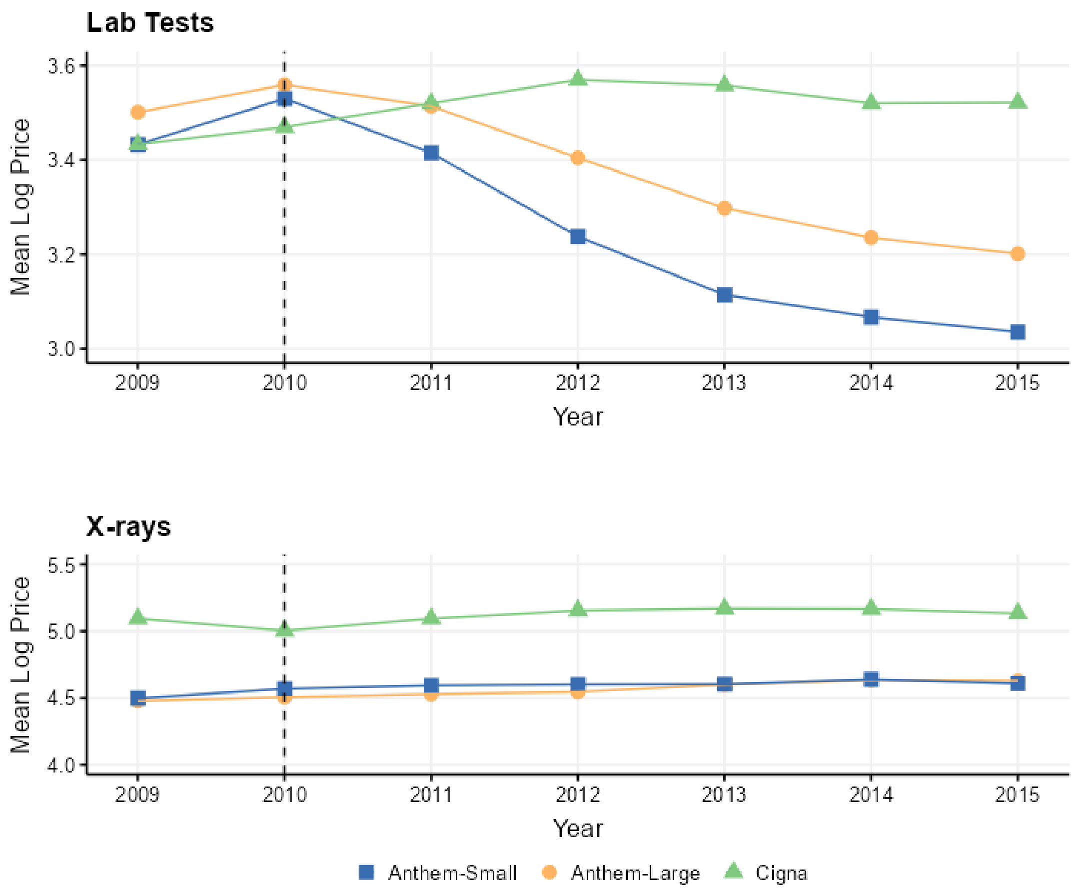
Appendix A: Additional Tables and Figures

Table A1: Demographic Comparison

	Cigna		Anthem	
	Mean	Sd	Mean	Sd
Age	40.89	17.55	41.90	17.38
Female	0.58	0.49	0.58	0.49
Charlson Score	0.36	0.92	0.33	0.88
Median Income	41681.27	7560.67	41013.26	7884.63
Bachelors	34.17	12.80	34.06	13.36
Modal HSA	30015		30015	
Modal ZIP	03301		03301	
Patients	419312		847104	

Notes: Table shows summary demographic information of Anthem and Cigna enrollees in the sample 2009-2015. The sample is composed of individual-years that appear in the primary analysis sample of lab and x-ray claims used throughout. Data on income and education by zip code are from the 2010-2014 American Community Survey. HSA denotes hospital service area.

Figure A1: Average Log Price Paid per Procedure by Group



Notes: Top panel shows the weighted average log lab price paid for claims associated with each of the three main insured groups described in the text. Weights are computed as each procedure's share of total volume over the sample period. The bottom panel shows the analogous plot for x-rays.

Table A2: Summary of Average Payment Changes and Resulting Savings

Test	2010 Price	2015 Price	% Change	Savings
80053 Comprehen metabolic panel	48.46	41.27	-14.83	-445.65
85025 Complete cbc w/auto diff wbc	34.25	32.09	-6.30	-114.44
80061 Lipid panel	55.25	40.21	-27.21	-775.12
84443 Assay thyroid stim hormone	69.72	59.42	-14.77	-352.73
83036 Glycosylated hemoglobin test	37.84	27.29	-27.89	-230.71
Aggregate	51.60	42.93	-16.55	-6125.96

Notes: 2010 price is the average amount paid for the indicated service across all claims associated with Anthem plans in the main sample. The 2015 price is the analog for 2015. Savings is computed as $\bar{p}_{15} * q_{15} - \bar{p_{10}} * q_{15}$, where $\bar{p}_{15}$ is the average price in 2015, $\bar{p}_{10}$ is the average price in 2010, and q_{16} is the quantity in 2015. For the aggregate measure, prices and the percent change are computed as weighted averages using spending share weights, and savings is the unweighted sum of total savings across all 200 lab procedures.

Table A3: Event Study Estimates

	Anthem-Small			Anthem-Large		
	(1)	(2)	(3)	(4)	(5)	(6)
Estimate Year						
2008	0.006			-0.095***		
	(0.03)			(0.019)		
2009	-0.024***	-0.089**	0.088***	-0.022*	-0.03*	0.091***
	(0.003)	(0.029)	(0.019)	(0.011)	(0.014)	(0.021)
2011	-0.144***	-0.171***	-0.12***	-0.07**	-0.103***	-0.038
	(0.017)	(0.012)	(0.023)	(0.022)	(0.025)	(0.027)
2012	-0.33***	-0.398***	-0.3***	-0.198***	-0.262***	-0.154**
	(0.038)	(0.007)	(0.038)	(0.058)	(0.053)	(0.059)
2013	-0.443***	-0.483***	-0.363***	-0.361***	-0.344***	-0.281***
	(0.024)	(0.011)	(0.036)	(0.039)	(0.043)	(0.048)
2014	-0.538***	-0.5***	-0.439***	-0.453***	-0.371***	-0.338***
	(0.02)	(0.019)	(0.036)	(0.032)	(0.036)	(0.042)
2015	-0.58***	-0.525***	-0.451***	-0.517***	-0.391***	-0.383***
	(0.035)	(0.015)	(0.042)	(0.027)	(0.031)	(0.038)
Model	DD	DD	DDD	DD	DD	DDD
Coefficient	year*lab	year*anthem	year*lab*anthem	year*lab	year*anthem	year*lab*anthem
R^2	0.54	0.39	0.46	0.55	0.45	0.5
N	733,088	3,412,048	3,631,900	6,639,634	8,228,967	8,862,894
Mean Dep Var	50.31	51.79	59.29	53.78	50.21	56.46

Notes: Table shows event-study estimates associated with equation (1) in the text. The dependent variable is the log of the total price paid for a particular procedure. In the $year * lab$ case, the sample is Anthem claims and the treated group consists of lab claims in the post period. In the $year * anthem$ case, the sample is entirely lab claims from both Anthem and Cigna, and the treated group consists of Anthem claims in the post period. In the $year * lab * anthem$ case, the sample contains both lab and x-ray procedures from both insurers, and the treated group consists of Anthem lab claims in the post period. Standard errors, clustered at the plan level are shown in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table A4: Difference-in-Differences Estimates: Price Paid Per Procedure (Poisson GLM)

	Anthem-Small			Anthem-Large		
	(1)	(2)	(3)	(4)	(5)	(6)
Estimate	-0.24***	-0.231***	-0.226***	-0.165***	-0.161***	-0.161***
	(0.04)	(0.012)	(0.03)	(0.027)	(0.027)	(0.026)
Model	DD	DD	DDD	DD	DD	DDD
Coefficient	post*lab	post*anthem	post*lab*anthem	post*lab	post*anthem	post*lab*anthem
Psuedo R^2	0.53	0.44	0.51	0.55	0.49	0.53
N	733,088	3,412,048	3,631,900	6,639,634	8,228,967	8,862,894
Mean Dep Var	50.31	51.79	59.29	53.78	50.21	56.46

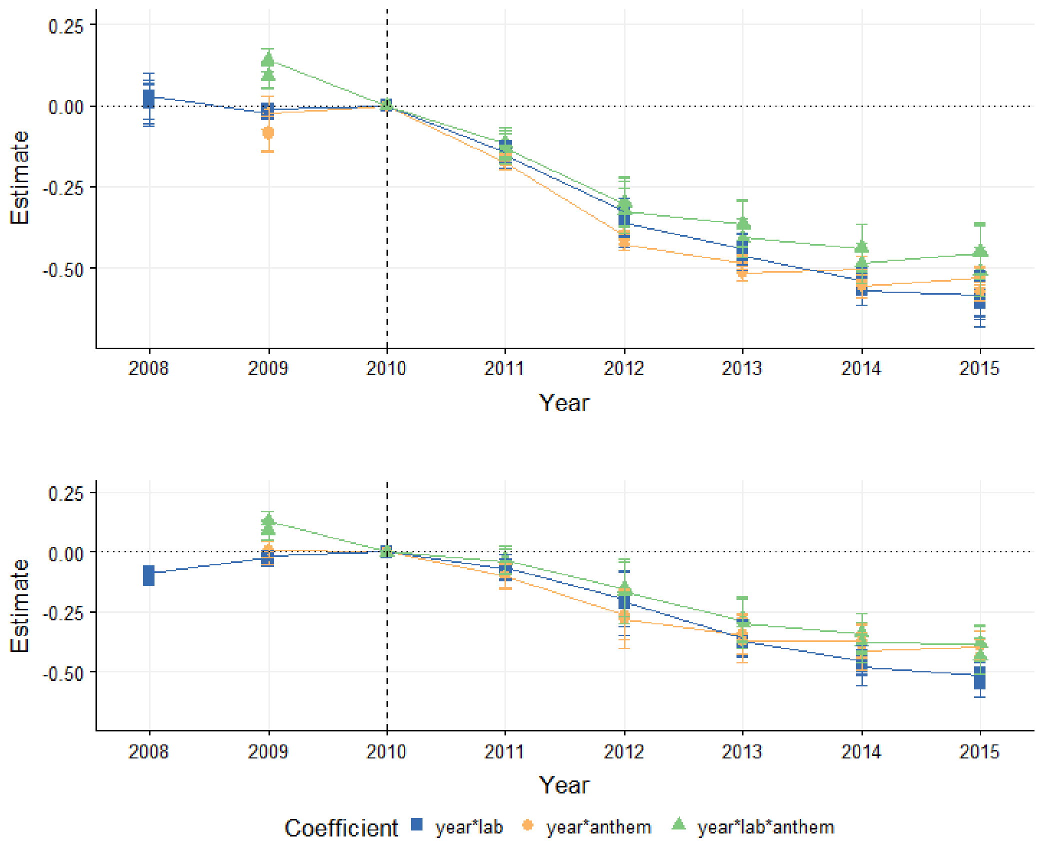
Notes: Table shows difference-in-differences and triple-differences estimates associated with equation (2) in the text, estimated as a GLM with a log link function. The dependent variable is the log of the total price paid for a particular procedure. In the $year * lab$ case, the sample is Anthem claims and the treated group consists of lab claims in the post period. In the $year * anthem$ case, the sample is entirely lab claims from both Anthem and Cigna, and the treated group consists of Anthem claims in the post period. In the $year * lab * anthem$ case, the sample contains both lab and x-ray procedures from both insurers, and the treated group consists of Anthem lab claims in the post period. Standard errors, clustered at the plan level are shown in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table A5: Difference-in-Differences Estimates: Price Paid Per Procedure (Additional Specifications)

Specification:	No Controls	HSA FE	Individual FE	Top 20 Procs
Panel A: Anthem Small Group				
Estimate	-0.374***	-0.375***	-0.342***	-0.415***
	(0.05)	(0.045)	(0.032)	(0.05)
R^2	0.51	0.54	0.74	0.54
N	733,088	733,088	733,088	419,807
Panel B: Anthem Large Group				
Estimate	-0.252***	-0.262***	-0.259***	-0.281***
	(0.045)	(0.045)	(0.037)	(0.052)
R^2	0.51	0.55	0.74	0.55
N	6,639,634	6,639,634	6,639,634	3,933,543

Notes: Table shows estimates for several alternative specifications of the difference-in-differences model (2) in the text. Here, the sample is all Anthem lab and x-ray claims and the treated group consists of lab claims in the post period. The first specification excludes controls, the second specification uses zip-code-level fixed effects, the third specification includes patient-level fixed effects, and the fourth specification is estimated on the sample of the 20 most common lab and x-ray procedures. Standard errors, clustered at the plan level are shown in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure A2: Event Study Estimates: Price Paid Per Procedure (Additional Specifications)



Notes: Event-study estimates for alternative specifications associated with A5. Each line reflects the estimates from a particular version of the general event-study equation, where the dependent variable is the log of the total price paid for a particular procedure. In the $year * lab$ case, the sample is Anthem claims and the treated group consists of lab claims in the post period. In the $year * anthem$ case, the sample is entirely lab claims from both Anthem and Cigna, and the treated group consists of Anthem claims in the post period. In the $year * lab * anthem$ case, the sample contains both lab and x-ray procedures from both insurers, and the treated group consists of Anthem lab claims in the post period. Standard errors are clustered at the plan level. Cigna claims begin in 2009, so models with Cigna are estimated on 2009-2015 data, while the Anthem-only model is estimated using 2008-2015 data.

Table A6: Heterogeneous Effect Estimates

	Small Group	Large Group
Treat	-0.024	-0.154***
	(0.029)	(0.022)
Treat × Age	-0.008***	-0.003***
	(0.002)	(0.001)
Treat × Charlson Index	0.042***	0.031***
	(0.006)	(0.005)
Treat × Female	-0.038***	-0.016**
	(0.006)	(0.005)
R^2	0.55	0.55
N	733,088	6,639,634

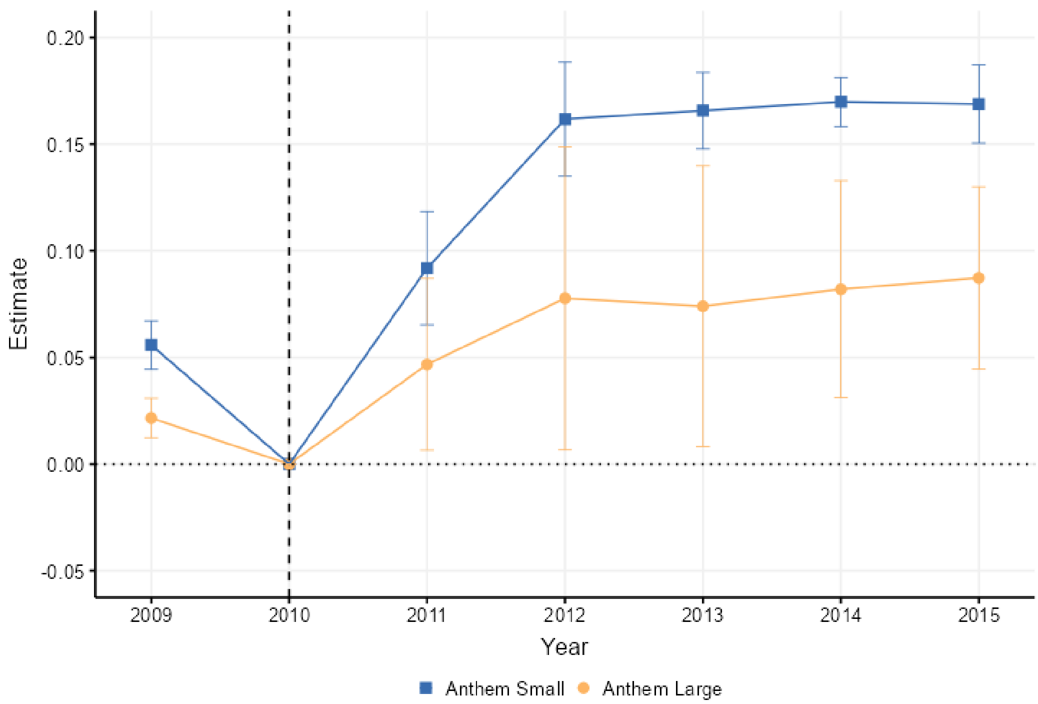
Notes: Table shows estimates for several interacted specifications of the difference-in-differences model (2). Here, the sample is all Anthem lab and x-ray claims and the treated group consists of lab claims in the post period. Standard errors, clustered at the plan level are shown in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

Table A7: Difference-in-Differences Estimates of Effect Independent Lab Use

Treated Group:	Anthem-Small	Anthem-Small	Anthem-Large	Anthem-Large
Estimate	1.369***	1.346***	0.095**	0.095**
	(0.196)	(0.197)	(0.03)	(0.03)
Individual FEs	No	Yes	No	Yes
R^2	0.19	0.63	0.17	0.59
N	3,412,048	1,692,867	8,228,967	4,924,108
Mean Dep Var	0.19	0.18	0.17	0.16

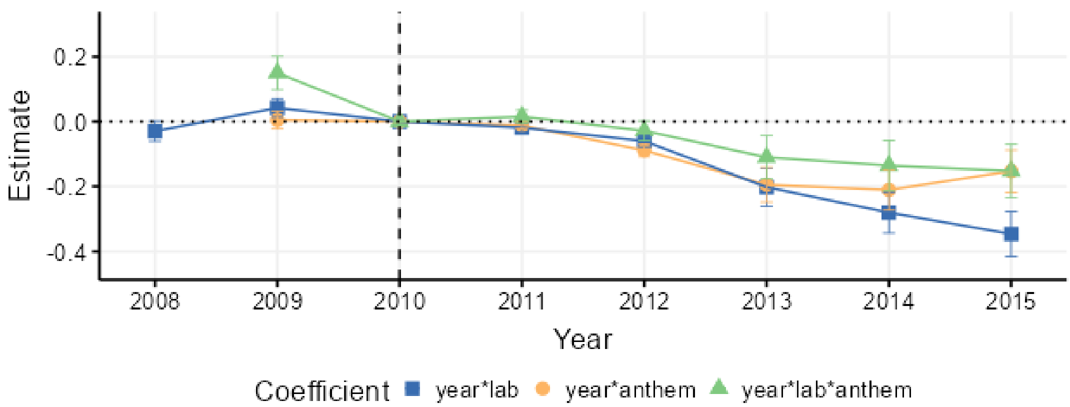
Notes: Table shows difference-in-differences estimates associated with equation (3) in the text, estimated as a logit GLM. The dependent variable is an indicator that is equal to one if the servicing provider is an independent lab and zero otherwise. The models with individual fixed effects are estimated on the sample of patients with a lab claim in both the pre period and post period.

Figure A3: Event Study Plots for Alternative Specifications



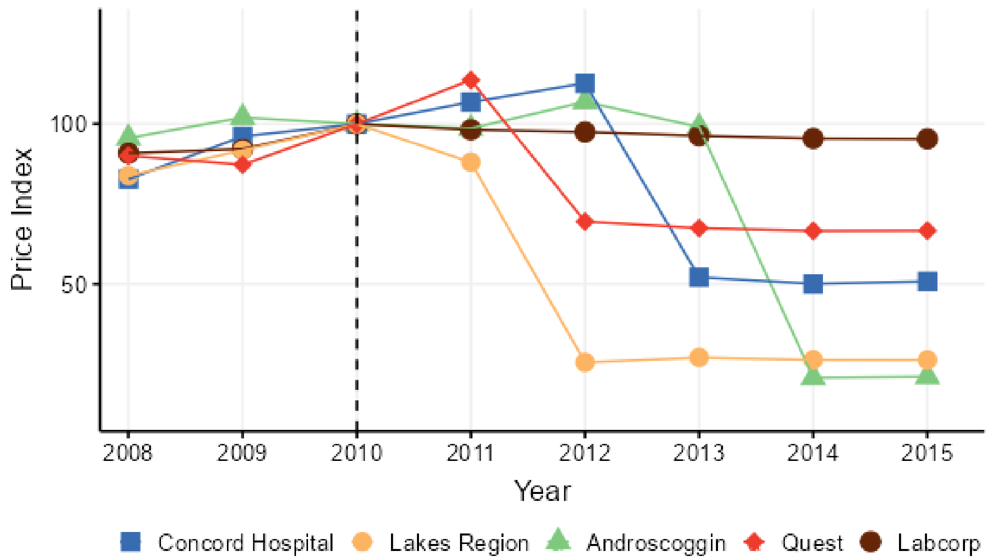
Notes: Event-study-version estimates associated with (3) in the text and Table 3. The dependent variable is an indicator that is equal to one if the servicing provider is an independent lab and zero otherwise.

Figure A4: Event Study Estimates: Negotiated Prices



Notes: Each line reflects the estimates from a particular version of the general event-study equation, where the dependent variable is the log of the total price paid for a particular procedure. In the $year * lab$ case, the sample is Anthem claims and the treated group consists of lab claims in the post period. In the $year * anthem$ case, the sample is entirely lab claims from both Anthem and Cigna, and the treated group consists of Anthem claims in the post period. In the $year * lab * anthem$ case, the sample contains both lab and x-ray procedures from both insurers, and the treated group consists of Anthem lab claims in the post period. Standard errors are clustered at the plan level. Cigna claims begin in 2009, so models with Cigna are estimated on 2009-2015 data, while the Anthem-only model is estimated using 2008-2015 data.

Figure A5: Price Index by Preferred Group



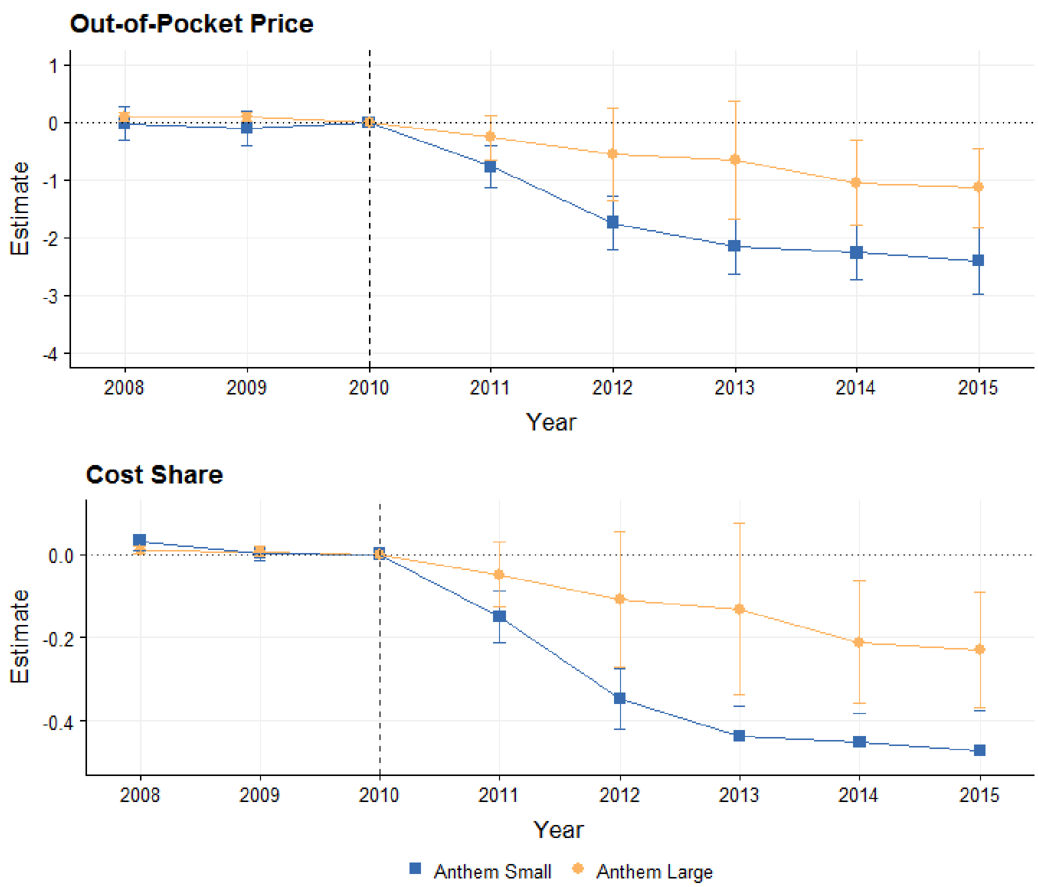
Notes: Figure shows Laspeyres-type price index estimates associated with equation (6) in the text for a handful of providers. 2010 is taken as the reference period for each group.

Table A8: Difference-in-Differences Estimates: Negotiated Prices

	Newly Preferred			Never Preferred		
	(1)	(2)	(3)	(4)	(5)	(6)
Estimate	-0.338***	-0.454***	-0.366***	-0.063***	-0.068***	-0.073**
	(0.032)	(0.034)	(0.038)	(0.013)	(0.013)	(0.026)
Model	DD	DD	DDD	DD	DD	DDD
Coefficient	post*lab	post*anthem	post*lab*anthem	post*lab	post*anthem	post*lab*anthem
R^2	0.7	0.78	0.72	0.75	0.8	0.77
N	1,551,904	1,324,122	1,630,399	5,132,746	4,137,466	5,533,451
Mean Dep Var	59.23	62.31	65.9	59.23	62.31	65.9

Notes: Table shows difference-in-differences and triple-differences estimates associated with equation (5) in the text. The dependent variable is the log of the total procedure price. Sample is restricted to hospital-based providers that perform both lab tests and x-rays. Standard errors, clustered at the plan level are shown in parentheses. * $p < .05$, ** $p < .01$, *** $p < .001$.

Figure A6: Event Study Estimates: Cost Sharing



Notes: Event-study-version estimates associated with equation (2) in the text and Table 6. In the top panel the dependent variable is the patient's out-of-pocket price. In the bottom panel the dependent variable is the ratio of out-of-pocket price to total price.

Appendix B: Additional Tables and Figures

Table B1: Summary of Top 50 Lab Procedures

CPT/HCPCS	Description	Mean Price	SD Price	Count	Type
80053	Comprehen metabolic panel	46.75	33.73	836,917	lab
85025	Complete cbc w/auto diff wbc	33.94	20.80	807,641	lab
80061	Lipid panel	49.14	32.28	758,076	lab
84443	Assay thyroid stim hormone	64.70	33.28	516,300	lab
83036	Glycosylated hemoglobin test	33.82	20.46	266,323	lab
80048	Metabolic panel total ca	38.56	63.53	261,400	lab
82306	Vitamin d 25 hydroxy	96.15	67.72	220,243	lab
85610	Prothrombin time	18.40	71.45	215,176	lab
87086	Urine culture/colony count	28.28	19.60	203,434	lab
81001	Urinalysis auto w/scope	17.09	16.64	187,549	lab
85027	Complete cbc automated	26.96	88.56	175,085	lab
88305	Tissue exam by pathologist	287.32	301.82	156,771	lab
84153	Assay of psa total	49.68	30.27	146,668	lab
81003	Urinalysis auto w/o scope	10.63	17.78	137,637	lab
80076	Hepatic function panel	35.20	25.76	135,601	lab
88175	Cytopath c/v auto fluid redo	85.25	43.05	131,352	lab
84439	Assay of free thyroxine	38.11	27.59	121,332	lab
88142	Cytopath c/v thin layer	61.12	34.81	116,714	lab
85652	Rbc sed rate automated	12.94	12.70	87,650	lab
87491	Chylmd trach dna amp probe	87.16	46.27	86,595	lab
80050	General health panel	48.90	38.08	85,835	lab
87591	N.gonorrhoeae dna amp prob	85.89	45.06	85,812	lab
87081	Culture screen only	25.76	17.52	85,424	lab
87070	Culture othr specimn aerobic	35.65	29.01	84,246	lab
82947	Assay glucose blood quant	18.11	33.36	82,398	lab
82043	Microalbumin quantitative	24.41	20.40	72,743	lab
84460	Alanine amino (ALT) (SGPT)	22.46	16.44	71,366	lab
82550	Assay of ck (cpk)	36.19	28.14	70,792	lab
82728	Assay of ferritin	52.24	32.69	70,776	lab
82565	Assay of creatinine	20.77	121.30	69,989	lab
83690	Assay of lipase	35.87	20.04	69,747	lab
82607	Vitamin B-12	45.21	29.93	67,140	lab
87186	Microbe susceptible mic	40.17	30.71	66,983	lab
84484	Assay of troponin quant	77.45	66.17	66,896	lab
87880	Strep a assay w/optic	31.22	21.49	62,494	lab
87077	Culture aerobic identify	36.63	24.84	61,620	lab
83735	Assay of magnesium	30.24	23.91	58,618	lab
83540	Assay of iron	24.70	16.66	58,138	lab
81002	Urinalysis nonauto w/o scope	6.52	6.85	54,928	lab
86140	C-reactive protein	26.20	20.28	54,189	lab
82570	Assay of urine creatinine	20.57	15.69	52,305	lab
84450	Transferase (AST) (SGOT)	22.17	15.04	50,941	lab
85730	Thromboplastin time partial	31.88	22.98	48,781	lab
84550	Assay of blood/uric acid	17.12	16.61	48,111	lab
87621	Hpv dna amp probe	88.61	54.48	47,728	lab
83550	Iron binding test	30.84	18.43	45,445	lab
81025	Urine pregnancy test	27.48	22.51	42,751	lab
84520	Assay of urea nitrogen	17.74	49.44	41,468	lab
85018	Hemoglobin	11.57	10.32	32,199	lab
89240	Pathology lab procedure	1147.61	2803.46	60	lab

Table B2: Summary of X-ray Procedures

CPT/HCPCS	Description	Mean Price	SD Price	Count	Type
71020	Chest x-ray 2vw frontal	latl	141.01	135.23	149,650
x-ray					
73630	X-ray exam of foot	124.11	88.60	51,866	x-ray
73610	X-ray exam of ankle	124.83	81.62	39,778	x-ray
73030	X-ray exam of shoulder	122.13	92.04	39,683	x-ray
73562	X-ray exam of knee 3	121.42	92.42	34,371	x-ray
72100	X-ray exam l-s spine 2/3 vws	141.68	83.57	32,246	x-ray
71010	Chest x-ray 1 view frontal	119.62	100.35	30,317	x-ray
74000	X-ray exam of abdomen	115.10	72.66	28,906	x-ray
73110	X-ray exam of wrist	126.62	104.84	28,345	x-ray
73130	X-ray exam of hand	134.61	94.86	27,324	x-ray
73140	X-ray exam of finger(s)	102.63	69.77	22,742	x-ray
73564	X-ray exam knee 4 or more	158.23	114.00	21,049	x-ray
73510	X-ray exam of hip	129.17	74.64	19,569	x-ray
72170	X-ray exam of pelvis	92.47	68.18	17,916	x-ray
73560	X-ray exam of knee 1 or 2	107.88	82.85	16,221	x-ray
72040	X-ray exam neck spine 3/¡vws	132.51	79.31	14,015	x-ray
72050	X-ray exam neck spine 4/5vws	202.77	112.59	12,644	x-ray
73080	X-ray exam of elbow	131.92	82.84	11,830	x-ray
73590	X-ray exam of lower leg	119.57	82.07	9,597	x-ray
74020	X-ray exam of abdomen	152.94	80.80	9,476	x-ray
72110	X-ray exam l-2 spine 4/¿vws	194.08	107.16	9,237	x-ray
73565	X-ray exam of knees	80.74	69.88	8,720	x-ray
77421	Stereoscopic x-ray guidance	338.15	195.44	7,790	x-ray
73090	X-ray exam of forearm	125.40	82.08	6,955	x-ray
72070	X-ray exam thorac spine 2vws	133.11	74.42	6,671	x-ray
74022	X-ray exam series abdomen	219.04	112.06	6,369	x-ray
71101	X-ray exam unilat ribs/chest	184.47	105.96	5,957	x-ray
73100	X-ray exam of wrist	79.17	52.30	5,698	x-ray
73500	X-ray exam of hip	69.17	50.96	5,446	x-ray
73000	X-ray exam of collar bone	113.22	77.58	5,236	x-ray
73660	X-ray exam of toe(s)	105.96	65.97	5,170	x-ray
73520	X-ray exam of hips	136.41	86.85	4,401	x-ray
73620	X-ray exam of foot	84.76	78.90	4,010	x-ray
73650	X-ray exam of heel	103.73	80.64	3,985	x-ray
73070	X-ray exam of elbow	98.11	84.97	3,877	x-ray
74220	Contrast x-ray esophagus	295.91	133.73	3,804	x-ray
70030	X-ray eye for foreign body	96.34	61.31	3,760	x-ray
73550	X-ray exam of thigh	120.14	82.39	3,238	x-ray
73060	X-ray exam of humerus	127.85	79.18	3,230	x-ray
73120	X-ray exam of hand	99.21	88.33	3,108	x-ray
73040	Contrast x-ray of shoulder	348.97	239.08	3,021	x-ray
72020	X-ray exam of spine 1 view	140.01	629.35	2,934	x-ray
23350	Injection for shoulder x-ray	332.45	232.41	2,878	x-ray
76098	X-ray exam breast specimen	146.86	212.69	2,809	x-ray
73600	X-ray exam of ankle	75.30	65.15	2,757	x-ray
71100	X-ray exam ribs uni 2 views	121.42	65.16	2,601	x-ray
73010	X-ray exam of shoulder blade	62.72	46.44	2,561	x-ray
70220	X-ray exam of sinuses	181.02	92.71	2,435	x-ray
72052	X-ray exam neck spine 6/¿vws	265.25	135.23	1,817	x-ray
72010	X-ray exam spine ap	lat	178.78	119.02	888
x-ray					

Figure B1: Anthem Site of Service Program Advertisement

Figure B2: Anthem Site of Service Program Advertisement

Anthem BlueCross BlueShield

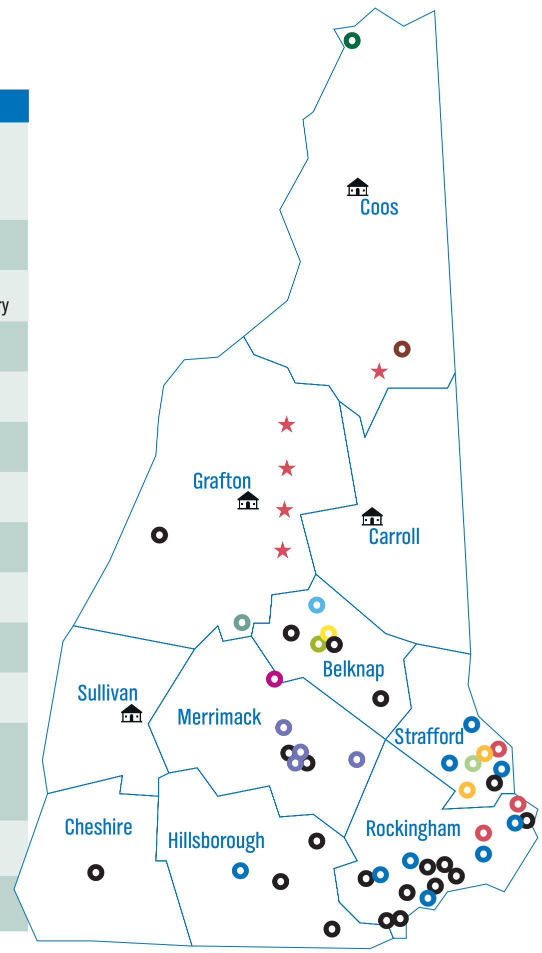
Find a high-quality, Site of Service lab location today!

Great news! We've got high-quality, Site of Service lab choices for you.

Please take a look at our map and contact list of labs in New Hampshire.
As a member, your visits will be covered 100% by your health plan.

For the most up-to-date information or for questions, please contact Anthem Member Services at 1-800-870-3122 or visit our website at www.anthem.com/siteofservicenh

Name	Phone number/Website
★ Ammonoosuc Community Health Services ★ *Lab services are done by Quest Diagnostics, limited to patients at this practice.*	ammonoosuc.org
Androscoggin Valley Hospital	1-603-752-2200 avhnh.org
Concord Hospital Independent Outpatient Laboratory	1-603-225-2711 concordhospital.org/services/laboratory
Franklin Regional Hospital Lab	1-603-934-2060 ext. 3200 lrgh.org
Granite State Lab	1-603-330-7057 granitestatelab.com
Interlakes Clinical Laboratory	1-603-279-2269 lrgh.org
Laboratory Corporation of America	1-855-277-8669 labcorp.com
Laconia Clinic Laboratory	1-603-524-5151 lrgh.org
Lakes Region General Hospital Lab	1-603-524-3211 lrgh.org
Newfound Clinical Lab	1-603-744-5441 ext. 1403 lrgh.org
Nordx	1-800-773-5814 nordx.org
Quest Diagnostics Incorporated *Quest provides at-home lab service in certain ZIP codes within Carroll, Coos, Grafton and Sullivan counties.*	1-866-697-8378 questdiagnostics.com
Upper Connecticut Valley Hospital	1-603-237-4971 ucvh.org
Wentworth-Douglass Hospital	1-603-742-5252 wdhospital.com



- Androscoggin Valley Hospital
- Concord Hospital Independent Outpatient Laboratory
- Franklin Regional Hospital Lab
- Granite State Lab
- Interlakes
- LabCorp
- Laconia
- Lakes Region
- Newfound Clinical Lab
- Nordx
- Quest Diagnostics Incorporated
- *Quest provides at-home lab service in certain ZIP codes within Carroll, Coos, Grafton and Sullivan counties.*
- Upper Connecticut Valley Hospital
- Wentworth-Douglass Hospital

Figure B3: Anthem Site of Service Program Advertisement

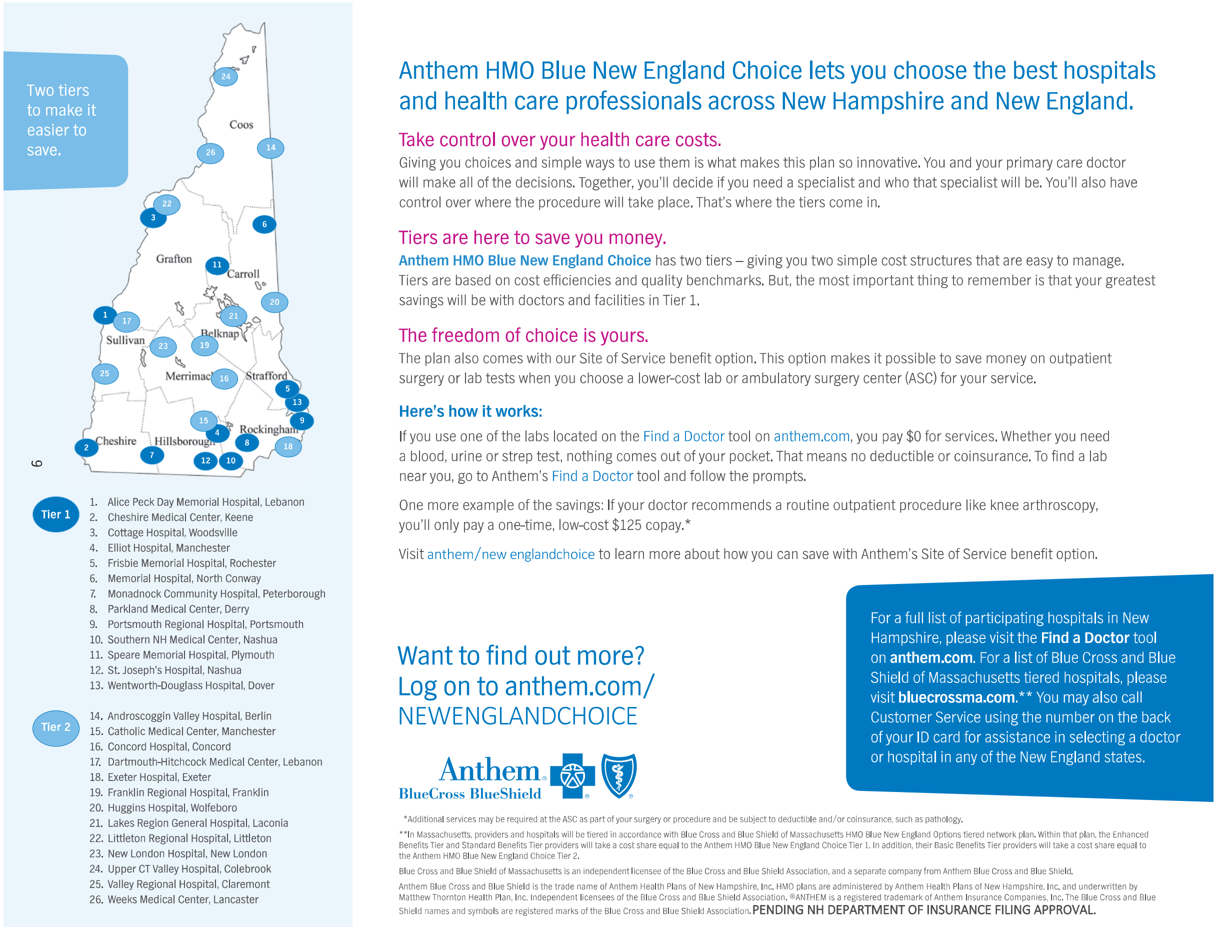


Figure B4: Anthem Site of Service Program Advertisement

Site of Service

Frequently asked questions

Here are answers to some common questions about the Site of Service benefit option:

Q: How does Anthem's Site of Service benefit option help lower costs?

A: It can reduce your out-of-pocket costs, as well as the overall cost of health care. By using labs found on Anthem's Provider Finder, which include all independent labs in New Hampshire and some hospital labs, you will not have a copay for your lab services. It's one of the ways we are working to help ensure you have access to affordable care when you need it.

Q: How do I find a zero-cost lab through the Site of Service benefit?

A: We have an independent lab flier with a map and contact list for state of New Hampshire members. However, for the most up-to-date listings, log on to **anthem.com** and select **Find a Doctor**. Then, follow the prompts.

Q: What if I don't use a zero-cost lab located on Anthem's *Provider Finder*?

A: You may use any lab you'd like. However, if you choose a lab that's not listed on Provider Finder, your share of the costs may be higher.

Q: Does an independent lab provide the same type and quality of service as a hospital outpatient lab?

A: Yes, independent labs specialize in these types of services. They have the same level of quality and offer the same types of services as other outpatient labs.

Q: Can my doctor use an independent lab for my lab work?

A: Yes, when your doctor orders lab work, ask him or her to order your lab work at an independent lab. Even if your doctor usually uses another hospital lab not found on Provider Finder, he or she can order your lab work at an independent lab.

Q: My doctor doesn't have a contract with any independent labs. How can he or she get one?

A: Your doctor can easily set up an account with independent labs (e.g., LabCorp, Quest Diagnostics®, Converge™ Diagnostic Services LLC or NorDx) through their websites.

While the account is being set up, your doctor can still order your lab work at an independent lab by giving you the laboratory order form. You'll need to take the form to the lab. Once your doctor has an account with the lab, the lab will automatically pick up your specimen(s) from the doctor's office.

Q: My doctor says he or she doesn't want to use one of the labs that I've requested under the Site of Service benefit option. What can I do?

A: Remind your doctor that under New Hampshire law, you have the right to choose where you want to receive services, including lab services. Your doctor is required to give you a laboratory order form to take to the lab you want to use.

Q: What are ASCs?

A: ASCs are ambulatory surgery centers that provide a wide range of same-day surgical services, such as a tonsillectomy or knee arthroscopy.

Q: How can I find an ASC?

A: Go to **anthem.com/stateofnhsaves** and select Ambulatory Surgery Centers, or call the Designated Service unit at 1-800-933-8415. To learn additional details about the ASCs, including location, hours, and affiliated doctors, please call the facility directly.

Q: I'm going to have surgery. How can I find out whether it can be done at an ASC?

A: Ask your doctor to recommend a surgeon who may be able to perform your surgery at an ASC.

Q: How much will I have to pay if I use an ASC?

A: If you use an ASC found on **anthem.com/stateofnhsaves** for same-day surgery, you'll have a $0 copay.*

* If you need additional care as part of the surgery or procedure you may have to pay coinsurance and/or a deductible (the amount you pay before the plan pays). For example, you may pay more for pathology or lab work if it is not sent to one of the labs found on Anthem's Provider Finder. Make sure your doctor at the ASC knows you have Anthem's Site of Service benefit option.

Appendix C: Conceptual Framework

Here I present a simple conceptual framework to motivate the empirical analyses presented in this paper. At the individual-choice level, I consider a standard discrete choice setup. Given a set of patients who are seeking medical service k at time t, each patient i chooses a provider r from among the set of available providers for this service R_k. Each patient is enrolled in a particular insurance plan j, which defines a mapping from providers and negotiated prices to out-of-pocket prices: $OOP_{jkr} = C_j(p_{jkr})$.

The utility associated with provider r depends on the out-of-pocket price, an out-of-pocket sensitivity parameter θ, characteristics of the patient and provider, X_{ikr}, and a random component, ϵ:

$$u_{ijkr} = \theta OOP_{ijkr} + g(X_{ikr}; \beta) + \epsilon_{ijkr} \tag{10}$$

Let s_{ijkr} denote the probability that i chooses provider r, where the functional form of s_{ijkr} is defined by distributional assumptions on (1). These choice probabilities then define a distribution of outcomes conditional on negotiated prices. Suppose also that negotiated prices are determined via the Nash bargaining framework of Gowrisankaran et al. (2015). Specifically, equilibrium prices depend on: (1) net surplus to the insurer of including a particular provider in their network, $V_{jkr}(p_{jkr}, C_j)$, and (2) the profit accruing to the provider from this network:

$$\pi_{jkr}(p_{jkr}, C_j) = \sum_{i \in I_{jk}} s_{ijkr}[p_{jkr} - mc_{kr}] \tag{11}$$

Equilibrium prices are then assumed to satisfy:

$$p^*_{jkr} = \text{argmax } V_{jkr}(p_{jkr}, C_j)^{\sigma} \pi_{jkr}(p_{jkr}, C_j)^{1-\sigma} \tag{12}$$

where σ is a bargaining power parameter. This model of underlying choice behavior and negotiations defines an equilibrium system from which aggregate quantities of interest emerge. The model also makes clear the channels through which cost-sharing design C_j is related to average spending per procedure, which is the focus of this paper. Specifically, the principal quantity of interest here is the total effect of switching from design j to design j', defined by:

$$TE_{j,j'} = \sum_{R_k} p^*_{j'kr} \times D_{kr}(p^*_{j'kr}, C_{j'}) - \sum_{R_k} p_{jkr} \times D_{kr}(p_{jkr}, C_j) \tag{13}$$

where $D_{kr}(p_{j'kr}, C_{j'})$ is the market share of provider r for service k.[21]

In this paper, I pursue a reduced-form analysis of total effects based on time-series variation in realized prices and quantities. An alternative approach is to estimate structural parameters in the equilibrium model above and then simulate counterfactual outcomes (Prager, 2016; Prager, 2020; Brown, 2019; Ackley, 2020; Ackley, 2022).

[21]Specifically, $D_{kr}(p_{j'kr}, C_{j'}) = \frac{1}{\sum_r \sum_i s_{ijkr}} \sum_i s_{ijkr}$