The Impact of Price (Charge) Transparency in Outpatient Provider Markets

Hunt Allcott, Kayleigh Barnes, Sherry Glied, Ben Handel, Grace Kim*

April 4, 2021

Abstract

Medical provider price transparency is often touted as a key policy for efficiently lowering health care spending, which is nearly 20% of GDP. Despite its many proponents, the impact of price transparency is theoretically ambiguous: it could lower health care spending via increased consumer price shopping or improved insurer bargaining position but could instead raise health care prices via improved provider bargaining or either tacit or explicit provider collusion. We conduct a randomized-controlled trial to examine the impact of a state-wide medical charge transparency tool in outpatient provider markets in the state of New York. In the experiment, individual providers' billed charges (list prices) were released randomly at the procedure X geozip level. We use a comprehensive commercial claims database to assess the impact of this intervention and find that the intervention causes a small increase in overall billed charges (+1%) but a relatively lower increase in the charges for procedures with many out-of-network claims (-2%). We find no evidence for quantity effects. We find larger price increases for specific categories that are almost always insured and less elective in nature, e.g. MRI (+6%) and radiology (+3%) and price decreases for categories that are less often insured and more elective in nature, e.g. psychology (-2%) and chiropractor (-3%) services. Taken together, these results are consistent with our intervention having a minimal effect on consumer price shopping but a meaningful effect driving provider price increases, especially for less elective services that are almost always covered by insurance, potentially reflecting perverse price effects resulting from tacit collusion or reduced information asymmetries.

JEL Codes: I11, I12, I13, I130, L11.

Keywords: health, price transparency, physicians, prices.

*Allcott: Microsoft, Harvard. hunt.allcott@nyu.edu. Barnes: UC Berkeley. Glied: New York University, NBER. Handel: UC Berkeley. Kim: New York University. The study was deemed not to have human subjects research by the New York University Institutional Review Board (IRB-FY2017-582). We acknowledge funding from the New York State Health Foundation under grant number 17-04903. Replication files are available from Hunt Allcott. We thank seminar participants at the NYU Wagner seminar and APPAM Fall Conference for comments, including Christopher Whaley, Kathy O'Regan, and Leanna Stiefel.

1 Introduction

Health care policymakers, insurers, and private companies have frequently discussed the transparency of health care pricing information as a way to reign in rising health care spending (Reinhardt (2006); Sinaiko and Rosenthal (2016); Volpp (2016)). Prices in health care markets are notoriously variable, opaque, and confusing and price transparency has the potential to reduce search costs and information asymmetries for consumers seeking cheap, high value health care. Most consumers purchasing services, especially those who intentionally or inadvertently purchase services outside their insurer networks, have no easy way to ascertain or compare prices charged by different providers. Recently, the federal government has sought to use regulatory levers to make healthcare prices more transparent, including upholding a CMS 2019 Final Rule that mandates hospitals release comprehensive information regarding negotiated rates, and was in effect as of 2021 (Wilensky (2019); Kullgren and Fendrick (2021); Glied (2021)).

Proponents of price transparency note the substantial potential benefits in (i) making price information available to consumers and (ii) reducing consumer search costs by making this information easily accessible. They also note the potential benefits for provider pricing: if consumers are more price elastic then providers may lower prices in this more competitive landscape. This may result either from providers choosing prices freely or as the result of insurer-provider bargaining if price transparency provides a competitive advantage to insurers.

However, despite these evident benefits, skeptics are concerned that the release of pricing information might lead to higher prices either via collusion among providers (tacit or explicit) or via providers gaining a competitive advantage from realizing that they are systematically under-pricing relative to similar quality peers. In a market with clear capacity constraints, as in health care, information may help motivate providers to raise prices under the realization that they can charge more without decreasing their quantity of services. Alternatively, if consumers equate price with quality, providers might raise prices to signal quality. Albaek et al. (1997) document exactly this phenomenon in the context of the Danish concrete industry, where the publishing of transaction prices produced a 15-20% increase in concrete prices, as producers stopped offering confidential discounts to purchasers. With these kinds of issues in mind, some policymakers and economists have urged caution in promoting price transparency (see, e.g., Cutler and Dafny (2011)).

In this study, we examine price transparency for providers across the state of New York, using a randomized controlled trial embedded in the information provision platform run by FAIR Health, a non-profit organization dedicated to promoting price transparency in health care markets. Prior to our intervention, FAIR Health provided market-level information on typical prices for a given procedure, but did not provide information for specific individual providers. We partnered with FAIR Health to implement a statewide randomized intervention providing individual-level provider billed charge information on their platform. We randomized whether this information would be provided for providers at the procedure-geozip level. For a given kind of procedure in a given market, all providers of that procedure are either randomized into our individual-level price information treatment or randomized out of the treatment. [1]

The design was set up specifically to capture market-level pricing and demand effects as well as the effects for specific kinds of individual providers. Our intervention applies to 107 procedures and all geozips in New York and was in place for two years so that we could assess its medium-run effects. The data captured all commercial claims in the FAIR Health data ware house related to these procedures and geozips and encompassed over 110 million claims and over 205,000 providers.We present a series of descriptive statistics showing (i) that our intervention is well-balanced across treatment and control arms (ii) that there is meaningful heterogeneity across procedure-geozip markets in market power, out-of-network claims, and ex ante price dispersion and (iii) that usage of the information tool we study averages just fewer than 10,000 uses per month.

Our randomized intervention allows us to cleanly identify the net price and quantity effects of our information-provision tool while also allowing us to study heterogeneous effects related to (i) initial absolute procedure prices (ii) initial prices relative to peers (iii) specific types of medical procedures and (iv) specific kinds of market structures. Importantly, since our information provision focuses on individual provider billed charges, rather than negotiated rates between insurers and providers, our analysis is especially relevant for the out-of-network services that these charges are germane to. However, since billed charges also feed into negotiated rates with insurers, and are meaningfully correlated with them (Batty and Ippolito (2017)), we also study services that are shoppable but typically received in-network. To our knowledge, this is the first randomized intervention of a price transparency tool that is specifically designed to address market-level effects as well as the effects on consumers and specific providers.[2]

Our randomized intervention directly guides our econometric approach to estimating our effects of interest. We use a triple difference-in-differences approach that compares key price and quantity outcomes for providers in treated procedure-location pairs to the same outcomes for providers in control procedure-location pairs. We control for procedure, geozip, and trimester fixed effects and also leverage time-series variation spanning the periods pre- and post-intervention. We conduct a number of robustness analyses to our main triple difference-in-differences specification including, e.g., a difference-in-differences version without time fixed effects, and find similar results for these alternatives.

[1] The trial was submitted to the AEA RCT registry.

[2] Several prior major studies in health economics rely on randomized controlled trials as a 'gold standard' for identification. See, for example, the 1974 RAND Health Insurance Experiment (HIE) studying the price elasticity of demand for health care and the Oregon Health Insurance Experiment studying the effects of expanding access to public health insurance (Manning et al. (1987); Newhouse (1996); Finkelstein et al. (2012)).

In our primary triple difference-in-differences specification, we find that, across all procedures and locations, providing individual-level provider charge information increases prices by 1.2% and has no statistically significant impact on quantity. We assess these impacts separately for providers whose prices were initially above or below the median for a given procedure in their geozip and find modest but larger increases in prices for providers who were initially below the median. In addition, we find no quantity impacts for providers who were initially high-priced as opposed to low-priced, suggesting that our intervention had no meaningful impact on the extent of consumer price shopping. We also find that providers who are in procedure markets that are above median market concentration, measured with procedure-geozip HHI, have slightly larger price increases than those below median, with no statistically significant quantity differences.

Given that our intervention provides information on billed charges, rather than insurer-contracted prices, we focus especially on out-of-network claims, for which billed charges are relevant.[3] We find that procedures with a high proportion of out-of-network claims have essentially no price change while procedures with a low proportion of out-of-network claims have a 2.5% price increase as a result of our intervention. This suggests that, for procedures where billed charges are much closer to final prices, there is no impact of our intervention, while for procedures that are more likely to be covered by insurance, prices increase. This could be, e.g., because providers serving out-of-network patients suspect that those patients will respond to prices in the medium to long run and thus be less willing to raise prices relative to other providers.

We also investigate the effects of our intervention for specific procedure categories. We find larger price increases for specific categories that are almost always insured and less elective in nature, including MRI (+6%) and radiology (+3%) services. We find price decreases for several categories that are less often insured and more elective in nature, including psychology (-2%) and chiropractor (-3%) services, though physical therapy services have a 1.6% price increase. Categories that we investigate that have reasonably precise zero price effects include CT scans, gastrointestinal, and eye care. Orthopedic services have a large point estimate (+3%) but a large standard error (2.5%) so we cannot rule out zero nor large effects for that category. While most category-specific quantity effects are fairly precise zeros, there are some notable impacts on radiology procedures (+6%) OB procedures (+4%) and physical therapy procedures (-7%). Since, ex ante, one would expect quantity effects to reflect consumer updating to more precise information, the quantity effects of price information revelation are theoretically ambiguous, even conditional on no price changes. Consequently, when quantities decrease, as with physical therapy, that is consistent with a theory where consumers believed out-of-pocket prices were lower than they actually are, and vice-versa for the

[3] Consumers going out-of-network could pay the entire billed charge as given in our intervention or some reduced version of that billed charge if negotiated down between themselves and the provider or their insurer and the provider. The latter case is relevant for consumers whose insurance covers some portion of out-of-network claims, which is more typical in generous PPO health plans. See, e.g., Bai and Anderson (2016, 2017, 2015, 2018) for an extended discussion of consumer payments when accessing out-of-network providers.

procedures with positive quantity impacts.

Taken together, these results are consistent with our intervention having (i) a minimal effect on consumer price shopping, except for the physical therapy category and (ii) a meaningful effect driving provider price increases, especially for less elective services that are almost always covered by insurance. These price increases are consistent with both tacit collusion between providers in an environment with greater provider-specific price information or with reduced information asymmetries that generally push providers towards realizing that they are under-charging relative to their peers.

Our results should be viewed with a number of caveats. It is important to note that the results of our intervention should be viewed in a short-run context since we measure the effects for two years. Many hypothesized impacts of price transparency relate to systemic, long-run impacts which we do not study here. In addition, our results are specific to the information provision tool provided by FAIR Health and how many providers and consumers use the tool. While website utilization average around 10,000 people per month during our sample period, the impacts of the intervention depend on the type of user. If providers carefully use the site to examine prices, there is potential for significant impacts on prices, while if users are primarily consumers, any impact is more likely to be on the quantity and price shopping dimensions. Finally, our intervention applies to billed charges, which, though relevant for out-of-network claims and potentially relevant via what they signal about negotiated rates, ultimately are a noisy signal of insurer-provider negotiated rates. Given this, it is possible that many providers will not find this information to be very valuable or they may misinterpret its relevance to their patient panels. Despite these potential difficulties our results shed light on important market-level issues related to price transparency that prior studies (who also share some of these difficulties) are not able to address and we are able to do so using a gold standard randomized design.

Relevant prior work in the literature on price transparency has mostly focused on the impact of insurer-provided price transparency tools on consumer price shopping, with equivocal results (Mehrotra et al. (2014)).Robinson and Brown (2013) and Robinson and MacPherson (2012) show that information provision about prices has a meaningful impact on the providers patients choose in the context of a reference-pricing payments model implemented in California where consumers have a lot of money at stake in a context where the potential for price differences is quite salient. Several studies focused on homogeneous services (e.g. lab services and MRIs) find some evidence of price shopping behavior in some populations (Christensen et al. (2017); Robinson et al. (2015); Sinaiko et al. (2016)). There is also evidence that consumers respond to price information in the context of tiered networks. Prager (2020) However, most studies find both relatively low use of price shopping tools by health plan members and, even when accessing such information, little impact on price shopping behavior (Desai et al. (2016); Mehrotra et al. (2014); Sinaiko and Rosenthal

(2016); Brot-Goldberg et al. (2017); Chernew et al. (2018); Cooper et al. (2017)). For consumers utilizing insurer-provided tools, the availability of price information could have large impacts on prices, but these impacts are mitigated by insurance coverage that shields true exposure to prices. Lieber (2017) Since these studies focus on insurer-provided tools, they typically don't address out-of-network price shopping, where consumers typically face larger price differentials and thus may be more responsive to price information.

While there have been quite a few studies on short-run consumer responses to price transparency tools there have been only a few studies that examine provider responses to price transparency. Robinson and Brown (2013) and Wu et al. (2014) both find some evidence that providers lower prices after the introduction of reference pricing / price shipping tools. Both of these studies focused on enrollees in specific-insurance plans. Whaley (2019) exploits the staggered deployment of an online transparency tool to a large pool of insured consumers and finds that robust consumer use of the tool can drive providers to reduce prices for homogeneous services but not for differentiated services. There are two prior studies on market-wide deployment of price transparency tools in New Hampshire. (Desai et al. (2021); Brown (2018)) These studies find that price transparency led to more aggressive bargaining by insurers that had medium-run impacts lowering the prices of high-priced hospitals. To our knowledge, there are no prior papers studying a market-wide deployment of a price transparency tool focused on individual provider prices and certainly none where price transparency information was implemented in a randomized controlled trial together with researchers.

The introduction of the New York Healthcare Online Shopping Tool (NY HOST) offers a unique opportunity to conduct such a trial and systematically and rigorously examine the effects of charge transparency on consumers and providers. The paper proceeds as follows. We provide an overview of the background and setting for the experiment in Section 2. Section 2.2 describes the experimental design and randomization. We discuss the the mechanism by which the information provided by the tool might change the shopping behavior by consumers and price-setting by providers in Section 2.3. Section 3 describes the empirical strategy. Section 4 provides an overview of the datasets utilized and construction of the datasets for analyses. Section 5 provides the results of the empirical tests of the impact of the tool on providers' charges at both the individual provider level and aggregated market level. Section 6 examines the mechanisms behind the results, and Section 7 concludes.

2 Background

2.1 About FAIRHealth

FAIRHealth is an independent non-profit organization that was established in 2009 as a replacement for Ingenix, a database owned by the insurance giant United Healthcare. FAIRHealth maintains the nation's largest data repository of privately billed health insurance claims.[4] Its principal purpose is to provide insurers with an unbiased source of information on usual, customary, and reasonable rates to support the adjudication of out-of-network claims. The FAIR Health database contains claims information from insurers covering approximately 75% of the privately insured population of New York State, including information on both fully-insured claims and claims administered by insurers on behalf of self-insured plans. FAIRHealth had an existing independent, publicly-accessible consumer price transparency tool that displayed aggregate estimates of the charge and insurer allowed amount for a given procedure in each geozip across the country. On September 12, 2017, FAIRHealth re-launched a revamped version of its website for New York State as the New York Healthcare Online Shopping Tool (NY HOST).[5] The rollout was accompanied by an extensive, multi-pronged marketing effort to raise awareness of and draw people to the new consumer facing website. A statewide advertising campaign by FAIR Health was estimated to have reached over 6 million consumers in New York State through traditional media, online advertising, and social media channels. The traditional media campaign included several components. In New York City and Albany, large billboards displayed ads in prominent places, including Times Square. Public service ads (featuring well-known personalities Larry King, Mandy Patinkin, and Nancy Grace) ran in New York City taxicabs; paid advertisements were featured in health clubs and shopping malls throughout the state; and magazine ads were featured in 22 national magazines (e.g. Harper's Bazaar, InStyle, Fortune, Food Network). The distribution of paid print advertising in malls and health clubs was focused on the most highly populated areas of the State, including New York City and Albany.

FAIR Health distributed press releases about the launch to websites with heavy internet traffic, such as Crain's New York and PR Newswire. A Facebook advertising campaign was estimated to have led to over 6 million impressions and reached 1.75 million people from September to July 2018, and Facebook click-throughs accounted for 28% of website hits. FAIR Health also ran a digital banner campaign from September to January 2018, which generated approximately five percent of website hits. FAIR Health also has a continuing active social media presence, with accounts on Facebook and Twitter, and frequent updates featuring its events, services and publications. According to FAIRHealth's analytics, direct searches (such as

[4]Information on FAIRHealth can be accessed via https://www.fairhealth.org.

[5]The consumer tool is publicly accessible on the FAIRHealth website via https://www.youcanplanforthis.org/

users typing the url), which may have been generated by these ongoing activities, generated the remaining website hits. (Kim and Glied (2021))

2.2 Experimental Design: New York Healthcare Online Shopping Tool (NYHOST)

In conjunction with the rollout of NYHOST, a randomized experiment was embedded within the website design of NY HOST. Based on data from FAIR Health, we identified 100 frequently performed procedures for professional services in New York State, spanning 30 different categories. Due to Current Procedural Terminology (CPT) code changes during the 2017 calendar year, another 7 CPT codes were added for a total of 107 procedure codes. This set of categories and procedures were selected because they were both common and had a high rate of out-of-network use.[6] Working with the FAIR Health web development team, we assigned 50 procedures for which specific provider-level charge information was featured in all the 31 FAIR Health constructed 3-digit zipcode (referred here as "geozip") in New York State. In addition to these 50 procedures, specific provider-level charge information also was released for a randomized set of geozip-procedure pairs for the remaining 57 common procedures across all 31 geozips, for a set of 947 procedure-geozip pairs. Each geozip was randomly allocated an additional set of procedures where provider-level charge information was displayed, with a range of 25 to 37 procedures in each geozip, for an average of 31 procedures. Thus, the website featured provider-level charge information, including the range of the billed charges, at the provider level for approximately 81 procedures in each geozip in New York State. The remaining procedure and 3-digit geozip combinations were randomized to the control group and had only aggregated median charge information posted on the website. The experiment ran from September 12, 2017 through August 30, 2019 (Figure 1), and during this time period, provider-level price information was only available on the randomized procedure-geozip pairs.[7]

2.3 How price transparency might affect prices and volume at the provider and market levels

The FAIRHealth price transparency tool reveals providers' billed charges, or the "list prices", in randomized procedure-geozip markets. The tool is available to the general public, including both consumers and providers.

[6]Because of CPT codes that were discontinued in 2017, specifically mammogram codes, we restricted our analysis to a "balanced panel" of procedures, which included the set of 104 procedure codes that were actively billed during the time period examined, CY 2016 through the second quarter of 2019.

[7]The experiment ended on August 2019, and for select procedures, provider-level price information was released in all geozips in New York State.

On the demand side, consumers could benefit from using a price-transparency tool in two contexts; (1) comparing out-of-network charges to the in-network price, and (2) comparing billed charges between out-of-network providers. In the first context, consumers might utilize the tool after being recommended or learning about a specific out-of-network provider. The consumer could access the information afforded by the tool in the decision to visit said specific out-of-network provider and incur the out-of-network price (charge), rather than visiting an in-network provider with the in-network applicable member cost-share. If that out-of-network provider's price exceeds the consumer's reservation price for the services (i.e. the billed charge is "too high" for the consumer), the consumer would remain in-network. Upon seeing the provider level price information, price elastic consumers may move away from the costly out-of-network providers and revert to in-network providers, which would lead to a shift from out-of-network to in-network volume. On the other hand, if the out-of-network price does not exceed the reservation price, despite the presence of lower-cost in-network providers, the consumer may select the recommended out-of-network provider. This model of consumers' choice of provider is consistent with evidence that consumers even in high-deductible plans choose high cost options for MRIs that were recommended by the referring provider, despite nearby cheaper options, highlighting the influence of referring providers on consumer choice (Chernew et al. (2018)). In the second context, consumers may choose to price shop in order to gain information on price information amongst out-of-network providers. Consumers might choose to shop amongst out-of-network providers in certain scenarios, such as inadequate provider networks or an absence of insurance coverage for select services.

Standard models of transparency focus on consumers and implicitly assume that producers already have full access to information, but in healthcare markets, providers are not typically aware of the prices in the market. A large trade literature and an army of consultants in the industry advise doctors on how to set their prices. While providers are aware of their own billed charges and the in-network rates that are offered to them by insurers, they typically lack information on their competitors' billed charges or negotiated rates. Providers might seek to utilize the transparency tool to gain additional information that would inform setting their billed charges. Provider driven changes as a response to the release of information provided by the tool are more likely to be demonstrated in price changes, rather than volume changes.

The change in providers' charges may depend on the degree of competitiveness in provider markets. In highly concentrated provider markets, providers may respond to price information in two different ways. Monopolistically competitive providers may raise their rates when they see competitors charging more. Although these providers may lose volume from consumers with high price elasticity, they may have gains in the form of higher revenues from the inelastic inframarginal patients who select an out-of-network provider. On the other hand, in highly concentrated markets with few specialists, providers may already be aware of their competitors' charges and have a limited response to the release of new price information. In competitive provider markets, providers' change in the billed price may vary based on the price information released by the tool and the anticipated gains from utilizing the tool. If a provider is in a competitive market with price elastic patients, then that provider may seek to lower charges to gain market share. If the provider is in a competitive market with price inelastic patients, then the incentive is to increase charges to maximize profits. Our discussion of the producer response is motivated by the antitrust literature, which states that producers can utilize publicly available information to set anti-competitive prices and circumvent traditional methods of curtailing collusion (Edlin (1997)). Even in industries dominated by online retailers where presumably search costs are close to zero, such as book retailers where Amazon is a major player, persistent price differentials exist (Chevalier and Goolsbee (2003)).

Given the theoretical ambiguity in whether the producer or the consumer price effects dominate, it is necessary to examine the impact of this tool empirically. Our measurement of changes in total volume and providers' prices can reflect the change in the equilibrium price and volume as a result of the price transparency tool. Assuming that volume is primarily driven by demand, we can assume that shifts in volume are primarily due to changes in patient demand for that particular provider. Changes in the billed charge are due to the providers' decision to update their chargemasters based on the information gleaned from the website. We are only able to assess total utilization of the tool and unable to access information on the type of user (whether the user is a patient or a provider) who might be accessing the tool. Thus, we can only deduce whether the consumer or the producer response dominates in the response to the price information by assessing the impact of the tool on the overall change in the charges and volume. The table below provides an overview of the estimated impact of the tool on total volume and the average volume-weighted price (billed charges) at the market level (procedureXgeozip) if the consumer response dominates the result, and if the producer response dominates the aggregate result in the market.

Table 1: Potential impact of the price transparency tool on market-level charges and volume

Price Elasticity	**Initial volume effect**	**Initial charge effect**	**Equilibrium volume at the market-level**	**Equilibrium charge at the market-level**	**Equilibrium out-of-network charge at the market-level**	**Average charge at the provider-level**	**Average out-of-network charge at the provider-level**
Price elastic consumers	Volume shifts to lower price providers.	Providers reduce prices.	Shift to lower price providers.	Lower charges.	Lower charges.	Lower charges.	Lower charges.
Price inelastic consumers	Volume shifts to higher price providers.	Providers raise prices.	Dependent on charge dispersion.	Higher charges.	Higher charges.	Higher charges.	Higher charges.
Monopolistic competitive producers	—	Providers raise prices	Shift to lower price providers.	Higher charges.		Higher charges.	Higher charges.

3 Empirical Strategy

Based on the experimental design laid out in Section 2.2, and the theoretical model laid out in Section 2.3, the econometric specifications utilized are difference-in-differences and our preferred triple difference-in-differences fixed effects models to assess the impact of the charge transparency tool in the randomized markets in the time period after the launch of the tool in September 2017.

3.1 Difference-in-differences specification

For notational simplicity, we first present the difference-in-differences estimation model we use to estimate the treatment effect on the modal billed provider charge:

$$ln(P)_{igpt} = \beta \cdot T_{gp} \cdot Post_t + \lambda_t \cdot \mathbb{1}(yrtri = t) + \gamma_{gp} \cdot \mathbb{1}(gp) + \varepsilon_{igpt}$$

The dependent variable $ln(P)_{igpt}$ refers to the log of the modal billed charge for provider i for procedure p in geozip g in time period t. The difference-in-differences estimate β specifies the treatment effect on the billed charge of provider i rendering services for a treated procedure and geozip combination (T_{gp}) after the launch of the tool. The treatment variable T_{gp} refers to the randomization, and represents a dummy variable equal to one in the procedure and geozip markets where the NYHOST tool revealed provider-level charge information. The time variable $Post_t$ refers to a dummy variable equal to 1 in the time periods after the launch of the tool in September 2017, and 0 before. The model includes controls for time fixed effects $\lambda_t \cdot \mathbb{1}(yrtri = t)$, procedure-geozip fixed effects $\gamma_{gp} \cdot \mathbb{1}(gp)$, and robust standard errors clustered at the provider and procedure-geozip level, ε_{igpt}.

We also estimate these analyses at the market-level. This specification captures volume differences between providers. The market-level regressions are specified at the procedure and geozip level and the difference-in-difference model is specified as follows.

$$ln(P)_{gpt} = \beta \cdot T_{gp} \cdot Post_t + \lambda_t \cdot \mathbb{1}(yrtri = t) + \gamma_{gp} \cdot \mathbb{1}(gp) + \varepsilon_{gpt}$$

The dependent variable $ln(P)_{gpt}$ refers to the log of the average volume-weighted actual billed charge for procedure p in geozip g in time period t. The difference-in-differences estimate β specifies the treatment effect on the market average charge for procedure p in geozip g after the launch of the tool. The treatment variable T_{gp} and time variable $Post_t$ has the same construct as in the provider-level model. The model controls for time and procedure-geozip fixed effects, and robust standard errors clustered at the procedure-geozip level, ε_{gpt}.

3.2 Triple difference-in-differences specification

Our preferred specification is a triple difference-in-differences estimate, which also accounts for time trends in procedure and geozip effects (Berck and Villas-Boas (2016)). Since the volume for each procedure can vary across geozips, the treatment group can vary in size across geozips. Because the randomization was conducted at the procedure and geozip level and the randomization units are not equal in size, geographic and procedure changes over time could bias our results because of compositional effects. The triple difference-in-differences specification removes any potential bias due to geozip or procedure-specific changes over time.

In the experimental design embedded in NYHOST, we randomized multiple procedures among the 3-digit geozips. In effect, we conducted 50 different experiments, with randomization for each experiment across the geozips. If we had analyzed the experiment as 50 different experiments, as is the case in the procedure-specific analyses below, time dummies would be included in each experiment. When all of the procedures are analyzed in a difference-in-differences framework, the coefficients on the time dummies from each of those 50 experiments are effectively constrained to be the same. To the extent that there is heterogeneity in the rate of change of charges among procedures (in the control group) over time, this constraint on the coefficients would fail to capture the degree of heterogeneity. We tested for this by adding interactions by procedure in the combined regression and testing for the joint significance of the interactions; we can reject the null hypothesis that the rate of change in charges across procedures is the same. Alternatively, the construction of the trial can be conceptualized as 31 different experiments, effectively where the geozips are randomized to the procedures. If we analyzed each geozip separately, we would include time dummies in each experiment. Because the zipcodes used in the randomization are of very different sizes (as is the case in comparing the geozip referring to Manhattan to geozips corresponding to rural areas in the state), it is unlikely that we will see the same rate of change across all procedure-geozip combinations. Instead, depending on the allocation of procedures to geozips, the interaction term on the experiment will capture a portion of the heterogeneity in changes in the control group.

Because of these considerations, we utilize a triple difference-in-differences specification that includes the "year by trimester" time variable interacted with the treatment effect to assess the impact over time. Our provider-level triple difference-in-differences economic specification is as follows.

$$ln(P)_{igpt} = \beta \cdot T_{gp} \cdot Post_t + \lambda_t \cdot \mathbb{1}(yrtri = t) + \gamma_{gp} \cdot \mathbb{1}(gp) + \kappa_p \cdot \mathbb{1}(yrtri = t) \cdot \mathbb{1}(p) + \alpha_g \cdot \mathbb{1}(yrtri = t) \cdot \mathbb{1}(g) + \varepsilon_{igpt}$$

The dependent variable $ln(P)_{igpt}$ refers to the log of the billed charge for provider i for procedure p in geozip g in time period t. The triple difference-in-differences estimate β specifies the treatment effect of the treatment dummy T_{gp} (equal to 1 for the randomized procedure and geozips), interacted with $Post_{gp}$ (equal to 1 for the trimesters encompassing the period after September 2017). The model includes controls for time fixed effects $\lambda_t \cdot \mathbb{1}(yrtri = t)$, procedure-geozip fixed effects $\gamma_{gp} \cdot \mathbb{1}(gp)$, procedure dummy variables interacted with the time dummy variables $\kappa_p \cdot \mathbb{1}(yrtri = t) \cdot \mathbb{1}(p)$, the geozip dummy variables interacted with the time dummy variables $\alpha_g \cdot \mathbb{1}(yrtri = t) \cdot \mathbb{1}(g)$, and robust standard errors clustered at the provider and procedure-geozip level, ε_{igpt}.

The market-level triple difference-in-differences economic specification is the same as the provider-level regressions, except with the procedure X geozip over time as the unit of observation.

$$ln(P)_{gpt} = \beta \cdot T_{gp} \cdot Post_t + \lambda_t \cdot \mathbb{1}(yrtri = t) + \gamma_{gp} \cdot \mathbb{1}(gp) + \kappa_p \cdot \mathbb{1}(yrtri = t) \cdot \mathbb{1}(p) + \alpha_g \cdot \mathbb{1}(yrtri = t) \cdot \mathbb{1}(g) + \varepsilon_{gpt}$$

The dependent variable $ln(P)_{gpt}$ refers to the log of the average volume-weighted billed charge for procedure p in geozip g in time period t. The triple difference-in-differences estimate β specifies the treatment effect of T_{gp} interacted with $Post_t$. The model includes the same set of controls, of time and procedure-geozip fixed effects separately and interacted with the time dummy variables, and robust standard errors clustered at the procedure-geozip level, ε_{gpt}.

We also generated event study graphs for which we estimated the treatment effect for each trimester. The econometric specification for the event study graphs with the triple difference-in-differences model includes a set of coefficients β_t, which reflects the treatment effect for each trimester.

$$ln(P)_{igpt} = \beta_t \cdot T_{gp} \cdot \mathbb{1}(yrtri = t) + \lambda_t \cdot \mathbb{1}(yrtri = t) + \gamma_{gp} \cdot \mathbb{1}(gp) + \kappa_p \cdot \mathbb{1}(yrtri = t) \cdot \mathbb{1}(p) + \alpha_g \cdot \mathbb{1}(yrtri = t) \cdot \mathbb{1}(g) + \varepsilon_{igpt}$$

This specification includes the treatment effect specified for each trimester, with each trimester dummy variable indicating if the observation falls within that trimester. The dependent variable $ln(P)_{igpt}$ refers to the log of the modal billed charge for provider i for procedure p in geozip g in time period t. The triple difference-in-differences estimate β_t specifies the treatment effect of the treatment dummy T_{gp} for each trimester, where the the time variable is included as a set of dummy variables, one for each trimester ($\mathbb{1}(yrtri = t)$). The controls in the models are the same as described previously.

4 Data

Our analyses draws on the nearly-comprehensive dataset of claims included in the FAIRHealth database. The data extract includes the entirety of the FAIRHealth database for claims in New York State zipcodes with dates of services between January 1, 2016 to June 30, 2019, a total of over 110 million claims (110,422,511 claims in total) with dates of service within this time period. The fields in these data include the National Provider Identifier (NPI), the FAIRHealth constructed 3-digit zipcode ("geozip"), date of service, procedure code (Current Procedural Terminology), billed charge, place of service code, the patient's gender, and the patient's age group.

We matched NPI data to the CMS Physician Compare and the National Plan and Provider Enumeration System (NPPES) files to obtain information on provider characteristics. There were 205,023 unique NPIs in the FAIRHealth data extract. We linked the provider NPIs represented in the FAIRHealth data to the information in the 2017 CMS Physician Compare Downloadable File in order to access provider-level information, including gender, years in practice, medical school, group size, and hospital affiliation. The 2017 CMS Physician Compare File contains information on the providers who are participating in the CMS quality program, which encompasses all eligible providers (EPs) that qualify or participate in the program.[8] Since providers were able to be credentialed in multiple specialties and practice in several different locations, the specialty and location was chosen as the first that appeared when sorted in alphabetical order. We also utilized U.S. Census Data to access information on population and to construct a urban/rural indicator for each geozip.

Since the experiment went into effect in September 2017, which was two-thirds of the way into the calendar year, we defined our time periods at the trimester level, or a third of a year, with four months in each trimester. The post randomization period spanned from September 2017 through June 2019 period, defined as the third trimester of 2017 through the second trimester of 2019. Since we had claims data from January 2016 through June 2019, we had incomplete data for the last trimester, defined as May 2019 through August 2019. We restricted our final analyses to the 10 trimesters for which we had complete claims data, which spanned the time period from January 2016 through April 2019, with comprehensive claims data for 5 trimesters in the pre-randomization period (from January 2016 through August 2017), and 5 trimesters in

[8]Table A.1 shows the procedure categories that were included for the study. Table A.2 shows the datasets utilized and time periods encompassed. Table A.3 shows the match between the FAIRHealth dataset and the CMS Physician Compare dataset. Approximately 58% of all of the provider NPIs in our FAIR Health data extract were captured in the 2017 CMS Physician Compare file.

the post-randomization period (from September 2017 through April 2019).

From the FAIRHealth claims data, we constructed an unbalanced panel dataset at the provider X geozip X procedure level across trimesters, which encompassed 3,598,866 observations. We computed the modal charge reported by each provider for each procedure in each trimester of the period studied. Given potential billing errors, we utilized the modal charge to capture the most frequently billed unit charge for a given procedure for each trimester as the list price for that provider. As robustness checks, we included the median and percentiles (95th percentile, 5th percentile) of the billed charge for each provider in a given trimester, procedure, and geozip. To account for outliers and billing errors (since several claims had billed charges that ranged as high as $70,000), we winsorized the provider-level panel dataset at the 95th percentile.

To account for the substantial provider churn in our dataset, we removed providers with fewer than five claims for a procedure in a given geozip and trimester. Next, to create a "balanced" panel of providers that would enable us to follow the same cohort of providers over time, we restricted the cohort to providers with at least one claim in each trimester, for a total of ten trimesters. We also removed providers with the credential identifying them as a physicians assistant or a nurse practitioner, since these practitioners often submit separate claims for procedures and services they may have rendered supporting services for, rather than as the primary provider. We also removed any CPT codes that were added or discontinued throughout the study period to generate a "balanced panel" of procedures for which charges were posted across all the trimesters included in our study period, and thus trackable over time.[9]

To assess the overall market impact of the tool, we created a panel dataset at the procedure and geozip level for each trimester, and constructed the aggregate volume-weighted charge and total volume for a given procedure and geozip in each time period. Because of the wide dispersion of charges even for the same procedure, we examined the logged values of the price and the quantity for both the market level regressions and the provider level regressions.

5 Results

5.1 Experimental Design

FAIR Health was created in 2009, and the organization created a consumer website in 2011 that displayed educational information but no price information. Consumer cost lookup was added in 2011 with market-level price data available. The website was resdesigned and relaunched in September 2017 with an extensive

[9] The codes were removed from the balanced panel because they were either discontinued or added to the CPT® (Current Procedural Terminology) list between 2016 and 2019. Those codes were 76641, 76642, 77052, 77056, 77065, 77066, 97161, 97162, and 97163.

marketing effort to raise public awareness of the tool (Figure 1). In New York State, consumers who accessed the FAIRHealth consumer website (youcanplanforthis.org) were able to access provider-level price information, which was embedded in the randomized procedures and geozip pairs. Figure 2 shows the aggregate price information that was displayed in the non-randomized procedure and geozip pairs. Figure 3 shows the consumer view of the tool when accessing price information on a procedure in a given geozip that was randomized to the treatment group. The consumer-facing website with the embedded randomization was released to the public on September 2017, and the randomization remained in place until August 2019, at which point the provider-level charge information for all of the selected procedures was released across the state in all geozips. After the completion of the experiment, FAIR Health continued to revise the website and has continued plans to expand the information displayed.

After the randomization was conducted and prior to the launch of the tool, balance checks were conducted to validate the randomization process, and compare market characteristics between the treatment and control groups, including measures of price, volume, and market concentration based on the claims with dates of service in 2016, the baseline period prior to the launch of the tool (Table 2). There was no statistically significant difference between the treatment and control groups on aggregate charges, the interquartile range of charges, volume of claims, volume out-of-network claims, insurer market concentration, and population density. Although most of the market characteristics were comparable, the control group had more concentrated provider markets at baseline (provider HHI of 860 in the control group compared to provider HHI of 732 in the treatment group), and the treatment group had slightly higher within market charge dispersion, defined as the standard devision of charges in a given procedure-geozip market (0.09 in the treatment group compared to 0.08 in the control group), and slightly higher charge dispersion at the 90th quantile (1.10 in the treatment group compared to the 1.09 in the control group). The "first stage" analysis assesses utilization of the price transparency tool prior to and after the launch of the revamped tool in September 2017. Figure 4 demonstrates the distribution of the utilization of the website, measured as the total number of searches, over years and months.[10] Web utilization of the tool by consumers appears to have been relatively consistent over time and low compared to the New York State overall population of over 19 million residents, with fewer than 15 searches on average for each procedure in a given month. (Kim and Glied (2021))

[10] Figure A.4 depicts the distribution of website utilization across the treatment and control groups by month between January 2016 and June 2019. The persistently higher number of searches in the control group can be attributed to the number of searches that occurred in New York City for the procedures in that geozip that were assigned to the control group. When we exclude the 3-digit geozip that corresponds to the borough of Manhattan (geozip = 100), the number of searches between the treatment and control groups is more comparable (Figure A.5). Because we randomized 57 different procedures across 31 geozips to the treatment and control groups, and the zipcodes are of different sizes in population density, random assignment to a highly populated zipcode can lead to a higher number of searches (Figure A.6).

5.2 Summary Statistics

We first assessed the market characteristics of charges to understand the underlying correlates of the level and dispersion of charges and the process of price adjustment. We assessed the market characteristics for the baseline period, which was defined as the calendar year 2016, to capture a year's worth of claims experience prior to the launch of the tool. We find that normalized modal charges are quite dispersed, with over half of providers' charges greater than the average charge for a given market, and with a substantial right skew even when truncating to normalized charges below 10 (Figure 5). The subset of procedures that were selected for the experiment had a relatively high percentage of out-of-network claims; although most procedures were performed out-of-network less than 10% of the time, some procedures have an out-of-network percentage that ranged as high as 40 percent (Figure 6). Figure 7 demonstrates the market concentration of providers for each procedure and geozip market. The provider market concentration for each procedure and geozip was calculated using the Herfindahl-Hirschman Index (HHI), with the market share for each provider represented in the dataset by the National Provider Identification number (NPI). Prior research suggests that concentration of provider markets is associated with higher charges (Roberts et al. (2017)), and the variation in the provider market HHI demonstrates substantial heterogeneity in provider market concentration and corresponding price dispersion.

Our balanced panel dataset for the provider-level models included all providers who had a minimum of 5 claims in each of the ten trimesters, for a total of 583,693 observations at the provider, geozip, procedure, and trimester level, for the set of 947 procedure-geozip combinations that were randomized. Each provider, at a given procedure and geozip, had an average of 25 claims each trimester, with an average billed charge of $420. We examined a variety of procedures, ranging from lower cost psychotherapy and physical therapy services to higher cost orthopedic and radiology services, and there is substantial heterogeneity in the billed charge. The billed charge ranged from $2 to $59,000, with a standard deviation of $1,279. There was also variation in the volume of services rendered, with a range in the volume of claims rendered for each from 1.5 to 5,582, with a standard devision of 50 claims (Table 3). [11]

At the market and trimester level, with the market defined for each procedure and geozip, there was significant underlying price heterogeneity, with an average charge of $910 with a standard deviation of $1,811 (Table 3). In our sample of procedures, there was an average of 1,489 claims rendered in a given procedure and geozip market in a single trimester, and of those claims, approximately 20% were rendered

[11] The billed charge for a given procedure varies widely across providers, and even for the same provider, the billed charge can vary over time since providers can update their chargemasters at will. There significant dispersion in the distribution of the charge updates (presented as the change in the log of the price in Figure A.7). Approximately 15% of providers updated their charges in the first trimester of 2017 and 2018, but providers continue to update their charges later in a given calendar year (Figure A.8). This underlying price heterogeneity points to meaningful scope for price changes.

by a provider who was out-of-network with a given insurance product.[12]

5.3 Provider-level outcomes

We applied both the difference-in-differences and triple difference-in-differences specification to assess the overall treatment effect on providers' charges across all procedures and geozips (Table 4). In our primary triple difference-in-differences specification, we find that, across all procedures and locations, providing individual-level provider charge information increases prices by 1.2% and has no statistically significant impact on quantity. In our event study, we also find a consistent positive difference in the modal price between the treatment and control groups in the post-randomization period (Figure 8). The difference-in-differences specification generated a null treatment effect on the modal billed charge. We find no meaningful quantity effects in general, in aggregate and for various forms of heterogeneity.

5.4 Tests for heterogeneity

Our main results show the impact of the tool in aggregate, and may mask heterogeneity in the effect on price and quantity. Our tests for heterogeneity stratified both the provider-level sample by provider and procedural characteristics. We assessed the impact of the tool separately for providers whose prices were initially above or below the median for a given procedure in their geozip and find modest but larger increases in prices for providers who were initially below the median. In addition, we find no quantity impacts for providers who were initially high-priced as opposed to low-priced, suggesting that our intervention had no meaningful impact on the extent of consumer price shopping. We also find that providers who are in procedure markets that are above median market concentration, measured with procedure-geozip HHI, have slightly larger price increases than those below median, with no statistically significant quantity differences (Table 5).

Given that our intervention provides information on billed charges, rather than insurer-contracted prices, we focus especially on out-of-network claims, for which billed charges are relevant.We find that procedures with a high proportion of out-of-network claims have essentially no price change while procedures with a low proportion of out-of-network claims have a 2.5% price increase as a result of our intervention. We find a similar pattern in the event study graphs when focused on the price trends for procedures with a low proportion of out-of-network claims when stratified by high and low price (defined as above or below median charge) providers (Figure 9). This suggests that, for procedures where billed charges are much closer to final

[12] We also calculated the insurer HHI as the sum of the square of the market share for each insurer within each each procedure and geozip market. Over 30 insurance companies in NYS represented in the claims data, with each insurer represented in the data by a FAIR Health "key" that kept the identity of each insurer confidential. The distribution of insurance market HHI shows that most markets are highly concentrated, and for most of the procedures in our sample, the insurer HHI was well above 4000 (Figure A.9).

prices, there is no impact of our intervention, while for procedures that are more likely to be covered by insurance, prices increase.

We also investigate the effects of our intervention for specific procedure categories. We find larger price increases for specific categories that are almost always insured and less elective in nature, including MRI (+6%) and radiology (+3%) services. We find price decreases for several categories that are less often insured and more elective in nature, including psychology (-2%) and chiropractor (-3%) services, though physical therapy services have a 1.6% price increase. Categories that we investigate that have reasonably precise zero price effects include CT scans, gastrointestinal, and eye care. Orthopedic services have a large point estimate (+3%) but a large standard error (2.5%) so we cannot rule out zero nor large effects for that category. While most category-specific quantity effects are fairly precise zeros, there are some notable impacts on radiology procedures (+6%) OB procedures (+4%) and physical therapy procedures (-7%) (Figure 6).

5.5 Market-level outcomes

Table A.11 shows the results from the market-level specifications across the balanced panel of providers and procedures. We find that the overall price effect is similar to the provider-level models, with no significant market level effects on overall volume. Although there are fluctuations in the coefficient of variation before the randomization went into effect, the decrease in the coefficient of variation is sustained in the post-randomization period with both the difference in differences (Figure A.12) and triple difference-in-differences specification (Figure A.13), suggesting a slight decrease in price dispersion in the treatment group. To test for heterogeneity in our market-level outcomes, we stratified the market-level dataset upon dimensions of market concentration, procedures with high vs. low out-of-network use, coefficient of variation, website utilization. We also tested for differential impact on services that typically require continuity of care and deemed "continuous" (such as psychotherapy and physical therapy), compared to services that do not require continuity of care and are "non-continuous" (e.g. radiology, orthopedic procedures). (Table A.14) Similar to the provider-level results, our most significant result is that in procedure markets with low out-of-network claims at baseline, there is a 2.9% price increase. We also investigated the effects of our intervention for specific procedure categories, and find a market-level price increase for MRI services (+8%) (Table A.15).

6 Discussion

Our findings support the hypothesis that the provider-response dominates in the overall response to this tool. Overall, provider-level prices and aggregated market prices increased more in the randomized procedureXgeozip markets than the non-randomized markets. Our findings that providers with a lower percentage of their services rendered out of network were more likely to raise their charges in the post-randomization period suggests that the tool yielded useful information for providers with limited out-of-network experience. The providers who were already rendering a larger proportion of their services out-of-network may already have been aware of their competitors' charges or had set their charges optimally. For providers with limited out-of-network experience, the presence of the tool enabled them to see the charge information posted by their competitors in a given market and increase their charges accordingly.

Overall, these results are consistent with our intervention having (i) a minimal effect on consumer price shopping, except for the physical therapy category and (ii) a meaningful effect driving provider price increases, especially for less elective services that are almost always covered by insurance. These price increases are consistent with both tacit collusion between providers in an environment with greater provider-specific price information or with reduced information asymmetries that generally push providers towards realizing that they are under-charging relative to their peers.

7 Conclusions

Although price transparency is a laudable goal in a healthcare market dominated by information asymmetries, there may be perverse price effects due to supply constraints, the inelastic nature of the demand for healthcare services and opportunity for providers to engage in price-setting. Our results suggest caution about price transparency if physicians more likely to leverage that information than consumers to set prices. Our results should be viewed with a number of caveats, including the limited time window during which we can study the effects of the intervention, and the application to billed charges, which, though relevant for out-of-network claims and potentially relevant via what they signal about negotiated rates, are not indicative of insurer-provider negotiated rates. Despite these potential difficulties, our results shed light on important market-level issues related to price transparency using a gold standard randomized design.

References

Albaek, S., Mollgaard, P., and Overgaard, P. B. (1997). Government-assisted oligopoly coordination? a concrete case. *The Journal of Industrial Economics*, 45(4):429–443.

Bai, G. and Anderson, G. F. (2015). Extreme markup: The fifty us hospitals with the highest charge-to-cost ratios. *Health affairs*, 34(6):922–928. doi: 10.1377/hlthaff.2014.1414; 26.

Bai, G. and Anderson, G. F. (2016). Us hospitals are still using chargemaster markups to maximize revenues. *Health affairs*, 35(9):1658–1664.

Bai, G. and Anderson, G. F. (2017). Variation in the ratio of physician charges to medicare payments by specialty and region. *JAMA*, 317(3):315–318. 10.1001/jama.2016.16230.

Bai, G. and Anderson, G. F. (2018). Market power: Price variation among commercial insurers for hospital services. *Health Affairs*, 37(10):1615–1622.

Batty, M. and Ippolito, B. (2017). Mystery of the chargemaster: examining the role of hospital list prices in what patients actually pay. *Health affairs*, 36(4):689–696.

Berck, P. and Villas-Boas, S. B. (2016). A note on the triple difference in economic models. *Applied Economics Letters*, 23(4):239–242.

Brot-Goldberg, Z., Chandra, A., Handel, B. R., and Kolstad, J. T. (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, Z. Y. (2018). Equilibrium effects of health care price information. *The review of economics and statistics*. doi: 10.1162/rest$_{a0}$0765$pmid:$.

Chernew, M., Cooper, Z., Larsen-Hallock, E., and Morton, F. S. (2018). No title. *Are health care services shoppable? Evidence from the consumption of lower-limb MRI scans.*

Chevalier, J. and Goolsbee, A. (2003). Measuring prices and price competition online: Amazon.com and barnesandnoble.com. *Quantitative marketing and Economics*, 1(2):203–222.

Christensen, H., Floyd, E., and Maffett, M. (2017). The effects of charge-price transparency regulation on prices in the healthcare industry.

Cooper, Z., Morton, F. S., and Shekita, N. (2017). Surprise! out-of-network billing for emergency care in the united states. *National Bureau of Economic Research Working Paper Series*, No. 23623.

Cutler, D. M. and Dafny, L. (2011). Designing transparency systems for medical care prices. *New England Journal of Medicine*.

Desai, S., Hatfield, L., Hicks, A., Chernew, M., and Mehrotra, A. (2016). Association between availability of a price transparency tool and outpatient spending. *JAMA*, 315(17):1874–1881.

Desai, S. M., Shambhu, S., and Mehrotra, A. (2021). Online advertising increased new hampshire residentsâ use of provider price tool but not use of lower-price providers. *Health affairs*, 40(3):521–528.

Edlin, A. S. (1997). Do guaranteed-low-price policies guarantee high prices, and can antitrust rise to the challenge? *Harvard law review*, pages 528–575.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., and Group, O. H. S. (2012). The oregon health insurance experiment: evidence from the first year. *The Quarterly journal of economics*, 127(3):1057–1106.

Glied, S. (2021). Price transparency promise and peril. In *JAMA Health Forum*, volume 2, page e210316. American Medical Association.

Kim, G. and Glied, S. (2021). Assessing utilization of a market-wide price transparency tool. *American Journal of Managed Care*.

Kullgren, J. T. and Fendrick, A. M. (2021). The price will be right. how to help patients and providers benefit from the new cms transparency rule. In *JAMA Health Forum*, volume 2, page e210102. American Medical Association.

Lieber, E. M. J. (2017). Does it pay to know prices in health care? *American Economic Journal: Economic Policy*, 9(1):154–79.

Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., Leibowitz, A., and Marquis, M. S. (1987). Health insurance and the demand for medical care: evidence from a randomized experiment. *American Economic Review*, 77(3):251–277.

Mehrotra, A., Brannen, T., and Sinaiko, A. D. (2014). Use patterns of a state health care price transparency web site: What do patients shop for? *INQUIRY*, 51:0046958014561496. doi: 10.1177/0046958014561496; 27.

Newhouse, J. P. (1996). Reimbursing health plans and health providers: Efficiency in production versus selection. *Journal of Economic Literature*, 34(3):1236–1263. Cited By :347.

Prager, E. (2020). Healthcare demand under simple prices: Evidence from tiered hospital networks. *American Economic Journal: Applied Economics*, 12(4):196–223.

Reinhardt, U. E. (2006). The pricing of u.s. hospital services: Chaos behind a veil of secrecy. *Health Affairs*, 25(1):57–69.

Roberts, E. T., Chernew, M. E., and McWilliams, J. M. (2017). Market share matters: Evidence of insurer and provider bargaining over prices. *Health affairs*, 36(1):141–148.

Robinson, J. C., Brown, T., and Whaley, C. (2015). Reference-based benefit design changes consumersâ choices and employersâ payments for ambulatory surgery. *Health Affairs*, 34(3):415–422.

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

Robinson, J. C. and MacPherson, K. (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.

Sinaiko, A. D., Joynt, K. E., and Rosenthal, M. B. (2016). Association between viewing health care price information and choice of health care facility. *JAMA Internal Medicine*, 176(12):1868–1870.

Sinaiko, A. D. and Rosenthal, M. B. (2016). Examining a health care price transparency tool: Who uses it, and how they shop for care. *Health affairs*, 35(4):662–670.

Volpp, K. G. (2016). Price transparency: Not a panacea for high health care costs. *JAMA*, 315(17):1842–1843.

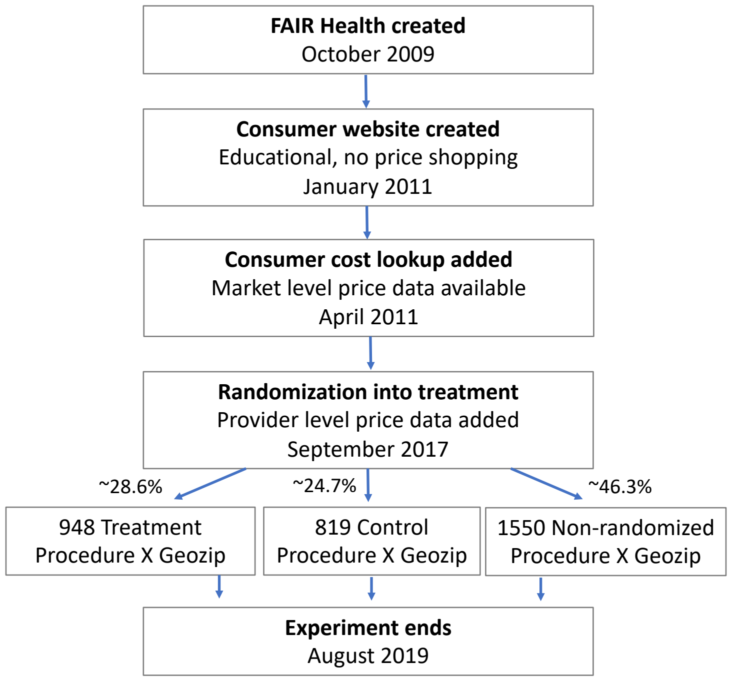
Whaley, C. M. (2019). Provider responses to online price transparency. *Journal of health economics*, 66:241–259.

Wilensky, G. (2019). Federal government increases focus on price transparency. *Jama*, 322(10):916–917.

Wu, S.-J., Sylwestrzak, G., Shah, C., and DeVries, A. (2014). Price transparency for mris increased use of less costly providers and triggered provider competition. *Health Affairs*, 33(8):1391–1398.

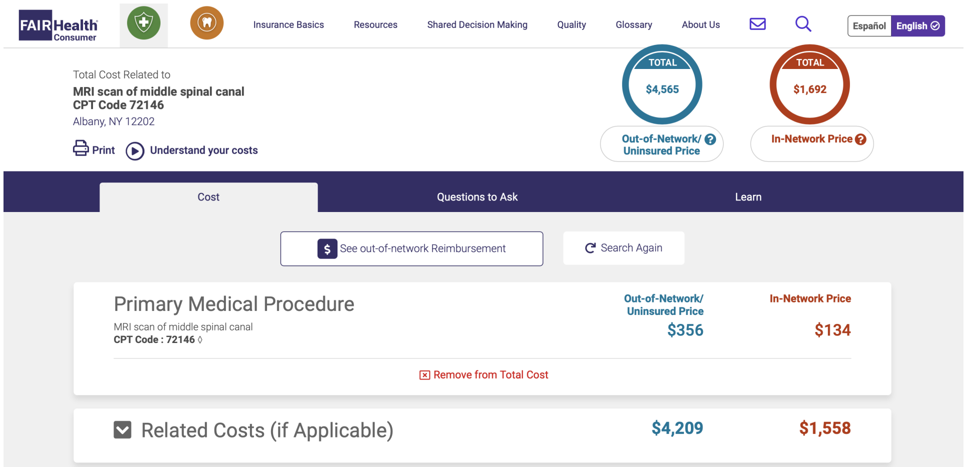
]Tables and Figures

Figure 1: Timeline and Experimental Design



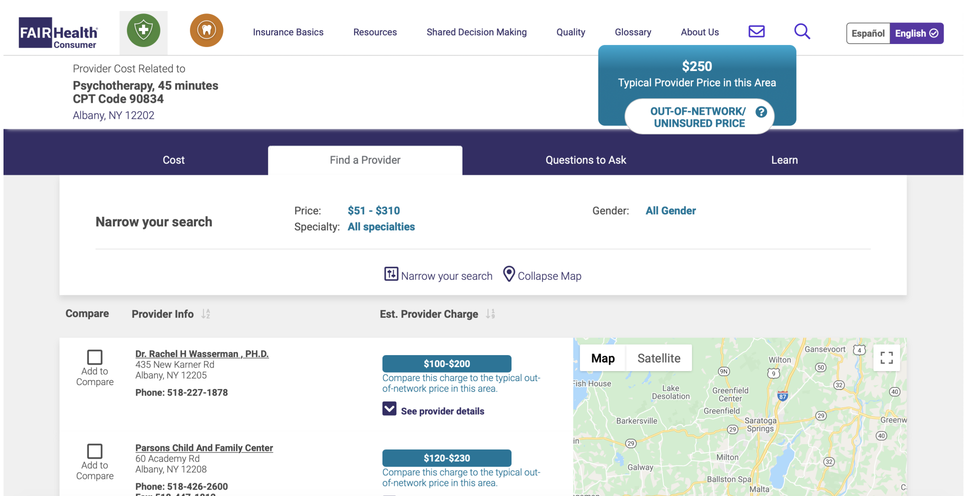
Notes: This shows the timeline of the implementation and randomization of the website.

Figure 2: NYHOST Website: Control Website Search Result



Notes: A snapshot of the FAIRHealth consumer shopping tool for a control procedure and geozip that does not contain provider-level information.

Figure 3: NYHOST Website: Treatment Website Search Result



Notes: A snapshot of the FAIRHealth consumer shopping tool for one of the randomized procedure and geozips that released provider-level price information.

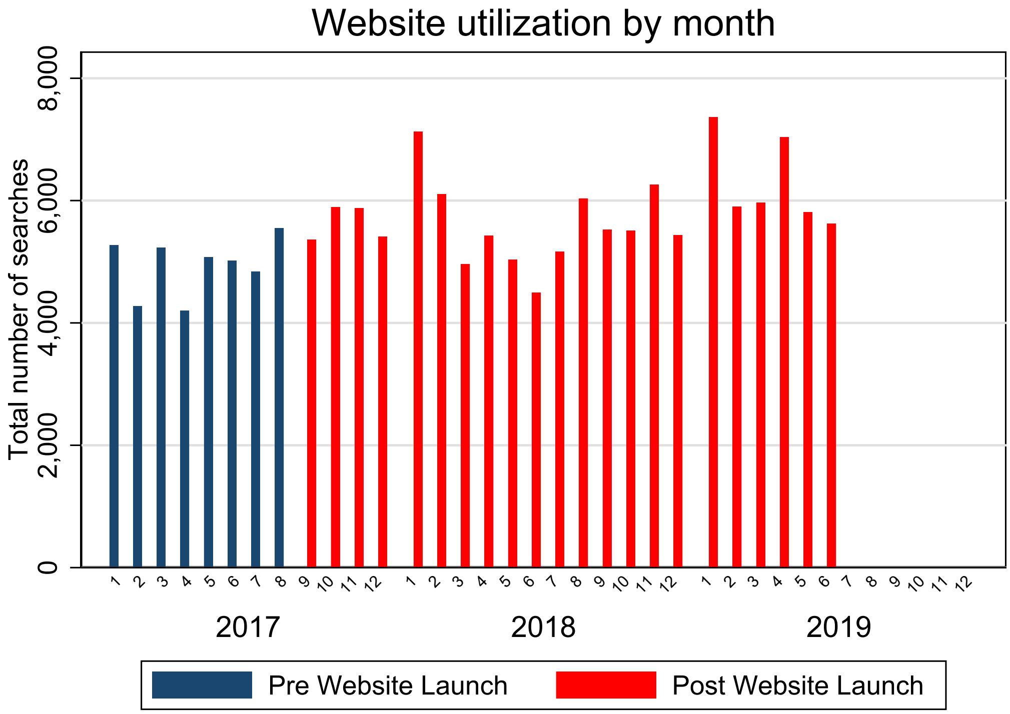
Tables and Figures

Table 2: Balance Check

Variable	(1) Treatment Mean/SD	(2) Control Mean/SD	T-test P-value (1)-(2)
Avg. charge	1,066.18	973.14	0.41
	(2,217.37)	(1,967.47)	
Avg. IQR charge	545.90	440.22	0.10
	(1,430.75)	(926.16)	
Avg. volume	1,049.87	1,215.78	0.55
	(4,557.15)	(5,409.81)	
Avg. volume OON	141.91	219.80	0.24
	(801.85)	(1,474.76)	
Within market charge disp. S.D.	0.09	0.08	0.03**
	(0.05)	(0.05)	
10th quantile of charge disp.	0.89	0.90	0.30
	(0.08)	(0.08)	
90th quantile of charge disp.	1.10	1.09	0.00***
	(0.07)	(0.06)	
HHI Insurers (2017 pre NYHOST)	4,680.94	4,639.07	0.61
	(1,533.54)	(1,469.85)	
HHI Providers (2017 pre NYHOST)	732.16	859.92	0.02**
	(861.70)	(1,099.24)	
Population density	7,215.97	7,354.20	0.87
	(15277.83)	(15033.83)	
N	734	622	
Clusters	734	622	

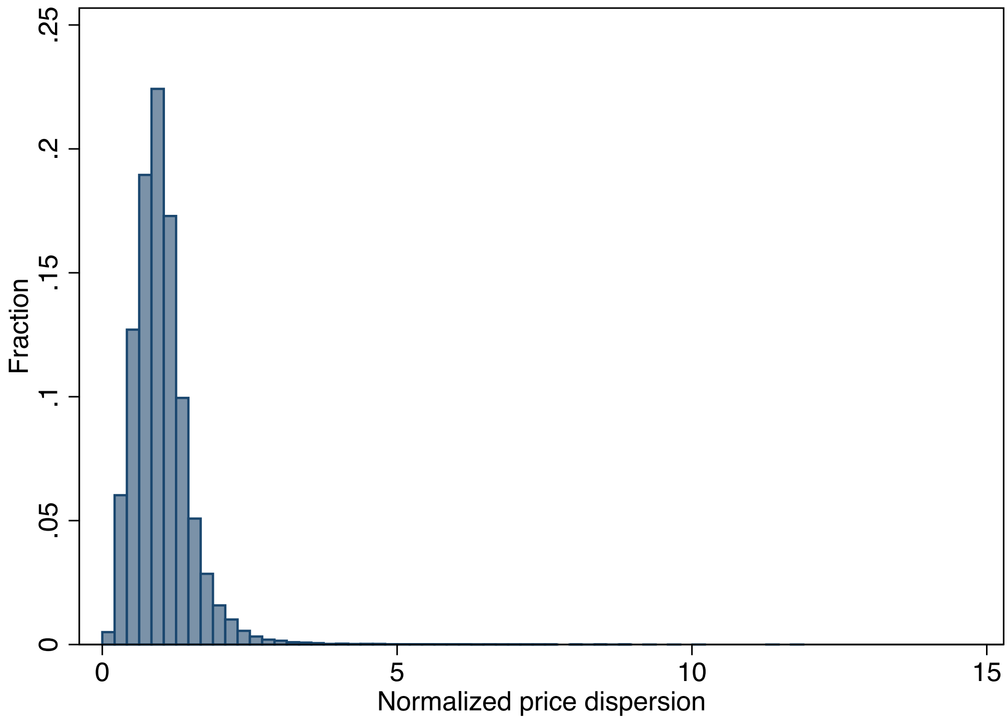
Data source: FAIRHealth.
Note: This table checks for balance of market level summary stats in the pre-period for treatment and control markets.

Figure 4: Website utilization by month



Notes: Data on the NYHOST web searches provided by FAIRHealth. Note: This figure plots total number of monthly searches across the pre- and post- period. Large spikes in 2016 and 2019 are associated with major website overhauls and marketing changes.The experiment was in effect from 9/2017 through 6/2019.

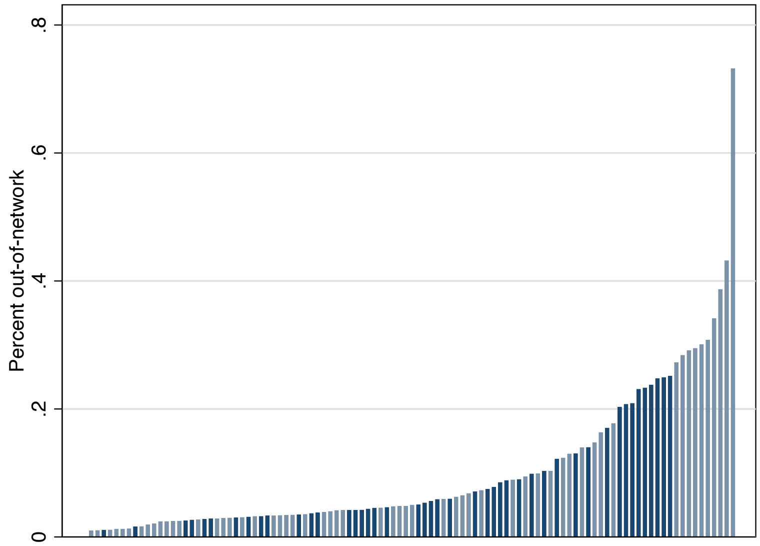
Figure 5: Normalized Charge Dispersion



,

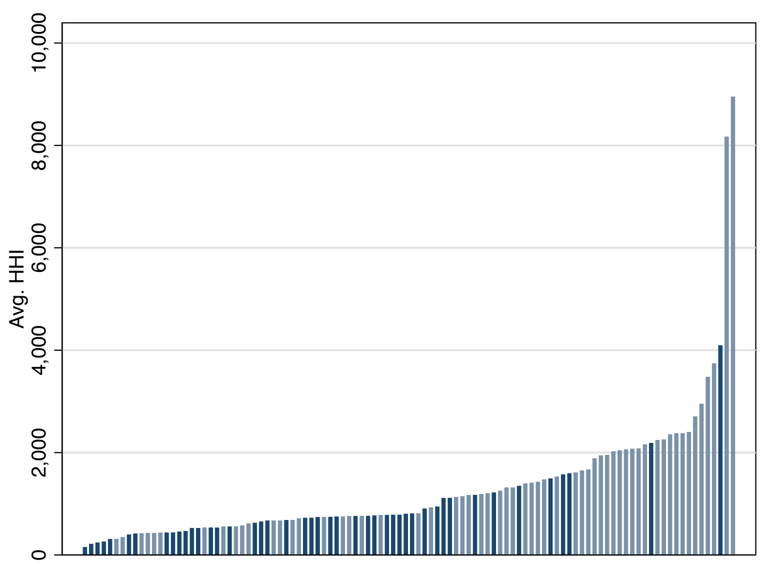
Data Source: FAIRHealth.
Notes: This graph presents the histogram of normalized modal charge dispersion by trimester. Normalized modal charge is calculated as the ratio of a provider's modal charge to the procedureXgeozip mean. The histogram is restricted to normalized charges below 10.

Figure 6: Percentage of Claims Out-of-Network, by Procedure Code



Notes: The graph represents the distribution of the percentage of out-of-network claims for procedure codes ordered lowest to highest. Dark blue bars indicate procedure codes used in the final analysis sample.

Figure 7: Physician Market Concentration



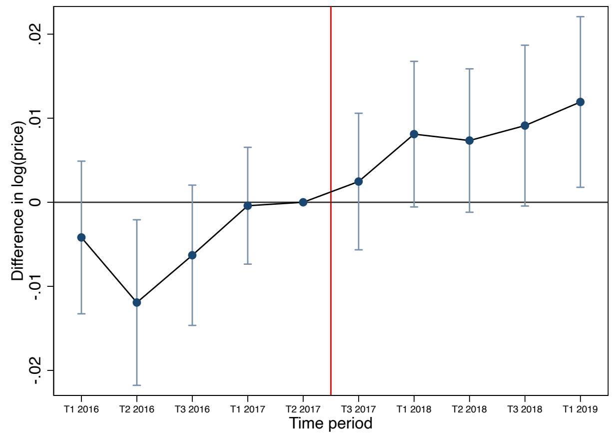
Notes: This figure represents the distribution of physician market concentration by procedure code in 2016. Provider market concentration was constructed as the HHI for each procedure and geozip, with the market share constructed as each provider, defined by the NPI.

Table 3: Provider and market summary statistics

	Mean	SD	Min	Max
Panel A: Provider summary statistics				
Volume	25.68	49.58	1.50	5,582.25
Unit Charge	420.43	1,278.58	1.78	59,000.00
N	583,693			
Panel B: Market summary statistics				
Avg. charge	910.70	1,811.49	9.00	29,792.60
IQR charge	735.27	1,644.23	0.00	25,186.25
Avg. volume	1,489.14	7,526.44	0.25	184229.75
Avg. volume OON	267.06	1,983.71	0.00	53,152.75
HHI Insurers	4,636.15	1,541.07	2,059.27	10,000.00
HHI Providers	869.18	1,137.99	4.70	10,000.00
Population Density	7,233.21	15,105.59	51.54	73,138.80
N	17,442			

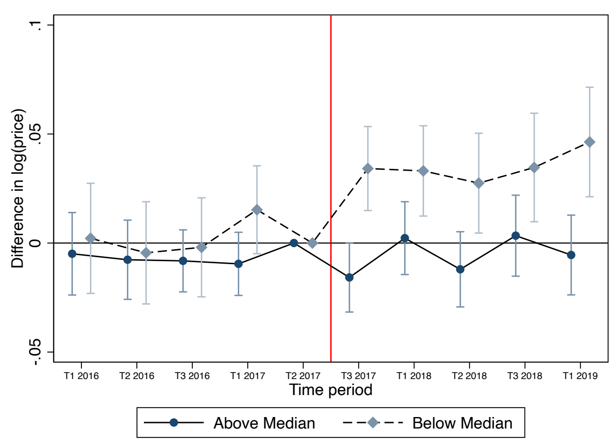
Data source: FAIRHealth.
Notes: This table shows the summary statistics for charge and volume information for the provider-level dataset at the NPI X procedure X geozip X trimester level as well as market level summary statistics of charges, volume and demographic factors.

Figure 8: Event Study: Difference between Treatment and Control with triple difference-in-differences specification



Data source: FAIRHealth.
Note: This figure plots coefficients from a regression of log(price) on an interaction between treatment and trimester and fixed effects for time (trimester-year), market (procedure X geozip), procedure X trimester and geozip X trimester. Treatment began at the start of trimester 3 in 2017. This event study corresponds to a triple difference regression. Standard errors are clustered at the procedure X geozip level.

Figure 9: Event study: Difference between Treatment and Control with triple differences-in-differences specification, for low out-of-network procedures



Data source: FAIRHealth.
Note: This figure plots coefficients from a regression of log(price) on an interaction between treatment and trimester and fixed effects for time (trimester-year), market (procedure X geozip), procedure X trimester and geozip X trimester for low OON procedures across above and below median price providers. Treatment began at the start of trimester 3 in 2017. This event study corresponds to a triple difference regression. Standard errors are clustered at the procedure X geozip level.

Table 4: Provider-level regressions: treatment effect of NYHOST

	log(Price)		log(Quantity)	
	(1)	(2)	(3)	(4)
	DiD	Triple Diff	DiD	Triple Diff
Treatment effect	0.0049	0.0123***	-0.0141	-0.0050
	(0.0044)	(0.0035)	(0.0101)	(0.0064)
Observations	583680	583680	583680	583680
Adjusted R^2	0.831	0.831	0.378	0.378
ProcedureXGeozip FE	Yes	Yes	Yes	Yes
Trimester FE	Yes	Yes	Yes	Yes
ProcedureXTime FE		Yes		Yes
GeozipXTime FE		Yes		Yes

Standard errors in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Data source: FAIRHealth.
Notes: This table contains coefficients from a regression of log(price) on an interaction between treatment and a post indicator with fixed effects for time (trimester-year) and market (procedure X geozip), corresposnding to a difference in difference regression or time (trimester-year), market (procedure X geozip), procedure X time and geozip X time corresponding to a difference in difference in difference regression. Treatment began at the start of trimester 3 in 2017. Standard errors are clustered at the procedure X geozip level.

Table 5: Provider level regressions: Heterogeneity tests for the treatment effect of NYHOST

	Provider Prices		Continuous Procedures		OON Procedures		Market HHI		Website Use	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	$>$Median	$<$Median	Non-Continuous	Continuous	$>$Median	$<$Median	$>$Median	$<$Median	$>$Median	$<$Median
Panel A: Charge treatment effects										
log(P) effect	0.0017	0.0087*	0.0199***	0.0099**	0.0079*	0.0234***	0.0174*	0.0106***	0.0162***	0.0149***
	(0.0033)	(0.0049)	(0.0077)	(0.0046)	(0.0044)	(0.0065)	(0.0089)	(0.0039)	(0.0051)	(0.0047)
Observations	313330	243076	156697	333755	407556	176124	87206	496470	294282	289377
Adjusted R^2	0.927	0.892	0.831	0.597	0.729	0.821	0.840	0.821	0.833	0.829
ProcedureXGeozip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ProcedureXTime FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GeozipXTime FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
* $p < .10$, ** $p < .05$, *** $p < .01$										
Panel B: Quantity treatment effects										
log(q) effect	-0.0032	0.0009	0.0149*	-0.0367***	-0.0135	0.0004	0.0188	-0.0063	-0.0344**	0.0022
	(0.0077)	(0.0084)	(0.0084)	(0.0127)	(0.0089)	(0.0087)	(0.0151)	(0.0073)	(0.0137)	(0.0076)
Observations	313330	243076	156697	333755	407556	176124	87206	496470	294282	289377
Adjusted R^2	0.419	0.391	0.183	0.268	0.331	0.274	0.332	0.371	0.336	0.397
ProcedureXGeozip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ProcedureXTime FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GeozipXTime FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Data source: FAIRHealth.

Notes: This table contains coefficients from a regression of log(price) on an interaction between treatment and a post indicator with fixed effects for time (trimester-year), market (procedure X geozip), procedure X time and geozip X time corresponding to a difference in difference in difference regression testing for heterogeneity. Treatment began at the start of trimester 3 in 2017. Standard errors are clustered at the procedure X geozip level.

Table 6: Provider level regressions: Treatment effect of NYHOST by procedure category

	(1) CT	(2) MRI	(3) RAD	(4) GI	(5) EYE	(6) ORTHO	(7) OB	(8) PSYCH	(9) PTOT	(10) CHIRO
Panel A: Charge treatment effects										
log(P) effect	-0.00173	0.06170***	0.02650*	-0.01831	-0.00298	0.03122	-0.01499	-0.01864***	0.01600*	-0.02449**
	(0.01868)	(0.01751)	(0.01590)	(0.01677)	(0.00812)	(0.02539)	(0.02400)	(0.00678)	(0.00824)	(0.01025)
Observations	33929	33464	55541	27321	65907	20418	13345	129138	159301	45316
Adjusted R^2	0.413	0.188	0.322	0.397	0.629	0.875	0.934	0.174	0.229	0.392
ProcXZip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
* $p < .10$, ** $p < .05$, *** $p < .01$										
Panel B: Quantity treatment effects										
log(Q) effect	0.01852	0.01496	0.05838***	0.02368	0.00889	-0.00380	0.04167*	0.00035	-0.06887***	-0.02006
	(0.02315)	(0.01877)	(0.01696)	(0.02376)	(0.01610)	(0.02658)	(0.02432)	(0.01539)	(0.02121)	(0.03044)
Observations	33929	33464	55541	27321	65907	20418	13345	129138	159301	45316
Adjusted R^2	0.213	0.140	0.125	0.222	0.195	0.248	0.244	0.379	0.085	0.167
ProcXZip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Data source: FAIRHealth.

Notes: This table contains coefficients from a regression of log(price) and log(quantity) on an interaction between treatment and a post indicator with fixed effects for time (trimester-year) and market (procedure X geozip) corresponding to a difference in difference regression for each procedure category. Treatment began at the start of trimester 3 in 2017. Standard errors are clustered at the procedure X geozip level.

A Appendix

A.1 Categories of service for procedures examined

Category	Randomized	Number of Procedures
Acupuncture	N	4
Allergy	N	4
Bone Density	N	1
Cardiology	N	2
Chemotherapy	N	1
Chiropractic	Y	4
CAT Scan (Radiology)	Y	4
Dermatology	N	4
ENT	N	5
Gastroenterology	Y	3
Infusion	N	1
Mammogram	Y	8
MRI	Y	5
Neuromuscular	N	1
Obstetrics & Gynecology	Y	4
Ophthalmology	Y	4
Orthopaedic	Y	6
Pain	N	5
Plastic Surgery	N	1
Psychotherapy	Y	5
Physical Therapy/Occupational Therapy	Y	8
Pulmonology	N	3
Radiology	Y	6
Sleep Medicine	N	2
Spine	N	4
Surgery	N	3
Urology	N	1
Ultrasound	N	4
Ultrasound-OB	N	2
Vascular Radiology	N	2

Notes: These categories were selected on the basis of encompassing procedures that were commonly serviced and non-emergent.

A.2 Description of the datasets utilized

Dataset	Description
FAIRHealth	2016-2019
FAIRHealth NYHOST Website data	2016-2019
CMS Physician Compare	2017
CMS National Plan and Provider Enumeration System (NPPES)	2017

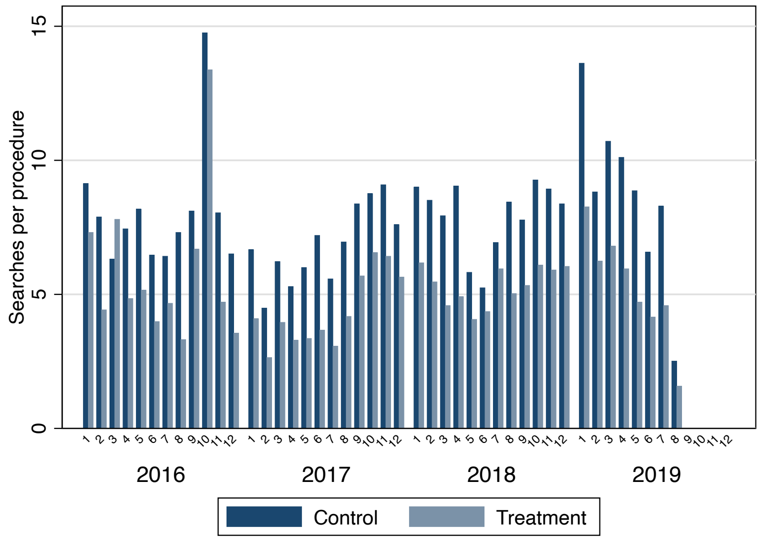
Notes: These datasets were utilized to conduct the analyses of the impact of the NYHOST price transparency tool. The CMS Physician Compare file utilized is the most recent dataset that was able to be accessed on 2/20/2020.

A.3 Data on provider characteristics

Dataset	# Distinct NPI in FAIRHealth Data	# Distinct NPI in CMS Compare (2017)
Total number of providers in each dataset.	205,258	1,142,428
Providers in both FAIRHealth + CMS Physician Compare	119,583	119,583
# Distinct NPI in FAIRHealth Data and CMS Physician Compare, with specialties associated with the MD/DO credential.	78,509	78,509
Number of providers represented in the balanced panel (subset of the total number of providers)	21,601	14,146

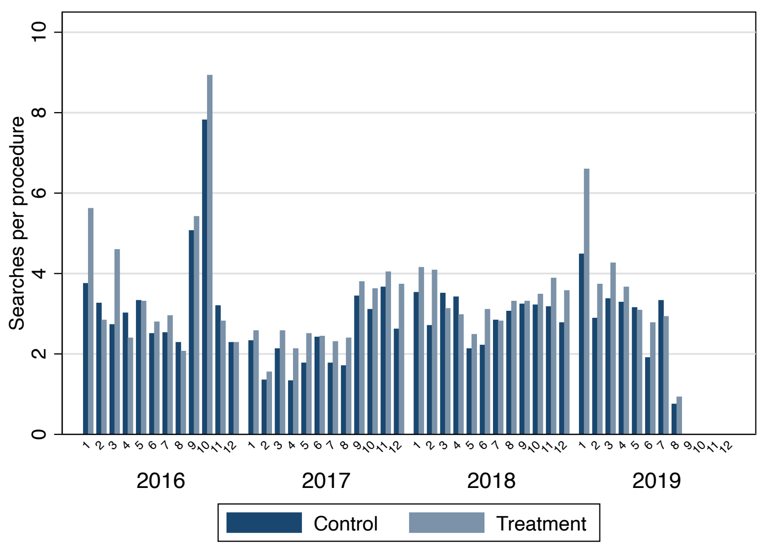
Notes: This table demonstrates the number of providers represented in the FAIR Health dataset, and the match of the NPIs in the FAIRHealth dataset with the 2017 CMS Physician Compare File. The CMS Physician Compare File includes information on providers, including credentials, medical school, gender, and affiliated hospitals.

A.4 Website Usage for Procedures and Geozips in Experiment



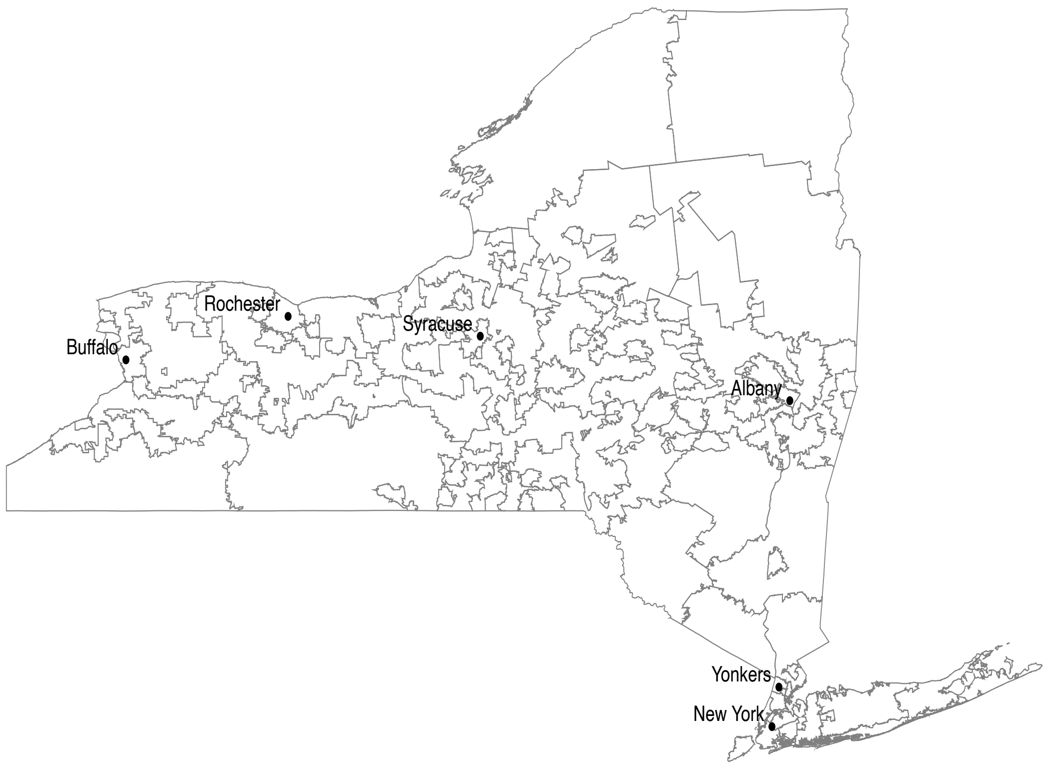
Notes: This figure plots the average monthly website utilization (data on NYHOST web searches provided by FAIRHealth) for procedures in treatment and control groups. The persistently higher searches in the control group can be attributed to the number of searches that occurred in New York City for the procedureXgeozip combinations in the control group.

A.5 Website utilization by month excluding geozip 100 (Manhattan)



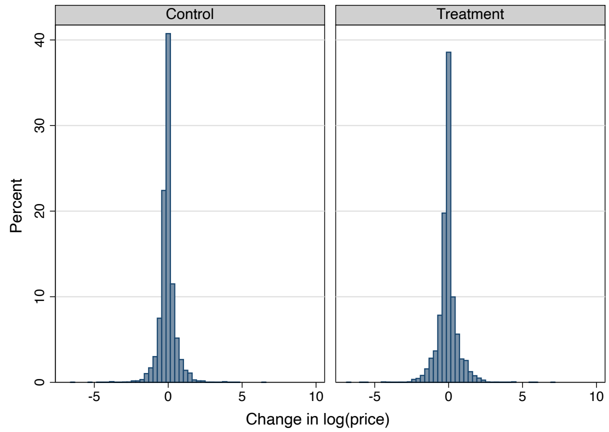
Data source: FAIRHealth.
Notes: This figure plots the average monthly website utilization (data on NYHOST web searches provided by FAIRHealth) for procedures in treatment and control groups when the geozip corresponding to Manhattan (geozip = 100) is excluded.

A.6 Map of New York Geozips



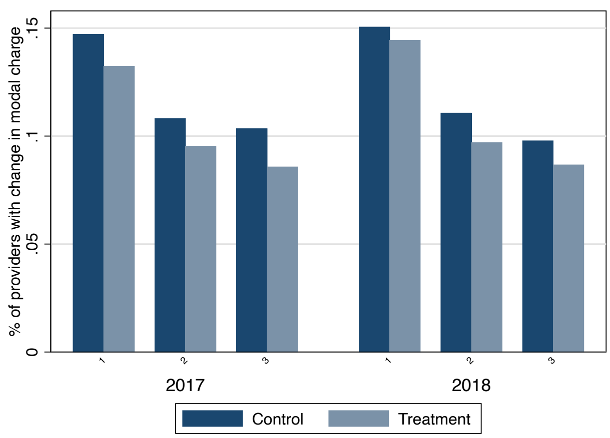
Notes: Map of New York State geozips, with select cities indicated.

A.7 Changes in log(Price) in Treatment and Control Groups



Note: This figure presents histograms of charge updates for providers in the treatment and control groups.

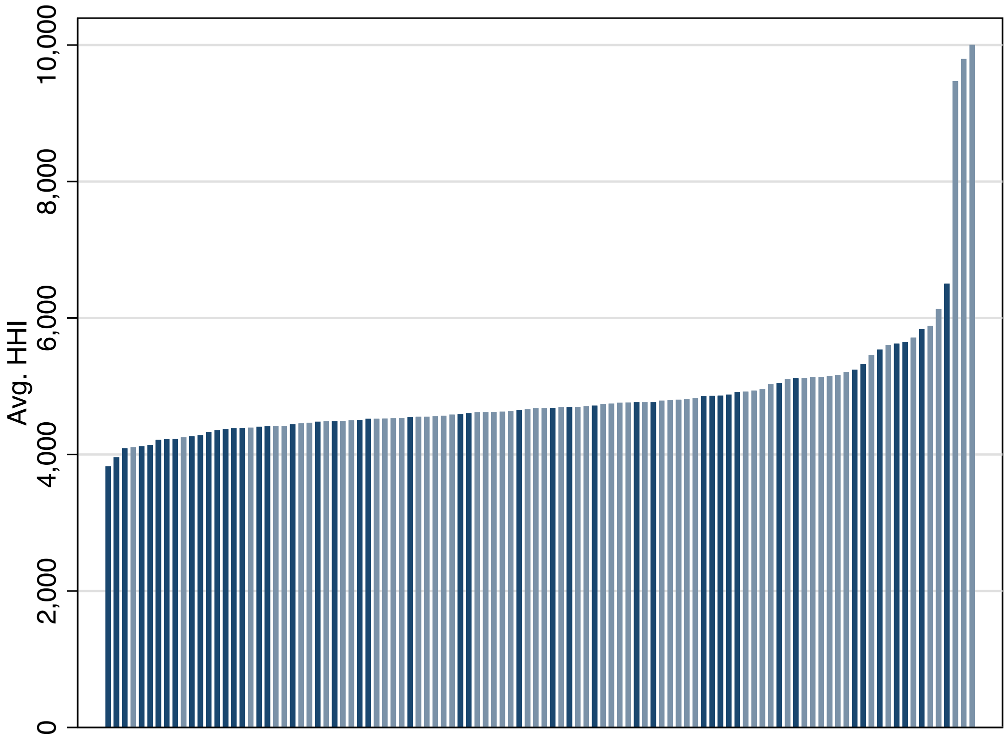
A.8 Changes in log(Price) over Time (2017-2018)



Data source: FAIRHealth.

Note: This figure presents histograms of charge updates for providers in the treatment and control groups. This displays the percentage of providers who update their charges each trimester.

A.9 Insurer Market Concentration



Notes: Insurer market concentration in 2016, by procedure code, with each insurer identified by a FAIR Health "key".

A.10 Provider-level results: Robustness test of the treatment effect on the percentile of the providers' charges

	(1)	(2)	(3)	(4)	(5)	(6)
	95th	95th	50th	50th	5th	5th
Treat*Post	0.003	0.012***	0.003	0.012***	0.007	0.016***
	(0.006)	(0.004)	(0.006)	(0.004)	(0.006)	(0.004)
Constant	5.016***	5.013***	4.948***	4.945***	4.839***	4.837***
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Observations	510815	510815	510815	510815	510815	510815
Adjusted R^2	0.848	0.848	0.846	0.846	0.815	0.816
ProcedureXGeozip Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Post	Yes	Yes	Yes	Yes	Yes	Yes
ProcedureXPost Dummies		Yes		Yes		Yes
GeozipXPost Dummies		Yes		Yes		Yes

Standard errors in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Data source: FAIRHealth.
Notes: Fixed effects models with procedureXgeozip and time fixed effects. The DD estimates demonstrate the overall impact of NY HOST on the log of the percentile of a provider's charge (5th, 50th, and 95th percentile) each trimester as the outcome variable. The time fixed effects are measured the "Post" dummy, signifying the time period after the experiment went into effect. The triple DD specification include postXprocedure dummy variables and postXgeozip dummy variables.

A.11 Market-level regressions: Treatment effect of NYHOST

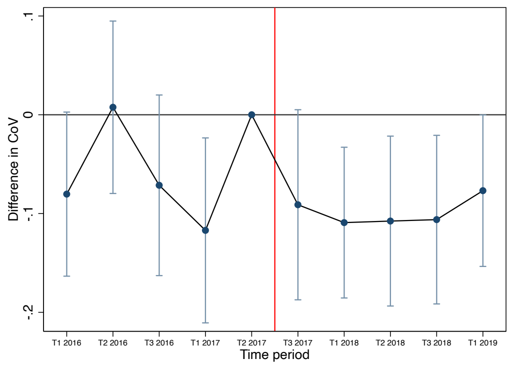
	log(Price)		log(Quantity)	
	(1)	(2)	(3)	(4)
	DiD	Triple Diff	DiD	Triple Diff
Treatment effect	0.0127**	0.0176***	-0.0111	0.0149
	(0.0063)	(0.0055)	(0.0272)	(0.0131)
Observations	15871	15864	15871	15864
Adjusted R^2	0.994	0.995	0.940	0.987
ProcedureXGeozip FE	Yes	Yes	Yes	Yes
Trimester FE	Yes	Yes	Yes	Yes
ProcedureXTime FE		Yes		Yes
GeozipXTime FE		Yes		Yes

Standard errors in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

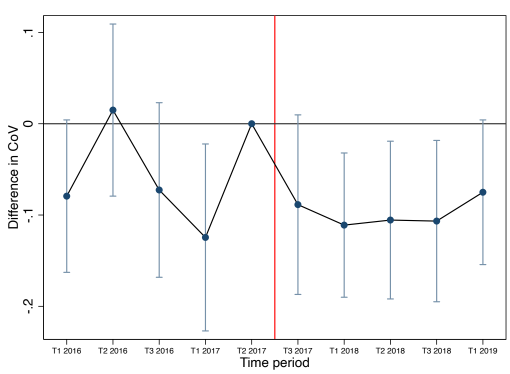
Data source: FAIRHealth.
Notes: This table contains coefficients from a regression of log(price) on an interaction between treatment and a post indicator with fixed effects for time (trimester-year) and market (procedure X geozip), corresposnding to a difference in difference regression or time (trimester-year), market (procedure X geozip), procedure X time and geozip X time corresponding to a difference in difference in difference regression. Treatment began at the start of trimester 3 in 2017. Standard errors are clustered at the procedure X geozip level.

A.12 Market-level regressions with difference-in-differences effect of NYHOST on market price coefficient of variation



Data source: FAIRHealth.
Note: This event study graph is generated from market-level regressions with the difference-in-differences specification.This figure plots coefficients from a regression of market charge coefficient of variation on an interaction between treatment and trimester and fixed effects for time (trimester-year) and market (procedure X geozip). Treatment began at the start of trimester 3 in 2017. This event study corresponds to a difference in difference regression. Standard errors are clustered at the procedure X geozip level.

A.13 Market-level regressions with difference-in-differences-in-differences effect of NYHOST on market price coefficient of variation



Data source: FAIRHealth.
Note: This event study graph is generated from market-level regressions with the difference-in-differences-in-differences (Triple DD) specification. This figures plots coefficients from a regression of market charge coefficient of variation on an interaction between treatment and trimester and fixed effects for time (trimester-year), market (procedure X geozip), procedure X trimester and geozip X trimester. Treatment began at the start of trimester 3 in 2017. This event study corresponds to a triple difference regression. Standard errors are clustered at the procedure X geozip level.

A.14 Market-level regressions: Heterogeneity tests for the treatment effect

	Coefficient of Variation		Continuous Procedures		OON Procedures		Market HHI		Website Use	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	>Median	<Median	Non-Continuous	Continuous	>Median	<Median	>Median	<Median	>Median	<Median
Panel A: Charge treatment effects										
log(P) effect	0.0164*	0.0189***	0.0120	-0.0046	0.0083	0.0286***	0.0196**	0.0141**	0.0148**	0.0282**
	(0.0088)	(0.0069)	(0.0089)	(0.0088)	(0.0072)	(0.0083)	(0.0097)	(0.0058)	(0.0060)	(0.0121)
Observations	8045	7801	7653	4750	7827	8037	7922	7922	13406	2159
Adjusted R^2	0.994	0.996	0.995	0.984	0.996	0.993	0.993	0.998	0.995	0.997
ProcedureXGeozip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ProcedureXTime FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GeozipXTime FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
* $p < .10$, ** $p < .05$, *** $p < .01$										
Panel B: Quantity treatment effects										
log(Q) effect	0.0201	0.0183	0.0612***	-0.0506	-0.0213	0.0600***	0.0194	0.0086	0.0148	-0.0159
	(0.0154)	(0.0226)	(0.0171)	(0.0488)	(0.0219)	(0.0170)	(0.0270)	(0.0088)	(0.0151)	(0.0183)
Observations	8045	7801	7653	4750	7827	8037	7922	7922	13406	2159
Adjusted R^2	0.990	0.985	0.984	0.988	0.989	0.981	0.969	0.996	0.985	0.994
ProcedureXGeozip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ProcedureXTime FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
GeozipXTime FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Data source: FAIRHealth.
Notes: This table contains coefficients from a regression of charges (constructed here as the log(price)) on an interaction between treatment and a post indicator with fixed effects for time (trimester-year), market (procedure X geozip), procedure X time and geozip X time corresponding to a difference in difference in difference regression testing for heterogeneity. Treatment began at the start of trimester 3 in 2017. Standard errors are clustered at the procedure X geozip level.

A.15 Market-level regressions: Treatment effect of NYHOST by category

	(1) CT	(2) MRI	(3) RAD	(4) GI	(5) EYE	(6) ORTHO	(7) OB	(8) PSYCH	(9) PTOT	(10) CHIRO
Panel A: Charge treatment effects										
log(P) effect	-0.01550	0.08102***	-0.02828	-0.00601	-0.00679	0.02856	0.00558	-0.00928	-0.00500	0.02081
	(0.02119)	(0.02502)	(0.02057)	(0.02529)	(0.01567)	(0.02107)	(0.01726)	(0.01009)	(0.01224)	(0.01465)
Observations	1240	1550	1766	930	1240	1858	1239	1549	1985	1219
Adjusted R^2	0.955	0.887	0.936	0.965	0.997	0.984	0.995	0.939	0.980	0.937
ProcXZip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
* $p < .10$, ** $p < .05$, *** $p < .01$										
Panel B: Quantity treatment effects										
log(Q) effect	0.06304*	0.08516***	-0.21411	-0.00588	0.06716**	-0.04532	0.11407**	0.01925	-0.03083	-0.09618
	(0.03725)	(0.02756)	(0.13690)	(0.02968)	(0.03316)	(0.03495)	(0.05126)	(0.03232)	(0.05479)	(0.09246)
Observations	1240	1550	1766	930	1240	1858	1239	1549	1985	1219
Adjusted R^2	0.980	0.974	0.613	0.989	0.982	0.977	0.968	0.990	0.965	0.970
ProcXZip FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimester FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Data source: FAIRHealth.

Notes: This table presents the regression results when stratifying the procedures by category. This table contains coefficients from a regression of log(price) and log(quantity) on an interaction between treatment and a post indicator with fixed effects for time (trimester-year) and market (procedure X geozip) corresponding to a difference in difference regression for each procedure category. Treatment began at the start of trimester 3 in 2017. Standard errors are clustered at the procedure X geozip level.