Transparency and Negotiated Prices:
The Value of Information in Hospital-Supplier Bargaining

Matthew Grennan[*]

University of Pennsylvania, The Wharton School & NBER

Ashley Swanson

University of Pennsylvania, The Wharton School & NBER

June 19, 2019

Abstract

Using data on hospitals' purchases across a large number of important product categories, we find that access to information on purchasing by peer hospitals leads to reductions in the prices hospitals negotiate for supplies. These effects are concentrated among hospitals previously paying relatively high prices for brands purchased in large volumes. Evidence from coronary stents suggests that transparency allows hospitals to resolve asymmetric information problems, but savings are limited in part by the stickiness of contracts in business-to-business settings. Savings are largest for physician preference items, where high-price, high-quantity hospital-brand combinations average 3.9 percent savings, versus 1.6 percent for commodities.

1 Introduction

Business-to-business markets make up a large part of the economy, but they often lack transparency. Suppliers negotiate different contracts with different buyers, potentially with widely varying prices, and a buyer typically has limited information regarding other buyers' contracts. Many business-to-business markets have seen the entry of information intermediaries who facilitate buyers' ability to benchmark the prices they negotiate.[1] In this paper,

[1]This trend is due in part to technology improvements that have made data easier to collect, distribute, and analyze. In addition to the hospital purchasing context, we are aware of business-to-business "transparency" benchmarking services emerging in areas as diverse as home appliances and television advertising.

we empirically examine the effect of transparency in the form of benchmarking information on negotiated prices between hospitals and their suppliers.

There is substantial variation in the prices of hospital supplies, including medical devices, across hospitals. For the top hospital supplies in our data, the average coefficient of variation across hospitals for the same exact brand-month is 0.18. This input price variation is approximately half the coefficient of variation for common procedure prices charged by hospitals in different markets (Cooper et al. 2018). It is also near the top of the range of coefficients of variation found in consumer goods markets (Scholten and Smith 2002).[2] In the short run, these costs come directly from hospitals' profit margins.[3] In the longer run, increasing supply costs tend to increase health care costs (see, e.g., Maeda et al. 2012).

Prior research in consumer goods markets has largely confirmed the intuition that information can facilitate search and better decisions for buyers with imperfect information regarding product quality or cost (Sorensen 2000; Jin and Leslie 2003; Hendricks et al. 2012; Bronnenberg et al. 2015) or supplier willingness to accept lower prices (Zettelmeyer et al. 2006; Scott Morton et al. 2011). However, the mechanisms via which information impacts consumer goods may not extend to business-to-business markets where there is often no search across sellers (when products are purchased directly from manufacturers), and negotiators are professionals employed by firms and thus with different expertise and incentives than the typical consumer.[4] Recent empirical research in business-to-business bargaining (Draganksa et al. 2009; Crawford and Yurukoglu 2012; Grennan 2013, 2014; Gowrisankaran et al. 2015; Lewis and Pflum 2015; Ho and Lee 2017) explains variation in prices across buyers using full information models, but in doing so also documents substantial heterogeneity in *bargaining ability* parameters, which could include variation in information available to negotiators.[5] Our work contributes to these literatures by extending our understanding of transparency to the business-to-business setting and by offering information as one explana-

---

[2] See also Kaplan and Menzio (2015) regarding price variation for consumer goods across stores.

[3] According to the American Hospital Association 2018 Trendbook, the average hospital operating margin in 1995-2016 was 4.4 percent (https://www.aha.org/system/files/2018-05/2018-chartbook-table-4-1.pdf). Craig et al. (2018) calculate the supplies in our data represent 26 percent of hospital operating costs.

[4] The theoretical and empirical literatures on information disclosure are large and reviewed in Dranove and Jin (2010). Recent studies from a variety of contexts including consumer healthcare spending (Lieber 2017) and choice of college major (Hastings et al. 2015) have found the effects of information to be increasing in prior uncertainty but attenuated by frictions in consumers' ability and incentives to put information to use. In our context of professionals, one might expect prior uncertainty to be lower but ability and incentives to use information to be higher than in the consumer context.

[5] There are two exceptions of which we are aware. Larsen (2018) estimates a bargaining game of two-sided incomplete information about valuations in the used car wholesale market; we argue that uncertainty over bargaining parameters better fits our context. Backus et al. (2016) study cheap talk in bargaining with asymmetric information for collectibles on eBay, but in our case price seems to dominate all other concerns for negotiators, decreasing the scope for the trade-offs involved in signaling that they document.

tion for the large unexplained heterogeneity documented in negotiated prices.

Our empirical analysis utilizes new data containing all purchase orders issued by more than 17 percent of US hospitals that subscribed to a web-based benchmarking database. Section 2 details the data and hospital purchasing context. The majority of our analysis focuses on price negotiations for coronary stents in 508 facilities with cardiac catheterization services.[6] We also estimate our main specification using data on 52 product categories that are important in terms of total spend or utilization.

Motivated by discussions with industry experts, Section 3 outlines two candidate mechanisms through which benchmarking information might have an impact in the hospital purchasing context: (1) an asymmetric information model in which hospitals face uncertainty about suppliers' costs or bargaining parameters, so that transparency reduces uncertainty and the equilibrium dispersion in negotiated prices; and (2) an agency model in which price transparency allows hospital managers to better observe purchasing agents' effort and, in turn, provide improved incentives to purchasing agents to reduce prices.

Section 4 clarifies our research design for testing empirical predictions from these models. We use two sources of variation to provide plausibly causal identification. First, the database is generated by monthly submissions of member hospitals' transactions, and new members are asked when they join to submit twelve months of retrospective, pre-benchmarking data. We use variation in timing of hospitals' joining the database to construct difference-in-differences estimators. Exogeneity of join timing with respect to price trends is supported by the institutional details of the setting, and by event studies that show no statistically significant divergence of pre-trends for coronary stents.

Second, we develop a set of tests focusing on *new brands* entering the market during our sample period. We compare prices between hospitals pre- and post-join immediately upon a brand's introduction, before either hospital type has access to information, in order to remove any persistent sources of bias around join timing. New brand introductions also offer a strategy to investigate theoretical mechanisms. The asymmetric information mechanism wherein hospitals use benchmarking information to learn about suppliers relies upon concurrent availability of data on others' prices, but the agency mechanism wherein hospitals

---

[6]Stents are especially useful as a place to focus for several reasons: They are representative of the "physician preference items" (PPIs) central to many policy discussions – PPIs are products where doctor usage decisions are relatively insensitive to price, making negotiating lower prices the main mechanism via which a hospital can obtain savings. Stents are also important in their own right, comprising two percent of sample hospital spend. Finally, stents typically have simple linear contracts, so the price we observe is the contracted price. As discussed in Appendices C and F, we observe no evidence of standardization (e.g., exclusive dealing or contracts based on market share) affecting prices or usage in our data, and we find no effect of benchmarking information on quantities. This motivates our focus on the effects of information on linear prices (rather than quantities) in the remainder of the paper.

use benchmarking information to better incentivize their purchasing negotiators relies on the fact that such information will be available in the future. Thus, the delayed release of benchmarking data after new brand entry allows us to separate the two.

Section 5 presents our main results. Focusing on stents, we find that access to the database information has heterogeneous effects across hospitals and brands. The average treatment effect of benchmarking information across all hospital-brand-months is small and noisy – we estimate no significant price changes for hospital-brands with low to moderate prices pre-join – but high-price hospital-brand combinations exhibit unit price declines of $27 upon accessing database information. The price declines are larger for hospital-brands with larger purchase volumes: $71 for high-quantity hospital-brands that are also high-price prior to joining the database, compared to $17 for hospital-brands with lower volumes. The specifications leveraging brand entry suggest that these effects are largely explained by a mechanism wherein benchmarking solves an asymmetric information problem, helping hospitals learn about suppliers. Evidence for the agency mechanism is noisier and less robust across specifications.

Each of the above-described results is an estimate of the treatment effect of benchmarking on *prices*, which will be an underestimate of the treatment effect of benchmarking on prices negotiated *in a given contract*. We find price effects are generated by increasing the likelihood of renegotiation and by generating larger price decreases conditional on renegotiation.

Taken together, our results suggest that transparency can lead to significant savings, but that these savings are dampened by stickiness of contracts and other costs of putting information to use in business-to-business settings. These findings contribute to a growing body of empirical bargaining studies showing evidence that "bargaining costs" play an important role in real-world negotiations (Backus et al. 2018; Jindal and Newberry 2018; Shelegia and Sherman 2018).

In Section 6, we extend our difference-in-differences analysis to consider the effects of access to benchmarking information for a wide variety of important product categories, from commodities (e.g., surgical gloves) to other physician preference items beyond stents (e.g., prosthetic hips). We find that the effects of information are broadly similar across product categories. Average treatment effects are often negative, but tend to be small and statistically insignificant. However, among hospital-brands in the top quartile of quantity and quintile of price at the time of join, treatment effects tend to be larger in magnitude and more often statistically significant. The largest effects we document are for the purchase of expensive physician preference items, where our treatment effect estimates indicate three to four percent savings due to benchmarking information among hospitals formerly paying the

highest prices.[7]

These estimates are of direct interest for considering the impact of information intermediaries on the prices buyers negotiate in previously opaque business-to-business markets. They also provide the first empirical evidence on the potential effects of transparency policies that have been proposed for medical technology markets, though we caution that benchmarking as implemented in our sample may have different effects than a broad policy change for many reasons. In Section 7, we conclude and discuss potential directions for future research on information in business-to-business bargaining.

2 Data

The primary data used in this study come from a unique database of all supply purchases made by approximately seventeen percent of US hospitals that joined a price benchmarking service during the period 2009-2014. The data are from the PriceGuide™ benchmarking service (hereafter, "PriceGuide data") offered by the ECRI Institute, a non-profit healthcare research organization. We observe unique, anonymous hospital identifiers and several hospital characteristics: Census region, facility type, and number of beds. For each transaction, we observe price, quantity, transaction month, and supplier for a wide range of product categories, including commodities (e.g., cotton swabs and gloves) as well as physician preference items (e.g., stents and orthopedic implants).[8]

Our analyses consider price negotiations between hospitals and suppliers for a large number of important product categories. The contracting environment is described in detail in Appendix C. Included products were the top 50 product categories by *either* total spending *or* transactions, for a total of 52.[9] As detailed in Appendix A, the data are at the stock-keeping-unit (SKU) level, requiring us to use manufacturer catalogs and classification algorithms to group SKUs that belong to the same manufacturer-brand.[10]

---

[7]For some product categories, we also find evidence that prices increase in the bottom of the distribution, consistent with a supplier response to an information externality across buyers in the asymmetric information framework. However, this evidence is sensitive to empirical specification, whereas the evidence of decreases in the top of the price distribution is quite robust.

[8]The reported data are of high quality because they are typically transmitted as a direct extract from a hospital's materials management database. Hospitals have strong incentives to report accurately because the analytics the benchmarking service's web portal provides are based on comparing the hospital's submitted data to that of others in the database.

[9]There are 80 "top" categories total, but we omit product categories that are too broad or with missing or inconsistent data. See Appendix A for details.

[10]Note that we use the term "brand" to refer to the "product" level at which prices are negotiated – e.g., Medtronic Resolute Integrity drug-eluting coronary stent. The use of "brand" is not meant to connote any particular marketing strategy. We use "product category" to refer to the "Universal Medical Device Nomenclature System (UMDNS) code" grouping included in the transaction files. The UMDNS system classifies devices by intended purpose and mechanism of action (e.g. drug-eluting coronary stents have

Table A5 summarizes the data for the 52 product categories of interest. Spending per month varies dramatically across product class: hospitals typically spend only $11 thousand per year on bandages, vs. over $1.1 million on coronary drug-eluting stents. Some product categories are used fairly universally – 618 sample hospitals purchased hypodermic injection needles – and some are used only in highly specialized facilities – only 249 sample hospitals purchased biological cardiac valve prostheses. As discussed in greater detail in Appendix A.3.1, we document similar price dispersion for commodities as we do for other medical/surgical products and PPIs – coefficients of variation within brand-time across hospitals are 0.195, 0.166, and 0.188, respectively. This suggests that opportunities for savings are similar, relative to pre-information prices, across product classes.

Our sample facilities are discussed in detail in Appendix A.3. Overall, our regression samples include 775 facilities spending an average of $1.8 million per month across 774 product categories. We return to the full set of important products in Section 6, but our main analyses focus on 508 sample facilities that purchase drug-eluting coronary stents.[11] Stent purchasing patterns are presented in Table A1. Briefly, the average sample facility spends $3.4 million per month on 1,143 categories of supplies, $73 thousand of which is dedicated to coronary drug-eluting stents. During our sample period 2009-2014, there are thirteen branded products sold by four manufacturers – Abbott, Cordis, Medtronic, and Boston Scientific, with Cordis exiting the market in 2011. The average hospital purchases 48 drug-eluting stents per month.

The hospitals in the purchase order data voluntarily joined a subscription service that allows them to benchmark their own prices and quantities against those of other hospitals in the database. We consider the effect of potential non-representativeness of our sample in our discussion of identification in Section 4. In Appendix A.1, we also compare our sample hospitals along observable dimensions to two outside samples of US hospitals – all American Hospital Association (AHA) member hospitals with cardiac catheterization, and Millennium Research Group's (MRG) Markettrack™ survey of a geographically representative sample of catheter labs. The west region is overrepresented in our benchmarking database sample, while the south is underrepresented. The average PriceGuide hospital is larger than the average US hospital with cardiac catheterization. Similarly, member facilities in our PriceGuide estimation sample purchased in higher volumes and obtained slightly lower prices than MRG hospitals in overlapping periods. Finally, the analyses in Appendix D.2 show that our main

---

UMDNS code 20383). Finally, we use "product class" to refer to broad groupings of product categories: commodities, physician preference items, and other medical/surgical products.

[11]The database includes a few other healthcare facilities such as clinics and surgical centers, but hospitals constitute 97 percent of the stent sample and 73 percent of the extended sample by count; more by spend. For the remainder of the paper, we use the terms hospital and facility interchangeably.

findings are qualitatively and quantitatively similar when we perform our analyses within hospitals in each Census region or within hospitals in the same bed size range.

2.1 Price Variation Across Hospitals and Brands

The left panel of Figure 1 displays the distributions of drug-eluting stent prices across hospitals and hospital-brands, respectively. It illustrates the wide variation in prices across PriceGuide sample hospitals in their pre-benchmarking transactions, with a standard deviation of $164 and mean of $1,615, for a coefficient of variation of 0.10. Hospital-product effects explain much of this variation with $R^2 = 0.89$ for the residual price variation (after brand-month detrending). Hospital effects in turn explain almost half of the hospital-brand variation, with $R^2 = 0.44$. Thus, our price variation is driven in part by some hospitals consistently paying more than others for all drug-eluting stents, and in equal part by some hospitals paying more for particular stents. As shown in Appendix A.1, these patterns are shared by the representative MRG sample, though the MRG sample has slightly higher prices on average. This implies that hospitals joining the database have slightly less to gain than the representative sample in terms of raw price differentials.

Figure 1: Price Distribution Across Hospitals, Brands, and Information State

(a) $\bar{P}_h$ and $\bar{P}_{hj}$, Pre-Information   **(b)** P_{hjt}, Pre- vs. Post-Information
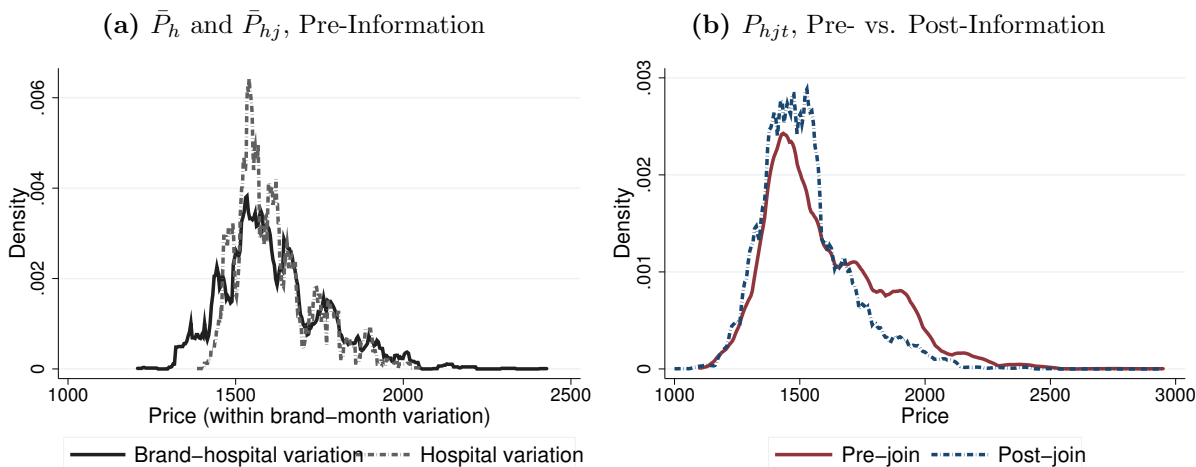
Authors' calculations from PriceGuide data. Left panel shows estimated hospital-brand and hospital fixed effects, obtained from regressions of price per stent, controlling for brand-month fixed effects. Right panel shows raw price distributions pre- and post- access to benchmarking information.

Observable hospital characteristics do not explain much of the variation in prices. Hospital bed size has no explanatory power. Total volume of stents purchased has more explanatory power: 10th decile hospitals by purchase volume (188 stents per month) achieve prices that are 7 percent lower than those obtained by 1st decile hospitals (7 stents per month).

However, we observe substantial dispersion in prices even conditional on facility size and purchase volume (see Appendix A.1 for details).

In a different data set, Grennan (2013, 2014) found evidence that heterogeneity in stent prices across hospitals could be explained in part by heterogeneity in physician brand loyalty, but this left a large residual heterogeneity in hospital-product *bargaining ability*.[12] Our analysis explores the possibility that part of this heterogeneity in bargaining abilities may be due to heterogeneity in information among hospitals, and that transparency in the form of benchmarking information on other hospitals' prices might affect this.

2.2 The Benchmarking Information Treatment

The information treatment considered in this study is one in which hospitals log in to a database and receive information about their relative performance in purchasing. The basic interface members access upon logging in presents graphical analytics for "potential savings" opportunities at the supplier level, defined as the total dollars that might have been saved in the previous year based on the hospital's volume of purchase and the mean/min prices paid by other hospitals at the manufacturer-SKU level. By clicking through, the hospital could observe these potential savings broken down by SKU, filter by geography and bed size, or even access the full (de-identified) purchase order microdata. We obtained clickstream data on the timing of all members' website logins, allowing us to reconstruct each member's benchmarking information set at each point in time.

In order to preview our approach and results in a simple graphical manner, the right panel of Figure 1 displays the histograms of prices paid for drug-eluting stents across the entire sample, splitting the sample into pre- and post-join observations. The raw data clearly suggests the primary impact of access to the benchmarking information: hospitals with information are much less likely to pay the highest prices. In the Sections that follow, we consider what theoretical mechanisms might drive this result in business-to-business negotiations as well as the research designs and regression specifications that will allow us to establish causal treatment effects and the mechanisms behind them.[13]

---

[12]In these and other studies of empirical bargaining, bargaining ability is parameterized by Nash weights in a structural model of full-information bargaining. These terms represent heterogeneity in prices after controlling for variation in competitive environment, captured by factors such as the outside option.

[13]We focus on the potential effect of information on negotiated *prices*. In Appendix F, we also estimate the effects of information on *quantities* and find no effect, consistent with stents being "physician preference items" where physician demand is based on strong preferences and is relatively insensitive to price.

3 Theory: Bargaining and Benchmarking Information

While knowledge of others' prices could potentially affect negotiations in many direct and indirect ways, the policy and economics literature on this setting (see, e.g., Pauly and Burns 2008), as well as our conversations with market participants, suggest that there are two primary mechanisms for how benchmarking information could be useful to hospital buyers. First, benchmarking could reduce asymmetric information about the price a supplier is willing to concede. Second, benchmarking could help solve the agency problem between the hospital and its procurement negotiators by providing a tool for the hospital to monitor negotiator performance relative to a market aggregate. In this Section, we outline models that capture each of these effects and use them to motivate our empirical predictions.

Our models build on the Rubinstein (1982) model of alternating offers bargaining.[14] The model has a single buyer negotiating with a single supplier over a per-unit surplus $V = wtp - c$ equal to the buyer's willingness-to-pay for a unit of the supplier's product, minus the supplier's marginal cost of manufacturing and distributing a unit of the product.[15] Beginning with the buyer, each player in turn makes a proposal for the division of the surplus. After one player has made an offer, the other must decide to accept or reject it and make a counteroffer in the next round. Players discount continued rounds of bargaining with discount factors $\delta^B \in (0, 1)$ for the buyer and $\delta^S \in (0, 1)$ for the supplier. In the institutional setting of bargaining over medical devices like stents, the typical negotiation occurs between a purchasing agent of the hospital and a sales representative of the device manufacturer. Each discount factor should be thought of as coming from a combination of negotiator skill and the incentives negotiators face as agents of their respective employers.

The unique subgame perfect equilibrium of this game is for it to end immediately with the buyer making an offer and seller accepting. The resulting complete information (CI) price is $p^{CI} := c + \delta^S \frac{1-\delta^B}{1-\delta^B \delta^S} V$. Thus, the observed variation in prices in our data could be generated

---

[14]This model underpins a large theoretical literature on bargaining (Rubinstein 1985; Binmore et al. 1986; Horn and Wolinsky 1988; Collard-Wexler et al. 2017) as well as a more recent empirical literature on bargaining (Draganska et al. 2009; Crawford and Yurukoglu 2012; Grennan 2013, 2014; Gowrisankaran et al. 2015; Lewis and Pflum 2015; Ho and Lee 2017). The predictions of the model extend well to empirical settings because the "discount factors" that parameterize bargaining strength in the Rubinstein model can be thought of more generally as proxies for a host of factors that might affect a real-world negotiation such as impatience, opportunity costs of time, laziness, or fear of negotiation breakdown.

[15]V_{hjt} (subscripts suppressed in text) should be thought of as the incremental value created by stent j for the set of patients for which the doctors at hospital h choose to use j over alternative stents or non-stent treatments, given physician preferences over all stents available at time t. For the sake of parsimony, we abstract here from price externalities across negotiations. Appendix D.3 provides a prediction regarding such externalities, and an empirical test of that prediction. Fully modeling a multilateral contract equilibrium (e.g. as in Collard-Wexler et al. (2017)) with the addition of asymmetric information is beyond the scope of this paper.

in a full information model by wide heterogeneity in discount factors and valuations across buyer-supplier pairs, in which case there may be no effect of benchmarking information.

3.1 Asymmetric Information about Bargaining Parameters

In order to introduce asymmetric information into the baseline bargaining framework, we follow Rubinstein (1985), in which hospital buyers have uncertainty about the bargaining parameter of a given supplier. The model departs from the complete information model outlined above in that the supplier is either of weak type with discount factor δ_w^S or strong type with discount factor δ_s^S $(1 > \delta_s^S > \delta_w^S > 0)$. The supplier knows his own type, but the buyer has only a subjective prior ω of the probability that the supplier is the weak type.[16]

Rubinstein (1985) shows that, in this asymmetric information (AI) game, if the buyer is sufficiently pessimistic about the seller being the weak type, then there exists a pooling equilibrium wherein the buyer simply offers what she would offer the strong type in a complete information game: p_s^{CI}, and both seller types accept this offer. If the buyer is more optimistic, then there exists a separating equilibrium wherein the buyer offers a low price p_w^{AI}, which the weak seller type accepts. But the strong seller type will reject this offer and counteroffer with p_s^{AI} (where $p_s^{CI} > p_s^{AI} > p_w^{AI}$), which the buyer accepts.

For simplicity, we begin by assuming that access to benchmarking information fully reveals a seller's type. Several empirical predictions for the effects of information on negotiated prices follow directly:

Prediction 1 (Direct Effect of Complete Information on High Prices) If information is costless, pessimistic buyers will always become informed. This information will cause the highest prices p_s^{CI} to fall to the complete information weak-supplier price p_w^{CI}, for those cases where the supplier was in fact the weak type.

Prediction 2 (Direct Effect of Complete Information on High Prices with High Quantity) If information is costly to obtain (e.g. searching and analyzing the data takes time that

---

[16]This model focuses on the case where uncertainty is embodied only in the discount factors and not the value over which negotiations occur, which is not directly testable without data on breakdown or beliefs because the surplus and bargaining parameters enter the price multiplicatively. However, this seems to fit the primary potential source of uncertainty in coronary stent negotiations, where doctor preferences are typically quite well known by those involved in the negotiation and marginal costs are small relative to the surplus created. It is also consistent with anecdotal evidence of little if any equilibrium breakdown in negotiations or destruction of surplus, which are central predictions of models of incomplete information about values (thanks to Brad Larsen for this observation). See Ausubel et al. (2002) for a review of the literature focused on informational asymmetries in values. Of particular note in that literature is Cramton (1992), which extends a model similar to the one here to a continuum of types and two-sided asymmetric information, and where information revelation arises endogenously through the timing of initial offers.

could be used on other productive activity), a pessimistic buyer will become informed whenever the expected benefit of information $\omega(p_s^{CI} - p_w^{CI})q$ exceeds the cost.

The above illustrative model and Predictions assumed that benchmarking information fully reveals the seller's type. However, the transparency introduced by price benchmarking could have a countervailing effect: sellers may be less willing to offer low prices if they know those prices will then be included in the database and potentially hurt them in subsequent negotiations with other buyers (Duggan and Scott Morton 2006; Grennan 2013). To show how this effect might arise, Appendix B provides more detailed analysis of the baseline asymmetric information model, plus a simple extension of the model in which the existence of benchmarking may create externalities across buyers in a multilateral environment.

The extended model allows for benchmarking data to provide a noisy signal of a seller's type to future buyers by revealing the current buyer-seller pair's price. Via this potential information externality, the presence of a future negotiation effectively puts a lower bound on the price the seller would be willing to accept in today's negotiation, and thus some of a seller's lowest prices may increase. Appendix B.1.2 presents one parameterization of the model to illustrate the mechanism of interest. The extended analysis yields an additional Prediction under a model of asymmetric information:

Prediction 3 (Effect of Information Externality Across Buyers on Low Prices) Via the benchmarking process, a negotiated price with one buyer becomes information available to other buyers, potentially introducing an externality in future negotiations. If this externality is large enough, it can provide the seller a credible threat not to accept low prices. Thus some prices may rise with the introduction of benchmarking information, particularly among buyers with relatively low prices and a relatively low quantity under negotiation.

Put plainly, Prediction 1 is that exposure to benchmarking information should lead to some of the highest prices falling (cases where the supplier was the weak type). Prediction 2 is that this effect will be more likely among those brands with the highest quantity used. Prediction 3 is that benchmarking may generate an externality such that some of the lowest prices increase, particularly among low-quantity buyers.[17]

---

[17]Alternative models that could generate similar empirical predictions might include models wherein one party has preferences over relative as well as absolute performance. See, e.g., Card et al. (2012) regarding pay transparency, in which workers learning they have relatively low salaries have reduced satisfaction and are more likely to leave their jobs.

3.2 Negotiator Agency

Another mechanism via which benchmarking information could be valuable to buyers would be through providing aggregate information to help the buying firm solve a moral hazard problem with its purchasing agent. We expect this mechanism to be relevant in the cardiac unit context. McConnell et al. (2013) present survey data documenting that hospitals' cardiac units vary substantially in their focus on performance measurement, and the Centers for Medicare and Medicaid Services recently found that cardiac and orthopedic units in hospitals responded to bundled payments (which entail higher-powered financial incentives) by improving contracting with suppliers.[18]

Extending the model presented thus far, suppose that instead of the hospital negotiator's bargaining parameter being exogenous, the price will be a function of the hospital agent's *choice* of discount factor δ^B. Further, suppose that in addition to uncertainty as to whether the supplier is a strong type or a weak type, there is an additional i.i.d. shock to the supplier's bargaining parameter that is buyer-specific (see Appendix B for details in the case where hospital h faces a supplier bargaining parameter equal to $\delta_h^S \in \{\delta_w^S \epsilon_h, \delta_s^S \epsilon_h\}$ for $\epsilon_h \sim U[0, 1]$). Supplier bargaining strength is then observable to the hospital negotiating agents, but not to the principals who manage them.

A moral hazard problem arises in this setting because bargaining effort is costly and provides the agent disutility. Under the usual assumption that the agent is risk averse, the optimal employment contract involves risk sharing between the principal and the agent. Holmstrom (1982) shows that if agents face some common parameter which is uncertain from the principals' perspectives (here, the portion of the bargaining parameter δ^S that reflects whether the supplier is a strong or weak type), then relative performance evaluation compared to some aggregate sufficient statistic can be used to write a stronger incentive contract with each agent.

In our interviews with industry participants, we did not encounter a single case where purchasing agent contracts were formal functions of a quantitative performance metric based on benchmarking information or otherwise.[19] However, we did learn that some hospitals use measured price decreases as part of a broader performance review or employee recognition program. When benchmarking information was available, we also heard cases of such data being used to quantify relative performance and opportunities for savings. This is in keeping with the spirit of the model above, motivating the following Predictions:

---

[18]See Calsyn and Emanuel (2014). The role of incentives in purchasing has also been examined in the broader government contracting context – e.g., in Bandiera et al. (2009), Italian public bodies' prices for generic goods vary with institutional characteristics and poor performance is attributed to passive wastefulness rather than corruption.

[19]See Appendix C.2 for details.

Prediction 4 (Monitoring Effect on Prices) If buyer negotiators are imperfect agents of the buying firm, then benchmarking information (observing the distribution of price realizations across hospitals) allows the principal to estimate whether the seller is the weak or strong type, and thus reduces the risk to which the agent is exposed in a higher-powered contract. The higher-powered contract induces more bargaining effort and a lower price than the case where only the buyer's own price is observed.

Prediction 5 (Monitoring Effect on Prices with High Quantity) Further, information will be used in this way when the expected benefit $\mathbb{E}[(p^{NoInfo} - p^{Info})]q$ exceeds the cost of information use.

In sum, Prediction 4 suggests prices will decrease on average upon the introduction of benchmarking; Prediction 5 suggests this effect will be more likely when greater purchase quantities are at stake.

3.3   New Brand Entry and Timing of Information Effects

Although the asymmetric information about supplier bargaining type mechanism and the negotiator agency mechanism can generate similar empirical predictions, an interesting feature that differs between the two mechanisms is the timing during which benchmarking information is valuable to the buyer. In the asymmetric information case, benchmarking is only useful if data on other buyers' prices for the same brand are *currently* available in the database at the time of negotiation. By contrast, even if there is no current data on others' prices for a given brand, agents may be incentivized today based on performance assessments taking place in the *future*.

This difference between the timing of information required for the two mechanisms is especially relevant when new brands enter the market. There will be no data in the benchmarking database on a brand for the first month or two it is on the market, and little data for the first few quarters. Thus, those who engage in their first negotiation for a new brand soon after its release do so without *current* benchmarking information, even if they have access to the database. This motivates our final theoretical prediction:

Prediction 6 (New Brand Entry Separates Asymmetric Information and Agency) For newly introduced brands, when they are first released to the market, differences between prices negotiated in the first (uninformed) round of negotiation and the second (informed) round of negotiation must be due to informing negotiators about the seller's bargaining parameter, rather than altering moral hazard.

<center>13</center>

Our empirical implementation of this idea identifies the effects of any contract redesign that negotiators are made aware of upon the firm joining the database *and* that incentivizes effort prior to benchmarking realizations. This structure, in which today's performance impacts tomorrow's information and, accordingly, compensation, is the approach taken in most explicit pay-for-performance schemes in health care markets (see James (2012)). In more general compensation schemes, relative performance evaluation can be part of employee compensation contracts with explicit bonuses (e.g., sales force compensation) or rewards can be focused on raises and promotion (see Lazear and Oyer (2012) for a review).

3.4 Other Considerations

In the interest of clearly illustrating the fundamental ideas behind the two theoretical mechanisms of interest, we have abstracted from some realities of hospital purchasing. Here we touch on some key features omitted from the model and how they affect the empirical analysis that is the focus of the paper.

First, to the extent that renegotiation is not frictionless, it will take time and effort to get to the negotiating table and come to a new deal – prices will be "sticky". This will tend to bias the short-run effect of information toward zero. We consider these dynamics in our empirical analyses using event studies and direct examination of recontracting.

Second, the same supplier salesperson may be in charge of negotiating contracts for a bare-metal and a drug-eluting stent, or for subsequent generations of a branded drug-eluting stent. To the extent that learning about types in the models above captures something that is specific and unchanging over time about that salesperson and the incentives she faces, there will be less asymmetric information and scope for learning, biasing the effect of benchmarking information toward zero.

Finally, while demand side effects of information are generally null or beneficial to buyers, to the extent that suppliers know when buyers join the benchmarking database (or transparency is imposed via public policy), then the model may omit supply side responses that may negate or overturn these effects through greater obfuscation (Ellison and Ellison 2009), or by facilitating collusion (Albek et al. 1996; Cutler and Dafny 2011). We return to this issue, and the extent to which our empirical estimates may capture these various supplier responses, in our discussion of results.

4   Identification of Information Treatment Effects

The key features of the data that allow us to estimate causal treatment effects of price transparency for the hospitals in our sample are: (1) that new members submit one year of

retrospective data when they first join the benchmarking database, and continue to submit monthly data thereafter; (2) that new members join over time in a staggered (and seemingly random with respect to price trends) fashion; and (3) that new brands enter the market at points in time that are seemingly uncorrelated with members' information states.

For hospitals that joined during the 2009-14 period, we observe data before and after they were first able to access the benchmarking information available in the database. The left panel of Figure 2 shows the time series of stent-purchasing hospitals joining the database between 2010 and 2014. Beginning in Q2 2010, 14 such hospitals joined the database in each quarter, on average.[20]

4.1   Using Join Date to Identify Price Effects

We leverage this variation using a series of difference-in-differences strategies. In our sample, all hospitals by definition access the benchmarking data at some point. The "control hospitals" for analyzing the price trend of hospital h in a window around h's joining the database are those hospitals $k \in \mathcal{H} \setminus h$ that subscribe either prior to or after that window. Under the standard assumption of parallel trends, we can isolate the treatment effect of joining the database on prices by comparing the price trends between treatment and control hospitals for their overlapping time periods.

The primary concern with this identification strategy is that timing of a hospital joining the database may be correlated with other contemporaneous factors that impact price trends at that hospital.[21] However, there are several institutional features which one might expect to dampen potential join-time bias for any particular focal brand. First, our conversations with industry participants indicate that it is unlikely for a hospital to join the database due to any single product category. PPIs such as stents are important purchases for hospitals, but subscription is costly and meant to cover a large number of product categories. Second, many determinants of price trends are specific to a product category market, limiting the degree of correlation in price trends across, for example, coronary stents and knee prostheses. Finally, hospital purchasing tends to be separated across groups of product categories, implying that, for example, organizational changes regarding purchasing in the catheter lab need not correlate with changes in orthopedic surgery. Thus any particular product category or brand

---

[20]The above figure focuses on hospitals purchasing coronary stents. Over the same period, 36 hospitals purchasing any important product in our data joined in each quarter.

[21]For example, a hospital may be inspired to join the database due to concerns about upward price trends, which could induce a positive bias in our results – we would be underestimating the counterfactual prices joining hospitals would face if they did not join. On the other hand, a joining hospital might concurrently be undertaking other initiatives intended to reduce input prices, such as hiring new personnel or contracting other outside consulting services. Conflating the effects of these other initiatives with the effect of access to the benchmarking information could induce a negative bias in our results.

Figure 2: Identifying Variation and Graphical Example of Identification

(a) Timing of Join and New Brand Entry

(b) Identifying Information Effects from New Brands
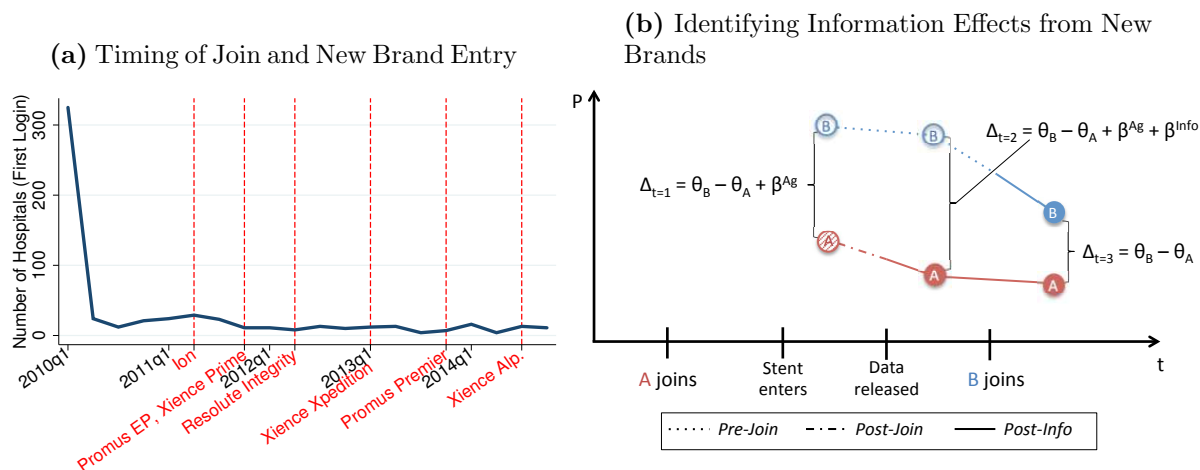
Left panel: authors' calculations from PriceGuide data, 2010-2014. "Join" defined by member's first associated login. New brand entry indicated with vertical red lines. Note: PriceGuide rolled out a new version of its web interface in the beginning of 2010 and re-invited all current members to "join"; members' whose first associated login is in Q1 2010 may have subscribed in 2009 or earlier. All such members' "pre-join" data is excluded from empirical analyses. Right panel: graphical illustration of "new brand" identification strategy.

within a category is likely to be a "bystander" to the join timing. This is consistent with the event studies in Section 5, which show no evidence of differential pre-trends in price in the months before hospitals join the database.

4.2 New Brand Entry, Mechanisms, and Bias

New brand entry provides another opportunity to identify the above information effect, and further allows us to identify a treatment effect of having joined the database but not yet having access to *concurrent* data on other hospitals' purchases. After new brand entry, there is a lag before members may access benchmarking data on the new brand due to lags in data submission and loading: in the months following new brand entry, the count of members purchasing that brand exceeds the count of members with transactions loaded in the database by 56 on average. Moreover, we observe transactions for new brands for some members before and after they join the benchmarking database: in the year following brand entry, 9 percent of members whose transactions are observed in the average month are pre-join (details in Appendix A.2). The time period for our study contains many drug-eluting stent brand introductions. In panel (a) of Figure 2, we note the timing of entry of seven new brands between 2010 and 2014, of the thirteen brands sold during this time period overall.

This variation allows us to identify a treatment effect of access to benchmarking infor-

mation via a mechanism that does *not* require concurrent access to data on other hospitals' purchases. In our analysis, we term this the agency ("Ag") effect to denote its relation to the mechanism outlined in Section 3 in which the benchmarking database allows hospitals to resolve a negotiator agency problem. Panel (b) of Figure 2 illustrates this identification strategy graphically. In this stylized example, we have hospital A joining the database well before the brand enters the market; hospital B joins after the brand enters. Once the brand enters, each hospital negotiates prices; hospital B is untreated, while hospital A is treated ("Ag") in the sense that it has joined but has no concurrent data on other hospitals. In the next period, after price data are submitted, loaded, and released to database members, non-member hospital B remains untreated, but hospital A receives another treatment ("Info") in the form of information on other hospitals' prices. In the final period, hospital B has joined the database and received the full treatment effect of access to benchmarking data ("Ag" + "Info"); hospital A retains both treatments in the final period as well.[22]

Entering brands also allow us to investigate potential bias due to timing of join. Any persistent bias associated with factors beyond information at hospitals that have joined will be included in the difference between pre- and post-join hospitals in the first few months after new brand introduction (β^{Ag}). Thus, our estimate of any "asymmetric information" effect where hospitals use information concurrently available in the database to negotiate better prices (β^{Info}) would be free of such bias.

5   Results: How Information Affects Negotiated Prices

In this Section, we estimate regressions based on the research design just described to more carefully measure and understand the effect of information suggested by Figure 1, accounting for time-invariant differences across hospitals (or hospital-brand combinations) and trends in prices over time. All of the regressions we present are extensions of a baseline specification implementing our difference-in-differences design around the timing of hospital access to benchmarking information.[23] Letting P_{hjt} denote the price observed for hospital h, brand j, and month t, our preferred specification controls for hospital-brand fixed effects $[\theta_{hj}]$, month

---

[22]Formally: in the final period, we identify the fixed hospital differences ($\Delta_{t=3}$); in the penultimate period, we identify the fixed differences plus the "agency" and "information" effects ($\Delta_{t=2}$); and in the first period, we identify the fixed differences plus the "agency" effect only ($\Delta_{t=1}$). These three differences allow us to separately identify the agency ($\beta^{Ag} = \Delta_{t=1} - \Delta_{t=3}$) and information ($\beta^{Info} = \Delta_{t=2} - \Delta_{t=3} - \beta^{Ag}$) effects.

[23]This includes information upon joining and when new brands enter. We show results estimated only from "timing of join" variation in Appendix A.3 and find our discussion unaffected.

17

fixed effects $[\theta_t]$, and separate linear time trends for each brand $[\gamma_j * (t - t_{min_j})]$:[24]

$$P_{hjt} = \beta^{Info} * \mathbb{1}_{\{post_{hjt}\}} + \theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j}) + \varepsilon_{hjt}. \tag{1}$$

Here, $\mathbb{1}_{\{post_{hjt}\}}$ is an indicator equal to one after a hospital first accesses benchmarking information for the given brand and zero prior, making the coefficient β^{Info} an estimator for the treatment effect. All of the regressions and results below extend this specification to allow for varying types of heterogeneity in this treatment effect. Results with alternative fixed effects and time trends are discussed as well.

5.1 Effects of Information throughout the Price Distribution

Our first result, shown in Table 1, regards the average treatment effect of information across all hospital-brand-months. Results are shown for a variety of different specifications of control variables. The estimates are significantly smaller when we control for hospital-by-brand (rather than hospital *plus* brand) fixed effects. This may be due to our effectively controlling for an unknown source of hospital-brand specific heterogeneity; or hospital-brand fixed effects may introduce attenuation bias towards zero, as there are some hospital-brands for which there are relatively few observations. We generally find that Version (3) treatment effects are smaller in magnitude and more precise than Version (4), so we focus on these results in the main text for the sake of brevity.[25]

Table 1: Average Treatment Effects of Information across All Hospital-Brand-Months

Version of Controls	[1]	[2]	[3]	[4]
β^{Info}	-12†	-21†	-3	-7
	(5)	(7)	(3)	(5)
Hospital+Brand FEs	Y	Y	N	N
Hospital×Brand FEs	N	N	Y	Y
Linear Brand Trends	Y	N	Y	N
Brand×Month FEs	N	Y	N	Y

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,453$ member-brand-months. Includes 508 members. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

[24]t_{min_j} is the first period in which we observe data for brand j: either the beginning of our sample or the month of entry of brand j into the market. To address concerns that linear trends do not adequately account for price trends at the beginning of a brand's life cycle, refer to Appendix D.1 for results with brand-month fixed effects, which are qualitatively similar.

[25]See Appendix D for all versions of heterogeneous treatment effect and mechanism results. The results are similar, with the primary difference being that effects in the top of the price distribution roughly double in size with hospital instead of hospital-brand fixed effects. This difference is due to a significant negative "agency" effect in the hospital fixed effect specifications, which does not appear in the specifications that control for hospital-brand fixed effects.

The preferred specification finds that prices decrease by only \$3 *on average* when benchmarking data are accessed. This average treatment effect (ATE) is also imprecisely estimated, with a standard error of \$3.[26] In keeping with the empirical predictions derived from theory, the remainder of our analyses will allow for heterogeneity in treatment effects depending on each hospital-brand pair's place in the price distribution for that brand at the time the hospital gains access to information. Upon member h's first login to the database ($t = \tau_h$), we compare P_{hj,pre_h} for each brand j purchased in the year prior to login ($pre_h := \{t \in [\tau_h - 13, ..., \tau_h - 1]\}$) to the full distribution of prices $\{P_{h'j,pre_h}\}_{h' \in \mathcal{H}}$ across all hospitals $\mathcal{H}$ during (pre_h). We assign each pair hj to a quintile k of the pre-join price distribution. We then re-estimate (6), such that coefficient β_k^{Info} represents the estimated treatment effect of accessing information in the benchmarking service for quintile k of the pre-information price distribution $\mathbb{1}_{\left\{P_{hj,pre_h} \in quantile\left(k, \{P_{h'j,pre_h}\}_{h' \in \mathcal{H}}\right)\right\}}$.

Panel (a) of Figure 3 shows the results. The treatment effects exhibit substantial heterogeneity depending on the pre-information price the hospital was paying for a brand relative to others. The treatment effects are statistically zero in all but the top quintile of the pre-information price distribution, where the effect is -\$27. This evidence is consistent with Prediction 1 that, absent benchmarking, pessimistic hospitals would pay suppliers high prices regardless of those suppliers' true bargaining parameters, leading those hospitals to negotiate lower prices after joining. Under the baseline asymmetric information mechanism, there is little reason to expect transparency to affect prices that are *relatively good*. In fact, the treatment effect for the bottom quintile of the price distribution is small and negative, providing little support for the externality discussed in Prediction 3.[27]

We also performed an event study analysis separately for each quintile of the price distribution. The results for the top quintile of the pre-information price distribution are shown in panel (b) of Figure 3.[28] The pre-trends in the six months pre-information are essentially zero, while there is a steady decline in prices after information access – a year after join, the treatment effect is -\$96 relative to the "Info" date. The estimates for the 6-12 months prior to information access are negative, though not significant. If one were to lend weight to the noisy point estimates, pre-trends prior to joining the database would lead the difference-in-differences estimates to be an understatement of the effects of information on price. This lack of pre-trends is strong suggestive evidence that the estimated treatment effects are due to accessing the information in the benchmarking data rather than to any potential sources

---

[26]Detailed tables and figures on the timing of the effect are available in Appendix D.1.

[27]A positive shift in the lowest part of the price distribution could be evidence of suppliers' reduced willingness to provide discounts in the presence of benchmarking, or of mean reversion. These results suggest that neither is a significant factor for stents.

[28]Results estimated using alternative fixed effects are available in Figures A11 and A12.
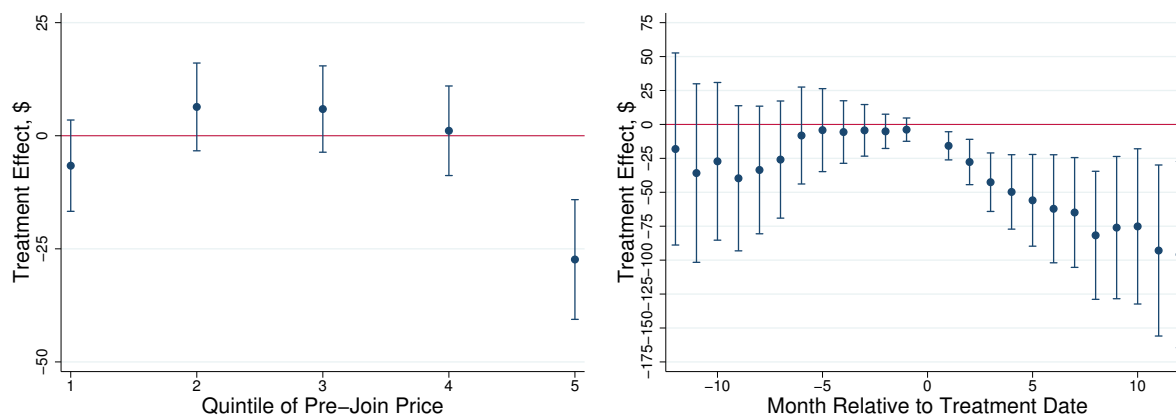
(a) DiD Estimates by Quintile **(b)** Top Quintile Event Study

Authors' calculations from PriceGuide data, 2009-2014. Panel (a): $N = 32,453$ member-brand-months; includes 508 members. Panel (b): $N = 23,016$ member-brand-months; includes 507 members, twelve months pre- and post-join only. Standard errors clustered at hospital-brand level.

Figure 3: Treatment Effect Estimates Throughout the Price Distribution

of bias due to join timing.[29]

For the sake of statistical power and for expositional simplicity, we return to estimating pre-/post-treatment effects, rather than breaking them down by month relative to information access. However, it is noteworthy that treatment effects become larger over the course of the year after information access. We see this as evidence of price "stickiness" as a friction that limits gains from transparency, and we return to this issue in Section 5.2.3.

5.2 Mechanisms: Where and Why Does Information Matter Most?

The above results establish that transparency in the form of access to benchmarking information leads to lower prices for hospital-brand cases where the hospital is in the upper quintile of the price distribution for that brand. In this Section, we test the further predictions from Section 3 to better understand the mechanisms behind these price reductions. We first allow for treatment effects to vary with purchase volume so that we may investigate whether hospital-brands with high expenditures at stake experience larger changes, in keeping with Predictions 2 and 5. Next, we utilize the fact that, for new brands, no benchmarking in-

---

[29]Indeed, the evidence of steeper negative price trends after join in the top quintile of the price distribution than there are in average prices suggests that, if there are factors that cause prices to decrease after join that are unrelated to information access, they must disproportionately impact hospital-brands whose prices are relatively high in the pre-period, a fact which would be unknown to parties whose behavior impacts prices without them accessing the database. For the reader who prefers a more skeptical interpretation, any remaining bias due to timing of join will be absorbed with our measure of the agency effect in our mechanism test, so that we are able to obtain a "clean" asymmetric information effect.

formation is available in the database until several months after brand entry; this allows us to separate the asymmetric information mechanism from the agency mechanism (Prediction 6). Finally, we decompose the estimated price effects into price effects conditional on renegotiation and price effects due to greater likelihood of renegotiation. The estimates are summarized in Table 2 here and discussed in turn below.[30]

Table 2: Treatment Effects of Information: Mechanisms

	(1) Treatment Effect Variation with Quantity Purchased:									
	Low quantity: $\beta^{Info}_{quintile,low^q}$					High quantity: $\beta^{Info}_{quintile,high^q}$				
$P_{quintile}$:	1	2	3	4	5	1	2	3	4	5
	-4	9	9	5	-17^{**}	-11	0	0	-9	$-71^{\dagger}$
	(6)	(6)	(6)	(6)	(7)	(9)	(8)	(7)	(8)	(13)
	(2) Agency vs. Asymmetric Information Mechanisms:									
	Future Info: $\beta^{Ag}_{quintile}$					Concurrent Info: $\beta^{Info}_{quintile}$				
$P_{quintile}$:	1	2	3	4	5	1	2	3	4	5
	-17	-3	2	7	13	-1	7	5	-1	$-30^{\dagger}$
	(11)	(12)	(10)	(12)	(18)	(6)	(5)	(5)	(5)	(7)
	(3) Renegotiation:									
	Pr renegotiation: $\mathbb{1}_{\{reneg_{hjt}\}}$					Upon renegotiation: $\beta^{Info}_{quintile,\mathbb{1}_{\{reneg_{hjt}\}}}$				
$P_{quintile}$:	1	2	3	4	5	1	2	3	4	5
	.01	.013	$.016^{*}$	.018	$.023^{**}$	-14	4	1	-13	$-76^{\dagger}$
	(.01)	(.01)	(.009)	(.011)	(.009)	(15)	(14)	(19)	(17)	(18)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,453$ member-brand-months. Includes 508 members. Standard errors clustered at hospital-brand level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

5.2.1 Costs of putting information to use: treatment effects and quantity

To the extent that using benchmarking information to identify opportunities and then engage in renegotiation (of supply contracts or employment contracts) is costly, Predictions 2 and 5 suggest that transparency will have the largest effect for hospitals and brands with high quantities involved. To investigate these predictions, we generate dummy variables $\mathbb{1}_{\{low^q_{hj,pre}\}}, \mathbb{1}_{\{high^q_{hj,pre}\}}$ that divide the sample into hospital-brands with monthly purchase volume below and above the 75th percentile in the months prior to join, and we estimate a model that allows for treatment effects to vary by pre-join price and quantity where β^{Info}_{k,low^q} now estimates the treatment effect, for quintile k, for lower volume brands; and $\beta^{Info}_{k,high^q}$ now estimates the treatment effect, for quintile k, for higher volume brands.

The estimates in Panel (1) of Table 2 show that the price treatment effect is largest for high-volume hospital-brands in the upper part of the price distribution. At -$71, the

---

[30]Detailed results with alternative fixed effects specifications are available for each of the below panels in Figures A13, A14, and A15 in Appendix D.1.

top-quintile treatment effect for high-quantity hospital-brands is more than four times the effect for low-quantity hospital-brands.

The fact that quantity matters suggests that costs of attention, analysis, or action act as frictions that sustain price variation. That more savings are not realized, even when large quantities are at stake, suggests that further frictions independent of information, such as strong physician brand preferences, could be important as well. As shown in Appendix D, we see similar patterns when we consider different sets of fixed effects and time trends, when we modify the regression sample to focus on only the twelve months pre- and post-information, when we limit the sample to hospitals only, and when we estimate effects within similar sets of hospitals. The results are also similar when we identify treatment effects based only on the information shock of database "join" as part of the expanded analysis discussed in Section 6. All told, these results imply partial price convergence: removing time trends and applying our treatment effect estimates to the pre-join price distribution decreases the standard deviation of price by 3.7 percent among low-quantity hospital-brands and by 6.4 percent among high-quantity hospital-brands.

5.2.2 New brands: agency and asymmetric information mechanisms

The β^{Info} estimates thus far have provided a treatment effect of access to the benchmarking information, subsuming both the agency and asymmetric information mechanisms that market participants put forth, as outlined in our Section 3. To separate these theories, we leverage the fact that almost all hospitals negotiate their first contract with a new brand by the first or second month after its introduction, but the resulting purchase order data will not begin to show up in the benchmarking database until month three or four. By month six, there are enough observations in the database for a hospital to develop a useful estimate of its place in the price distribution for that brand. We use this to estimate a regression in the spirit of Figure 2 that allows for heterogeneous treatment effects by price quintile and by these two information states. We estimate β_k^{Ag} by interacting the price quintile k treatment effect with an indicator for all hospital-months after the hospital joins and logs in to the benchmarking database $\mathbb{1}_{\{post_{ht}^{join}\}}$; and we estimate a "clean" β_k^{Info} by including a further interaction with an indicator that is equal to one upon a hospital's first login more than six months after the introduction of that brand $\mathbb{1}_{\{(t-t_{min_j})>6\}}$.

The estimates in Panel (2) of Table 2 suggest that the asymmetric information effect explains a substantial portion of the effect of information on prices. To the extent that one remains concerned about endogenous timing of join, recall from Section 4 that any associated bias will be captured in β^{Ag}, but not β^{Info}. Thus, our most robust finding is that of a statistically and economically significant β^{Info}, concentrated among those paying

22

the highest prices before obtaining information, and consistent with the use of concurrent information in bargaining. These results are most consistent with the theory of asymmetric information in bargaining based on Rubinstein (1985).

One implication of this result is that asymmetric information may be among the effects driving the heterogeneity found in bargaining parameter estimates in studies using full information Nash Equilibrium of Nash Bargaining models. It suggests that these information and incentive issues should be kept in mind when thinking about the factors driving bargaining outcomes, including as potential sources of changes to bargaining parameters in counterfactuals with negotiated prices.

5.2.3 Price changes with "sticky" contracts

All of the price coefficient estimates reported thus far can be described as capturing the combined effects of information on the probability that price negotiation occurs *and* on prices arrived at during each price negotiation. We consider this to be the treatment effect of interest for policy, as it estimates the overall value of access to benchmarking information for decreasing the total spend of hospitals on medical inputs over time, taking into account the stickiness of prices in real-world hospital-supplier contracting. However, in the main estimation sample, renegotiations take place in 9 percent of observations (member-brand-months with any transactions), and prices decrease on average by \$25 at each renegotiation, meaning that we would expect small price changes to occur if information led to larger price decreases at each renegotiation *or* if information increased the likelihood of renegotiation.

We consider these two effects separately by flagging hospital-brand-month observations in which renegotiation is observed $\mathbb{1}_{\{reneg_{hjt}\}}$.[31] We then estimate the effect of information on the rate of renegotiation using the usual price quintiles specification, but with the indicator for renegotiation as the dependent variable. The effect of information on price, conditional on renegotiation, is obtained from the same regression, run only on the subset of data where the renegotiation indicator equals one.

The results in Panel (3) of Table 2 show the effect of information on the likelihood of renegotiation is only statistically significant (at the five percent level) in the top quintile of

---

[31]We sort transactions for each hospital-brand by month and group observations with the same price together within month. We then flag each hospital-brand-month as including a renegotiation event if we observe that prices change *and* that the price change "sticks" for two cumulative months after the renegotiation event (or until the final observed purchase for that member-brand). The results are qualitatively similar (though larger in magnitude) using a less conservative method that flags all months in which average prices change. Of course, with transactions data, we cannot observe if a renegotiation took place and price remained the same, but the baseline level of these events is differenced out in our estimation strategy. We also take some comfort that our measure results in frequency of renegotiations similar to the annual contract structure that is common in the industry.

price, where information increases the probability of renegotiation by 2.3 percentage points, or about one quarter the baseline probability of renegotiation. Point estimates in other quintiles are positive but smaller and not significant at conventional levels. To the extent this is not simply a statistical coincidence, it could be due to a slight increase in efforts to get to the negotiating table or change in the frequency with which renegotiation results in zero price change among those with information.

By contrast, the effect of information on price conditional on renegotiation is about -$75, nearly three times the -$27 effect on price paid. Thus, the impact of transparency in the form of benchmarking information is substantially affected by renegotiation frictions.

6 Generalizing the Results – All Important Products

While the above results from the coronary stent sample are useful for investigating mechanisms via which savings are achieved for an important product category, a natural question arises: what happens to the remaining 98 percent of hospital spending? To investigate this question, we extend our analyses to 52 product categories that are important in terms of high spending or transaction counts. We organize product categories based on the likely importance of physician preferences in determining their usage. Class 1 is "commodities": relatively inexpensive products that are unlikely to be chosen primarily by physicians; e.g., surgical gloves. At the other extreme, Class 3 is "physician preference items": high-tech medical devices, mainly coronary and orthopedic products that are the primary implanted device in their corresponding surgical procedures; e.g., coronary stents. Class 2 is intermediate: other medical/surgical products used during invasive procedures, but explicitly excluding PPIs.[32] See Appendix A.3 for sample details.

In this expanded set of analyses, we estimate equation (6) and its variants within each product category, using join timing to identify treatment effects of information. Appendix E reports estimated average treatment effects and treatment effects by price and quantity for each product category, including several alternative regression specifications, and for the full sample as well as a restricted sample focusing on products purchased in at least 100 hospitals on average in each year (where price quintile estimates have greater statistical power). Here, we summarize the key insights using the full sample. Figure 4 plots the average treatment

---

[32]Our typology overlaps substantially with the Food and Drug Administration's classification system. Class I devices, such as gloves, are deemed to be low risk and are therefore subject to the least regulatory controls. Class II devices, such as catheters, are higher risk devices with greater regulatory controls to provide reasonable assurance of the devices' safety and effectiveness. Class III devices, such as replacement heart valves and coronary stents, are the highest-risk devices and must typically be approved by FDA before they are marketed (United States Food and Drug Administration 2018). For product categories that did not obviously fit into one of our classes, we relied on the FDA class directly.

effect (left panel) and high-price, high-quantity treatment effect (right panel) for each of the 52 product categories, normalized by mean price to facilitate comparisons across groups, and organized by product class. Within each product class, treatment effects are displayed in ascending order, followed by an "X" marker indicating the spending-weighted average treatment effect for the class. The table below the Figure shows spending-weighted averages of the estimated treatment effects for the top and bottom quintiles of the price distribution across all product categories within each class, in percentage terms.
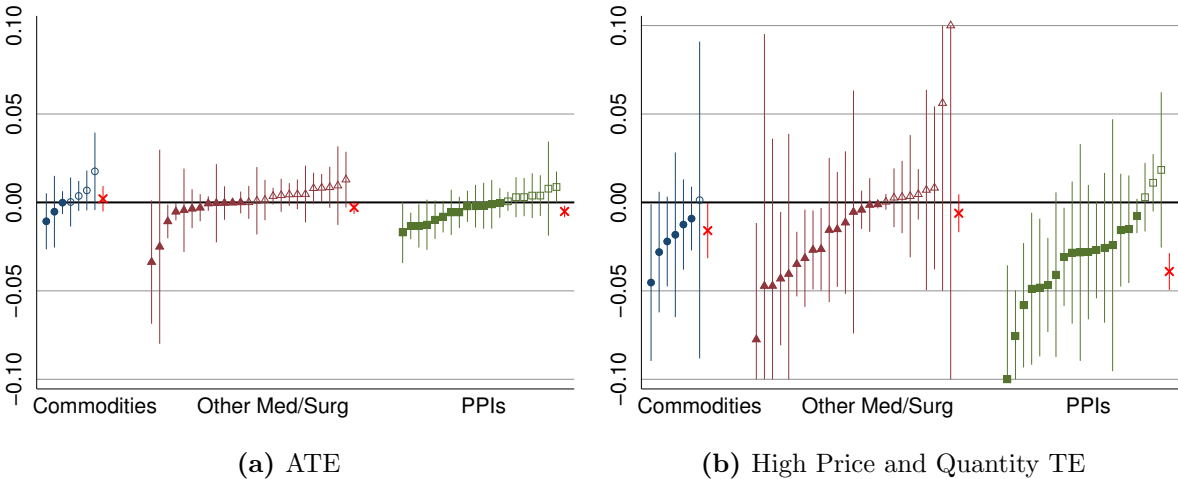
| | (a) ATE | (b) High Price and Quantity TE |

| | ATE | TE by Price Quintile (All Q) | | TE by Price Quintile (Low Q) | | TE by Price Quintile (High Q) | | Expected Savings | |
		1	5	1	5	1	5	$\mu(\$/\text{h-yr})$	$\sigma(\$/\text{h-yr})$
(1) Commodities	0.002	0.013*	-0.013†	0.021	-0.012**	0.005	-0.016**	63	2,828
	(0.004)	(0.008)	(0.005)	(0.016)	(0.006)	(0.008)	(0.008)	(306)	(1,028)
(2) Other Med/Surg	-0.003	0.003	-0.017†	0.004	-0.021†	0.002	-0.006	-254	1,974
	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(428)	(279)
(3) PPIs	-0.005†	0.014†	-0.033†	0.014†	-0.033†	0.014†	-0.039†	-1,869	5,492
	(0.002)	(0.003)	(0.004)	(0.003)	(0.004)	(0.005)	(0.005)	(1,281)	(893)

Authors' calculations from PriceGuide data, 2009-2014. $N_{hjt}^{(1)} = 516,582$; $N_{hjt}^{(2)} = 1,344,515$; $N_{hjt}^{(3)} = 703,544$. $N_h^{(1)} = 748$; $N_h^{(2)} = 701$; $N_h^{(3)} = 601$. Reported specifications include hospital-brand and brand-year fixed effects; alternative fixed effects shown in Appendix. Standard errors in category-specific regressions clustered at hospital-brand level. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure 4: Treatment Effect Estimates for Important Product Categories

We observe several general patterns similar to those from stents. In panel (a), we see that the average treatment effects are negative for the majority of product categories, but relatively small and rarely statistically significant. As shown in the table below the Figure, the aggregate average treatment effect is largest for PPIs, at -0.5 percent. In panel (b), we see that among hospital-brands in the top quartile of quantity and quintile of price at the time of joining, the treatment effects are larger in magnitude, almost always negative, and sometimes statistically significant, particularly among PPIs. In aggregate, we observe the largest price decreases under transparency for hospitals purchasing PPIs at formerly high prices (3.3 percent savings) and in relatively high quantities (3.9 percent savings).

Our finding of negative treatment effects in the top of the price distribution for PPIs is quite robust to decisions regarding sampling and regression specification. As discussed in greater detail in the Appendix, the numerous specifications reported in Tables A10-A12 consistently document a negative and significant average PPI estimate for high-price, high-quantity hospital-brands, usually document a negative and significant estimate for high-price hospital-brands, and report treatment effect estimates that are broadly similar to our preferred estimate of 3.3-3.9 percent savings on PPIs under transparency.[33]

In contrast to stents, the expanded analysis also reveals price *increases* in the bottom part of the price distribution for some product categories. As summarized in the table below Figure 4, we document a marginally significant price increase of 1.3 percent in the bottom quintile of the price distribution for the average commodity; the analogous result for PPIs is 1.4 percent and is estimated more precisely. The detailed category-level results in Table A10 in the Appendix show that this pattern is particularly strong for several prosthesis categories in the PPI class. These results are consistent with an externality effect of transparency as set forth in Prediction 3 and accordingly merit further discussion. Focusing on PPIs, the price increases in the bottom of the distribution are smaller than the price decreases in the top of the distribution (3.3 percent). The PPI-average point estimates are the same for buyers purchasing in relatively low quantities and for buyers purchasing in relatively high quantities (1.4 percent); our theoretical framework predicted larger effects for low-quantity buyers, so this is suggestive evidence against the externality model, but standard errors are such that we cannot rule out an economically meaningful difference in effects.

Also in contrast to the robust results we report for the top of the price distribution, the results for the bottom part of the price distribution are somewhat fragile. As detailed in Table A12, many of the treatment effect estimates in the first price quintile are sensitive to whether they are estimated in a pooled regression of all price quintiles together (left panel, indicating an average PPI effect of 1.4 percent for the bottom price quintile) vs. separate regressions of each price quintile (right panel, indicating an average PPI effect of -0.7 percent for the bottom price quintile). The primary difference between these specifications is that the pooled regression imposes the same brand-year trends across hospitals in all quintiles of the price distribution, whereas the non-pooled version allows brand-year trends to vary based on hospitals' position in the price distribution. Thus the difference in the first quintile results

---

[33]Table A10 compares our preferred full sample specification with all price quintiles estimated simultaneously and hospital-brand and brand-year fixed effects (left panel), to the same specification estimated on a restricted sample of brands with ≥ 100 hospitals purchasing in the average year (right panel). Table A11 presents summary versions of the full sample (left panel) and restricted sample (right panel) regressions in Table A10, aggregated to the product class level, with alternative control sets. Table A12 compares our preferred full sample specification (left panel) to an alternative full sample specification with hospital-brand and brand-year fixed effects, but with each price quintile estimated in a separate regression (right panel).

may be due to differential brand-specific price trends for low-price hospitals. For example, if low-price hospitals are those that have relatively high bargaining parameters, it may be the case that those hospitals achieve high discounts regardless of changes in the market environment and accordingly exhibit less steep downward trends than other hospitals.

The broad robustness of the treatment effects among high-price (and high-price, high-quantity) hospital-brands to sample and specification choices leads us to focus on them as the primary effect of transparency that we find in this study. Thus, our main takeaway across all of these top categories is that high-price hospital-brands within PPIs experience 3.3 percent savings (or 3.9 percent savings among high-quantity hospital-brands) after obtaining access to benchmarking data; savings are limited for other hospitals and products. We offer these results with the caveat that we observe suggestive evidence of a positive externality among previously low-price hospital-brands for some PPIs, though this evidence is sensitive to specification, and the secondary hypothesis of larger increases of low prices in cases with low quantities is not borne out in any specification.

The final two columns in the table at the bottom of Figure 4 show the mean and standard deviation of expected annual savings across all hospitals. Savings are calculated based on the richest specification, with separate treatment effects for each hospital-brand based on its position in the pre-join price and quantity distribution. Savings are aggregated to the hospital-product category, then averaged over product categories for each class. The results reinforce that, on average, hospitals can expect modest savings of $1,869 on PPIs, but there is a large amount of heterogeneity across hospitals. A standard deviation improvement takes expected savings to $7,361 per hospital-product category-year. The favorable parts of the savings distributions for commodities and other surgical supplies offer substantial opportunities as well, with annual category savings of over two thousand dollars one standard deviation from the mean.

7   Conclusion

This paper conducts the first empirical study of the impact of transparency on price negotiations in business-to-business markets. Our empirical context is hospital supply purchasing, an area where there has been keen interest in information as a way to decrease hospital supply costs. Using new data on all purchase orders issued by over 17 percent of US hospitals from 2009-14 and difference-in-differences research designs, we find that hospitals that gain access to benchmarking information see subsequent savings on the brands for which they were previously paying relatively high prices. These estimates provide evidence on the potential economic impacts of the rise in benchmarking data services marketed towards buyers

27

in business-to-business markets. They also suggest that information is a potentially important driver of heterogeneity in negotiated prices, with implications for the growing structural empirical literature in bargaining.

Our tests of the mechanisms behind these information effects imply that the value of information is attenuated by the costs of putting the information to use. The evidence suggests that there are costs consistent with time-constrained negotiators (gains are focused in high-quantity items where the most money is at stake) and also stickiness of business-to-business contracts (long term contracts may not be renegotiated for some time). The latter friction is a fundamental feature of many business-to-business markets. However, the time and effort cost of accessing and/or using information could be reduced as technology improves. As both information and analytics are increasingly important in the modern economy, this suggests a path for future research.

We examined two potential theories for how benchmarking information might be used in a business-to-business setting – asymmetric information about seller bargaining parameters and buyer-side negotiator agency. We found robust evidence for the asymmetric information theory, but noisy evidence for agency. Within the asymmetric information framework, we find strong evidence that transparency leads to partial price convergence through decreases in the top of the price distribution. We also find some evidence that prices increase in the bottom of the distribution, though these are sensitive to empirical specification. We see modeling frictions in the use of information, the potential for information to interact with within-firm agency frictions in negotiation, and externalities in multilateral bargaining with asymmetric information as three especially interesting areas for future theory development.

While our results suggest that intermediaries who increase transparency may indeed lower the prices hospitals pay for a wide variety of medical supplies, our detailed analysis of mechanisms focuses on coronary stents. Variation across product markets in terms of supply side competition, complexity of contracts (for example, nonlinearities or multi-product bundling), and the particular mechanisms through which information affects prices thus represents rich opportunities for future empirical analysis of information in business-to-business bargaining. Such work would require expanding the empirical toolkit to analyze complex contracts when the contract terms themselves may not be observed.

Finally, while this paper takes a small step towards understanding the implications of information in business-to-business markets, more work is necessary to evaluate what we would expect as benchmarking information diffuses into wider use or in policy proposals for greater transparency in medical device markets. For example, a more structural model, combined with variation in market structure and information penetration, could explore the potential roles of supply side phenomena such as obfuscation or collusion.

References

Albek, S., Mollgaard, P., and Overgaard, P. (1996). Law-assisted collusion? The transparency principle in the Danish competition act. *European Competition Law Review*.

Ausubel, L., Cramton, P., and Deneckere, R. (2002). Bargaining with incomplete information. In Aumann, R. and Hart, S., editors, *Handbook of Game Theory*, volume 3, chapter 50, pages 1897–1945. Amsterdam: Elsevier Science.

Backus, M., Blake, T., Larsen, B., and Tadelis, S. (2018). Sequential bargaining in the field: Evidence from millions of online bargaining threads. NBER Working Paper No. 24306.

Backus, M., Blake, T., and Tadelis, S. (2016). On the empirical content of cheap-talk signaling: An application to bargaining. forthcoming in Journal of Political Economy.

Bandiera, O., Prat, A., and Valletti, T. (2009). Active and passive waste in government spending: Evidence from a policy experiment. *American Economic Review*, 99(4):1278–1308.

Binmore, K. (1985). *Game-theoretic Models of Bargaining*, chapter Bargaining and Coalitions. Number 13. Cambridge University Press.

Binmore, K., Rubinstein, A., and Wolinsky, A. (1986). The Nash bargaining solution in economic modelling. *The RAND Journal of Economics*, 17(2):176–188.

Binmore, K. G., Shaked, A., and Sutton, J. (1989). An outside option experiment. *Quarterly Journal of Economics*, 104:753–770.

Bronnenberg, B., Dube, J.-P., Gentzkow, M., and Shapiro, J. (2015). Do pharmacists buy Bayer? Informed shoppers and the brand premium. *Quarterly Journal of Economics*, 130(4):1669–1726.

Calsyn, M. and Emanuel, E. J. (2014). Controlling costs by expanding the Medicare Acute Care Episode demonstration. *JAMA Intern Med*, 174(9):1438–1439.

Card, D., Mas, A., Moretti, E., and Saez, E. (2012). Inequality at work: The effect of peer salaries on job satisfaction. *American Economic Review*, 102(6):2981–3003.

Collard-Wexler, A., Gowrisankaran, G., and Lee, R. S. (2017). Nash-In-Nash bargaining: A microfoundation for applied work. *forthcoming in Journal of Political Economy*.

Cooper, Z., Craig, S., Gaynor, M., and Van Reenen, J. (2018). The price ain't right? Hospital prices and health spending on the privately insured. *Quarterly Journal of Economics*. Accepted.

Craig, S. V., Grennan, M., and Swanson, A. (2018). Mergers and marginal costs: New evidence on hospital buyer power. NBER Working Paper No. 24926.

Cramton, P. (1992). Strategic delay in bargaining with two-sided uncertainty. *Review of Economic Studies*, 59:205–225.

Crawford, G. and Yurukoglu, A. (2012). The welfare effects of bundling in multichannel television. *American Economic Review*, 102(2).

Cutler, D. and Dafny, L. (2011). Designing transparency systems for medical care prices. *The New England Journal of Medicine*, 364:894–895.

Draganksa, M., Klapper, D., and Villas-Boas, S. B. (2009). A larger slice or a larger pie? An empirical investigation of bargaining power in the distribution channel. *Marketing Science*.

Dranove, D. and Jin, G. (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, 48(4):935–63.

Duggan, M. and Scott Morton, F. (2006). The distortionary effects of government procurement: Evidence from Medicaid prescription drug purchasing. *Quarterly Journal of Economics*, 121:1–31.

Ellison, G. and Ellison, S. (2009). Search, obfuscation, and price elasticities on the internet. *Econometrica*, 77(2):427–452.

Gowrisankaran, G., Nevo, A., and Town, R. (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *The American Economic Review*, 105(1):172–203.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review*, 103(1):145–177.

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

Grennan, M. and Swanson, A. (2018). Diagnosing price dispersion. Working paper.

Hastings, J., Neilson, C. A., and Zimmerman, S. D. (2015). The effects of earnings disclosure on college enrollment decisions. NBER working paper 21300.

Hendricks, K., Sorensen, A., and Wiseman, T. (2012). Observational learning and demand for search goods. *American Economic Journal: Microeconomics*, 4(1):1–31.

Ho, K. and Lee, R. S. (2017). Insurer competition in health care markets. *Econometrica*, 85(2):379–417.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, 13(2):324–340.

Horn, H. and Wolinsky, A. (1988). Bilateral monopolies and incentives for merger. *RAND Journal of Economics*, 19:408–419.

James, J. (2012). Health policy brief: Pay-for-performance. *Health Affairs*.

Jin, G. and Leslie, P. (2003). The effect of information of product quality: Evidence from restaurant hygiene grade cards. *Quarterly Journal of Economics*, 118(2):409–451.

Jindal, P. and Newberry, P. (2018). To bargain or not to bargain: The role of fixed costs in price negotiations. forthcoming in Journal of Marketing Research.

Kaplan, G. and Menzio, G. (2015). The morphology of price dispersion. *International Economic Review*, 56(4).

Larsen, B. (2018). The efficiency of real-world bargaining: Evidence from wholesale used-auto auctions. NBER Working Paper 20431.

Lazear, E. P. and Oyer, P. (2012). *The handbook of organizational economics*, chapter Personnel economics. Princeton University Press.

Lewis, M. and Pflum, K. (2015). Diagnosing hospital system bargaining power in managed care networks. *American Economic Journal: Economic Policy*, 7(1):243–74.

Lieber, E. M. J. (2017). Does it pay to know the prices in healthcare? *AEJ Economic Policy*, 9(1):154–179.

Maeda, J. L., Raetzman, S. O., and Friedman, B. S. (2012). What hospital inpatient services contributed the most to the 2001-2006 growth in the cost per case? *Health Services Research*, 47(5):1814–1835.

McConnell, K. J., Lindrooth, R. C., Wholey, D. R., Maddox, T. M., and Bloom, N. (2013). Management practices and the quality of care in cardiac units. *JAMA Internal Medicine*, 173(8):684–692.

Noether, M. and May, S. (2017). Hospital merger benefits: Views from hospital leaders and econometric analysis. Technical report, Charles River Associates.

Osborne, M. J. and Rubinstein, A. (1990). *Bargaining and markets.* Academic Press, San Diego, CA.

Pauly, M. V. and Burns, L. R. (2008). Price transparency for medical devices. *Health Affairs*, 27:1544–1553.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):97–110.

Rubinstein, A. (1985). A bargaining model with incomplete information about time preferences. *Econometrica*, 53(5):1151–1172.

Schneller, E. S. (2009). The value of group purchasing - 2009: Meeting the need for strategic savings. *Health Care Sector Advances, Inc.*

Scholten, P. A. and Smith, S. (2002). Price dispersion then and now: Evidence from retail and e-tail markets. *The Economics of the Internet and E-commerce*, 11:63–88.

Scott Morton, F., Silva-Rosso, J., and Zettelmeyer, F. (2011). What matters in a price negotiation: Evidence from the u.s. auto retailing industry. *Quantitative Marketing and Economics*, 9(4):365–402.

Shaked, A. and Sutton, J. (1984). The semi-Walrasian economy. Discussion paper 84/98 (theoretical economics), Suntory Toyota International Centre for Economics and Related Disciplines, London School of Economics.

Shelegia, S. and Sherman, J. (2018). Bargaining at retail stores: Evidence from Vienna. Working paper.

Sorensen, A. (2000). Equilibrium price dispersion in retail markets for prescription drugs. *Journal of Political Economy*, 108(4).

United States Food and Drug Administration (2018). Medical devices. https://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Overview/ClassifyYourDevice/de

Waldman, P., Armstrong, D., and Freedberg, S. P. (2013). Deaths linked to cardiac stents rise as overuse seen. *Bloomberg Business*.

Zettelmeyer, F., Scott Morton, F., and Silva-Risso, J. (2006). How the internet lowers prices: Evidence from matched survey and automobile transaction data. *Journal of Marketing Research*, XLIII:168–181.

ELECTRONIC APPENDICES – NOT FOR PRINT PUBLICATION

A   Data Appendix

The raw transactions data contain 82.5 million observations for 2,111 members across 3,375 product categories and 1.9 million SKUs. Our main analyses focus on coronary stents; our expanded analysis includes 52 other important product categories. For stents, we assign product IDs to each branded product by examining each manufacturer's catalog and searching for brand and model names within the free-text item description field in the transaction database. The procedure for assigning brand IDs in the expanded sample is detailed in Appendix A.3 below.

Table A1 presents details for our main analysis sample of facilities purchasing coronary stents. The top panel summarizes facilities purchasing coronary stents: the average sample facility submitted transactions in 41 months, spending \$3.4 million per month on 1,143 product categories. The bottom panel examines stent purchasing patterns by type: the average sample facility spent \$80,000 on 59 stents per month. 82 percent of stents purchased were drug-eluting, as opposed to the older, bare metal technology. Stents comprised approximately 2 percent of overall monthly spending.

Table A1: Summary Statistics of Stent Purchase Orders

	All Stents [N=508]		Drug-Eluting [N=508]		Bare Metal [N= 504]	
	Mean	SD	Mean	SD	Mean	SD
Months of Data	41.4	21.4	41.4	21.4	41.6	21.4
Product Categories	1143.1	313.4	1143.1	313.4	1147.2	308.2
Total Spend/Month (\$m)	3.4	3.2	3.4	3.2	3.4	3.2
% Hospital & Health System	0.97	0.17	0.97	0.17	0.97	0.18
Stent Spend/Month (\$k)	80.4	73.8	73.1	69.2	7.4	7.0
Stent Qty/Month	58.7	52.7	48.4	46.1	10.4	9.8
Stent Brands/Month	3.8	1.4	2.3	0.9	1.5	0.7
Stent Mfrs/Month	2.1	0.7	1.9	0.7	1.4	0.6

Authors' calculations from PriceGuide data.

Table A2 shows each statistic separately by hospital bed count; larger hospitals generally submitted more months' data and, as logic would indicate, purchased more stents per month for a greater total monthly expenditure. Hospitals with $\geq$ 500 beds spent more than double the amount that the smallest hospitals did on stents per month. The vast majority of transactions in our data are for drug-eluting (as opposed to bare metal) stents; in the remainder of our description of the data and in our results, we focus on drug-eluting stents.

34

Table A2: Summary Statistics – Stent Hospitals Only

Bed Size	Members	Months of Data	Monthly Exp. ($ k)	Monthly Quantity	% DES
0-99	52	31.4	59.0	45.0	82.0
		(19.4)	(56.6)	(44.4)	(10.8)
100-199	102	40.0	45.5	33.5	81.6
		(20.3)	(43.3)	(31.5)	(12.1)
200-299	117	43.4	55.6	40.7	77.4
		(22.0)	(45.9)	(33.5)	(14.3)
300-399	83	41.0	73.5	53.6	79.7
		(20.7)	(46.9)	(33.1)	(11.6)
400-499	47	41.4	128.9	93.5	79.6
		(21.3)	(92.1)	(65.5)	(12.2)
500+	107	45.9	135.2	97.7	81.1
		(22.0)	(94.2)	(65.6)	(9.5)

Authors' calculations from PriceGuide data.

The heterogeneity in prices observed in our sample is not well-explained by hospital characteristics that might seem a priori to be important for negotiation. For example, we observe no clear relationship between hospital size and stent prices. See the left panel of Figure A1, in which we display a box plot of drug-eluting stent prices for each category of bed count.[34] Mean prices are, if anything, increasing in bed count, though the differences are not statistically significant. Part of this (lack of) relationship could be due to the heterogeneity in purchasing behavior across hospitals with similar bed counts – e.g., small cardiac specialty hospitals may purchase stents in greater quantities than similarly-sized acute care hospitals. In the right panel of Figure A1, we also show box plots of stent prices for each decile of monthly stent purchasing volume. Here, we do see a relationship between "size" and price – the price distribution for hospitals with the smallest purchasing volumes is spread slightly upward relative to that of the hospitals with the largest volumes, so that low-volume hospitals' prices have larger means and variances than high-volume hospitals. For drug-eluting stents, 10th decile hospitals' prices are 7 percent lower than those obtained by 1st decile hospitals. These differences are economically and statistically significant; however, the price distributions for the high-volume and low-volume hospitals overlap substantially, so that there is a great deal of unexplained hospital price heterogeneity conditional on purchasing volume.

---

[34] "Prices" are hospital fixed effects obtained from a regression of price on hospital and brand-month fixed effects.

Figure A1: Distribution of Prices Across Hospitals

(a) By Bed Count   **(b)** By Stent Volume Decile

excludes outside values

	Drug-Eluting Stent Prices by Size Category (Regression Results)															
	$\beta \mathbb{1}\{h \in \text{bed size } x\}$						$\beta \mathbb{1}\{h \in \text{decile } x\}$									
$0-99$	$100-199$	$200-299$	$300-399$	$400-499$	$500+$	1	2	3	4	5	6	7	8	9	10
1,785	1,801	1,852	1,778	1,805	1,829	1,916	1,864	1,830	1,832	1,766	1,784	1,801	1,807	1,787	1,774
(38)	(27)	(30)	(25)	(32)	(22)	(54)	(40)	(44)	(36)	(26)	(31)	(29)	(25)	(25)	(32)

Authors' calculations from PriceGuide data. Estimated mean hospital fixed effects within bed size categories and decile of monthly purchase volume. Hospital fixed effects obtained from regression of price on hospital and brand-month fixed effects, pre-join data only. Mean estimates from regression of fixed effects on indicators for size. Standard errors from nonparametric bootstrap of entire procedure, resampling at hospital level.

A.1 Representativeness of the Benchmarking Database Sample

The facilities in the purchase order data voluntarily joined a subscription service that allows them to benchmark their own prices and quantities to those of other members in the database and thus may not be a random sample of US hospitals. In particular, subscription is costly, so we expect hospitals with greater concerns about supply costs to be overrepresented in the database – for example, in a survey of database members, "cost reduction on PPIs" and "cost reduction on commodities" were the first and second (and nearly tied) most commonly cited reasons for joining. This is in accord with our own conversations with purchasing managers who cite a broad array of reasons and product areas as motivations for benchmarking. Here, we discuss representativeness of sample facilities.

The left panel of Figure A2 compares the distribution of sample hospitals across US census regions to American Hospital Association (AHA) member hospitals with cardiac catheterization labs. The Figure also compares our sample to that in the Millennium Research Group's (MRG) Markettrack™ survey of catheter labs. The MRG survey is explicitly intended to

provide an accurate picture of market shares and prices by US region. The Figure shows that, relative to both comparison samples, the west region is overrepresented in the benchmarking database sample, while the south is underrepresented. The average sample hospital is also larger than the average US hospital with cardiac catheterization capabilities – the right panel of Figure A2 shows that the sample contains disproportionately fewer hospitals in the < 200 beds range and disproportionately more hospitals in the ≥ 500 beds range, relative to AHA hospitals that would purchase stents. We do not have access to bed size for the MRG sample, but we do find that the member facilities in our estimation sample purchased in significantly higher volumes than MRG sample hospitals (60 vs. 33 stents per month).

Figure A2: Distribution of Benchmarking Database vs. Comparison Hospitals

(a) Across Census Regions **(b)** By Bed Size

Database sample computations from PriceGuide data, 2009-2014. AHA sample computations from American Hospital Association Annual Survey of Hospitals, 2012; hospitals with catheterization labs defined as those listed as having in-hospital adult or pediatric interventional or diagnostic catheterization services. MRG sample computations from Millennium Research Group survey of catheter labs, January 2010-June 2013.

For January 2010 through June 2013, we also have access to stent price data from the MRG survey. This allows us to further check the representativeness of the sample of hospitals joining the price benchmarking database. Figure A3 provides further details on the full distribution of prices across hospitals in the two samples of 143 pre-join database sample hospitals vs. 107 MRG hospitals. The prices paid in the two samples are statistically close to one another, with the average prices paid (controlling for brand-time trends) in the MRG sample slightly higher (mean \$1,666, s.d. \$149) than those paid by hospitals in the estimation sample (mean \$1,631, s.d. \$120) during the period before they joined the benchmarking service. These pre-join hospitals do have a slightly larger upper tail of high prices, with an 80th percentile of \$1,743 versus \$1,730 in the MRG sample, but this

difference is not statistically significant. The representation of larger facilities with slightly better negotiation outcomes ex ante in our sample may be due to small hospitals' limited ability to afford access to the database, though we would expect a countervailing effect to come from large hospitals' greater ability to purchase custom benchmarking services from consulting firms.

Figure A3: Pre-Join Distribution of Prices Across Hospitals: Comparison to MRG Sample

	mean	sd	cv=$\frac{\mu}{\sigma}$	$20^{th}_{\%ile}$	$50^{th}_{\%ile}$	$80^{th}_{\%ile}$	N_H
Pre-join $\bar{p}_h$ (within brand-month)	1631	120	0.07	1539	1607	1743	143
MRG $\bar{p}_h$ (within brand-month)	1666	149	0.09	1563	1651	1730	107

Authors' calculations from PriceGuide data and MRG survey. Sample restricted to Jan. 2010-Jun. 2013 when MRG data available. Hospital average prices with brand-month trends removed.

Despite potential selection into our sample, the estimation strategy we employ in this paper is internally valid in that it exploits the existence of pre/post data and staggered join dates within the sample of joiners (and uses no non-joiners) to estimate the effect of access to benchmarking information. These estimates are of direct interest, capturing the benefit of benchmarking for facilities that seek out such services.

A.2 Information Flags

In order to construct each member's information set upon joining the database and later upon new brands' entry, we linked the transactions data with several additional datasets relating each individual login ID with associated members and login activity. The first of these is a "clickstream" dataset containing timestamped observations of unique IDs' login

activity, to the minute.[35] The second is a membership "hierarchy" file linking individual database members with parent accounts for those cases where members are part of a set of health care organizations purchasing membership jointly. The third file associates each login ID with the given individual's direct-linked member organization and broader access level – i.e., the individual with ID X may work with member 1 but have access to the data for all members 1, 2, and 3 under the same parent organization. For individuals with higher-level access, data for all associated members is automatically reported to them when they log in to the database. Accordingly, when we observed a click for a particular login ID, we associated that click with all linked member organizations for which that login had access.

We used the above-described datasets to determine each date of login for each member. The goal of this exercise was to determine when benchmarking data on a given brand would enter each member's information set. For non-entering stent brands and all product categories in the expanded analyses, this is the date of the first observed login for each member. For entering stent brands, this is the date of the first observed login for each member after six months post-entry. This is to account for the lag between transactions occurring for new brands and transactions being submitted by member facilities, loaded into the database, and viewed by members logging in. The left panel of Figure A4 displays the steady increase over time in the count of members for which transactions for the average entering brand are observed, and demonstrates the lag with which members' transactions become available in the database for benchmarking purposes.[36] In any given month in the year following new brand entry, there are on average 56 more members we observe having transactions for new brands than there are members whose transactions data are actually loaded into the database. We also observe transactions data for members that have not yet joined the benchmarking database – in the year following brand entry, 9 percent of members observed in the average month are pre-join. To see this more concretely, the right panel of Figure A4 displays the trend in the number of hospitals purchasing the average new brand, overall and for pre-join hospitals in particular. For each new brand, we observe 10-15 hospitals in the pre-join state within a short window after brand entry, and the time period for our study contains many meaningful brand introductions. This is precisely what allows us to separately identify effects of *joining* the database per se, versus actually having access to information *on a particular brand*, on prices. In Figure 2 in the main text, we note the timing of entry of seven new drug-eluting stent brands between 2010 and 2014 (of the thirteen brands sold

---

[35]Each login ID corresponds to a unique individual's account within a member. For example, a given database member may have had login accounts for a number of different purchasing managers, administrators, and department clinicians.

[36]There may be an additional lag before joining hospitals become informed if they do not frequently log in to the database.

during this time period overall).

Figure A4: Transactions Observed After New Brand Entry

(a) Cumulative Members Purchasing

(b) Members Purchasing, by Join Status

Authors' calculations from PriceGuide data.

We use the linked login and transactions data to calculate each member-brand's position in the pre-information price and quantity distributions. All calculations are specific to the first *informed* login for the given member-brand, as defined above. Following the approach used in the database to aggregate data across all members' transactions and present summary statistics to members logging in at a point in time, we calculate percentiles of the price distribution using all members' *most recent* transactions for the same brand, in the past year, that would have been loaded into the database prior to the login. We calculate percentiles of the quantity distribution using all members' total quantity purchased per month for the same brand in the past year. Across all specifications, we consistently include only those observations that can be used to estimate the richest specification with interactions based on point in the price and quantity distributions – this requires that we observe pre-information data for the given member-brand.

This estimation sample for drug-eluting stents includes 395,271 transactions for 508 members and thirteen brands in 72 months between January 2009 and December 2014. Seven of the included drug-eluting stent brands entered the market during this time horizon. We collapse the transaction-level data to perform all analyses at the member-brand-month level (with mean price as the dependent variable). We do this in order to avoid overweighting member-brands that tend to have multiple transactions per month. The regression sample used in our main analyses focused on drug-eluting coronary stents contains 32,453 member-brand-month observations.

40

A.3   Other Important Products

For our work in Section 6, we generated analysis files for a large number of additional product categories, defined by their UMDNS codes. We restricted to the top 50 categories by spending or number of transactions, yielding 80 categories total. From these, we excluded categories that were too broad (e.g., "food item").[37]   Next, creating the final analysis file required two key steps for each product category: (1) rationalizing the multiple units of measure in which different transactions' quantities were reported, in order to analyze price for a common quantity across transactions; and (2) generating brand IDs, in order to appropriately control for brand-specific price trends.

Regarding the first step, although many medical and surgical product categories are sold by the unit (a single coronary stent, e.g.), others are sold in pairs, boxes, cases, etc. The transactions data indicates this distinction in the "unit of measure" field, and further notes how many subunits are in each unit of measure using a "conversion factor" field. In order to perform our analyses on the cleanest and most internally-consistent transactions data possible, we transformed all transactions into price per single unit and quantity of single units purchased. We also excluded UMDNS codes with inconsistent or missing quantity data.[38]

Regarding the second step, the absence of a brand ID in the database creates a problem of sparsity, in which many SKUs are purchased by only a small number of hospitals, or in only a small number of months. The method we employed to identify brands within the coronary stents data involved examining manufacturer catalogs, finding likely brand names, searching for similar strings within the item description field, and validating SKUs for those brands against the catalog numbers. This was infeasible in the expanded analysis for several reasons: the average category has 30 manufacturers and more than 2,000 SKUs; many manufacturers' websites were found to be difficult to navigate, particularly once we extended the analysis beyond high-dollar physician preference items; and the item description field was often uninformative as to brand. Hence, we used an algorithmic approach to assign brand IDs for the other product categories.

The algorithm proceeded as follows: within each manufacturer, for each $k = 0, ..., K$, where K is the maximum SKU length for the manufacturer under consideration, we generated

---

[37]We did this based on "reasonableness" of the observed price variation – categories for which the coefficient of variation in price exceeded 10 were excluded – and selected categories by hand that seemed excessively broad based on their UMDNS names (e.g., "office supplies"). The list of product categories excluded by hand is 8889, 99936, 88885, 88884, 88883, 88695, 88539, 88311, 88073, and 16101.

[38]Specifically, we excluded UMDNS codes for which the conversion factor (e.g., ten units per box) was missing more than 1/3 of the time, or for which the modal unit of measure (e.g., "box" vs. "case") accounted for less than 50 percent of the data.

a candidate brand ID as the left-most k digits of each SKU. We then regressed price on dummies for each unique candidate brand ID, interacted with each hospital ID. We recorded the resulting R^2 for each k. The goal of this exercise was to identify brands that generate price variation in the data: looking within a given hospital, price variation proxies for contract variation. The R^2 is always increasing in k, so the algorithm is meant to choose k^{opt} that maximizes R^2 while keeping the number of identified "brands" relatively manageable. Table A3 below compares the algorithmic approach to the time-intensive hand-coding approach for coronary DES stents and for one manufacturer of surgical staplers whose catalog was available online. The first three rows show the number of brands implied by unique SKUs, vs. the hand-coding of manufacturer catalogs, vs. the algorithm. We also show the R^2 of a regression of price on hospital-brand dummies. The remaining rows show the detailed brand count and R^2 for each SKU digit count k. For the 52 products in our expanded analysis, we generally found that R^2 would climb steeply, then level off after a few SKU digits. For example, within stent manufacturer 1, R^2 increases from 0.42 at one SKU digit to 0.6 at three SKU digits, but only increases to 0.65 thereafter. Similarly, within linear surgical staplers, R^2 climbs from 0.61 to 0.95 as we move from one to three digits, eventually reaching a maximum of 0.97. These patterns are consistent with our intuition that the first few SKU digits would indicate a particular product line, while the last few SKU digits would indicate finer (and likely unpriced) distinctions such as color or size.

The machine learning literature has many methods for explicitly penalizing excessive complexity in classification and other modeling exercises. We chose $k^{opt} = min\{k|R^2(k) > 95\%*R^2(K)\}$; that is, we erred on the side of allowing for more sparsity and less classification error. For stents, this implies about ten times as many algorithmically-identified "brands" than our hand-coding would indicate; for staplers, the algorithm identifies 47 products while hand-coding indicates 53. Across all product category-manufacturer combinations in Table A3, the algorithmic brands explain as much price variation as the hand-coded brands, and mitigates the sparsity problem significantly. For stents (staplers), the number of algorithmic brands is 88 percent (49 percent) smaller than the number of unique SKUs.

The goal of our expanded analyses is to determine whether the patterns documented for stents are similar to those we observe for other important purchase categories. To that end, we employed the same regression specifications as in the stents analyses: (1) We estimate the average treatment effect of joining the database, for which we find no significant result in stents. (2) We estimate the average treatment effect of joining the database among all high-price hospital-brands. (3) We estimate the average treatment effect of joining the database specifically among high-price, high-quantity hospital-brands. For the latter two analyses, we find economically and statistically large effects in stents. We estimate treatment effects of

Table A3: Performance of Brand-Assignment Algorithm

| | Stents $[N_m = 4]$ | | | | | | | | Staplers $[N_m = 1]$ | | | |
| | Mfg 1 | | Mfg 2 | | Mfg 3 | | Mfg 4 | | Surgical | | Surgical Linear | |
	$|J|$	R^2	$|J|$	R^2	$|J|$	R^2	$|J|$	R^2	$|J|$	R^2	$|J|$	R^2
SKU	261	.65	510	.87	47	.6	350	.73	53	.99	40	.97
Hand-coded	5	.66	5	.82	1	.54	2	.69	29	.99	24	.96
Algorithm	47	.62	54	.84	23	.58	11	.69	20	.95	27	.95
$k = 1$	1	.42	1	.42	1	.54	4	.69	8	.84	7	.61
$k = 2$	2	.44	2	.73	2	.54	6	.69	20	.95	15	.7
$k = 3$	5	.6	5	.82	2	.54	8	.69	31	.98	27	.95
$k = 4$	5	.6	8	.82	8	.56	9	.69	40	.99	31	.95
$k = 5$	13	.6	14	.83	12	.56	11	.69	51	.99	38	.97
$k = 6$	29	.61	54	.84	23	.58	23	.69	52	.99	40	.97
$k = 7$	47	.62	112	.85	46	.6	43	.69	53	.99	40	.97
$k = 8$	137	.63	271	.86	47	.6	91	.7	53	.99	40	.97
$k = 9$	261	.65	510	.87	47	.6	129	.71	53	.99	40	.97

join only: due to the lack of precise brand data across all product categories, we do not use brand entry as an additional information shock.

A.3.1 Estimation Sample for Other Important Products

Table A4 shows the effect of each sampling restriction described above on the composition of facilities in our data. The first column summarizes the full set of 2,110 facilities in our data, the second shows the set of facilities purchasing the 52 product categories of interest, the third removes facilities missing conversion factor data or purchasing in non-standard units of measure, and the fourth removes facilities missing "brand" data and/or login data. The fifth column summarizes the final regression sample of facilities for which we were able to flag hospital-brands' positions in the pre-join price and quantity distribution. At each restriction, we tend to remove small facilities and non-hospitals. The final regression sample contains 775 facilities, 73 percent of which were hospitals or health systems, spending $1.8 million per month on 774 product categories.[39]

Figure A5 below compares the distribution of the expanded sample of database members to the general acute care hospitals in the AHA survey data. The patterns presented above

---

[39]One might expect "Restriction 1" to remove only a small number of facilities, as all health care facilities would be expected to purchase commodity products. The removal of a large number of facilities in that cut is due to our reliance on the UMDNS code to identify product categories – if some transactions in a facility's materials management database were recorded in such a way that UMDNS codes could not readily be assigned upon database upload, then it will likely not appear in our analytic sample.

Table A4: Summary Statistics from Purchase Order Database

	Full Sample [N=2,110]		Restriction 1 [N= 1,605]		Restriction 2 [N=1,491]		Restriction 3 [N= 1,388]		Regression Sample [N= 775]	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Months of Data	31.2	21.2	34.5	21.1	35.0	21.2	36.7	20.9	41.4	21.4
Product Categories	462.1	502.7	605.5	496.2	649.3	487.8	670.1	488.3	774.0	487.9
Total Spend/Month ($m)	1.1	2.7	1.5	3.1	1.6	3.1	1.6	2.6	1.8	2.7
% Hospital & Health System	0.48	0.50	0.61	0.49	0.65	0.48	0.67	0.47	0.73	0.44

Authors' calculations from PriceGuide data, 2009-2014.

for stent purchasers are largely present here as well: database members in our sample are underrepresented in the south and somewhat larger than the average US general acute care hospital.

Figure A5: Distribution of Benchmarking Database vs. Comparison Hospitals – All Facilities Purchasing "Important" Products

(a) Across Census Regions

(b) By Bed Size

Database sample computations from PriceGuide data, 2009-2014. AHA sample computations from American Hospital Association Annual Survey of Hospitals, 2012; includes all general acute care hospitals.

The hospital-brand-month level regression sample data are summarized for each product category in Table A5. The Table also shows aggregates at the product class level.[40] All

---

[40]As discussed in Section 6, Class 1 is "commodities": relatively inexpensive products that are unlikely to be chosen primarily by physicians; e.g., surgical gloves. Class 3 is "physician preference items": high-tech medical devices, mainly coronary and orthopedic products that are the primary implanted device in their corresponding surgical procedures; e.g., coronary stents. Class 2 is intermediate: other medical/surgical products used during invasive procedures, but explicitly excluding PPIs. Note that each row's data is within the set of hospitals with nonzero purchase in that category or class; i.e., average spending on bandages is within 481 hospitals whose bandage transactions we observe, whereas average spending on all commodities will include 267 hospitals whose bandage transactions we do not observe. As noted previously, the fact that we do not observe all facilities purchasing every commodity product is due to our reliance on the UMDNS code to identify product categories – if some transactions in a facility's materials management database were

44

together, top commodities account for 2 percent of hospital spend, vs. 21 percent for top PPIs, and 13 percent for other medical/surgical products. This is consistent with the large discrepancies in their pricing: the average unit price of a commodity is \$74, vs. \$2,032 for PPIs. Interestingly, we see similar degrees of price dispersion for commodities as we do for other medical/surgical products and PPIs – coefficients of variation across hospitals are 0.195, 0.166, and 0.188, respectively – hinting that physician preference alone cannot account for observed dispersion. Finally, note that the drug-eluting stent sample in Table A5 differs from the sample used in the main analyses in the draft because of the focus on join timing for identification of information effects: hospital-brands such that the brand enters after the hospital joins the database have no pre-information data and thus fall out of the analysis. However, average monthly spend and quantity for this subset of stent-purchasing hospitals is similar to our main stent sample.

B Mapping of Bargaining Setting into Models of Asymmetric Information and Agency

Here, we elaborate on the concise theory sections of the main text.

B.1 Asymmetric Information about Supplier Bargaining Parameters

We follow Rubinstein (1985) to model uncertainty among hospital buyers about the bargaining parameter of a given supplier. The model departs from the complete information model in Rubinstein (1982) in that the supplier is either of weak type with discount factor δ_w^S or strong type with discount factor δ_s^S ($1 > \delta_s^S > \delta_w^S > 0$). The supplier knows his own type, but the buyer has only a subjective prior ω of the probability that the supplier is the weak type.

The equilibrium split of this surplus depends on both the type of the supplier and the prior of the buyer as follows: Rubinstein (1985) shows that there exists a cutoff ω^* such that if the buyer is sufficiently pessimistic about the seller being the weak type ($\omega < \omega^*$), then the buyer simply offers what she would offer the strong type in the complete information game of Rubinstein (1982):

$$p^{AI}(\omega < \omega^*) := c + \delta_s^S \frac{1 - \delta^B}{1 - \delta^B \delta_s^S} V, \tag{2}$$

and both seller types accept this offer. However, if the buyer is more optimistic about the

recorded in such a way that UMDNS codes could not readily be assigned upon database upload, then it will likely not appear in our analytic sample.

45

Table A5: Summary of important product categories

ID	Name	Class	$\overline{Spend}_{hy}$ %	$\overline{Spend}_{hy}$ \$	N_{jht}	N_h	N_j	$\overline{Q}_{hy}$	$\overline{p}_{hjt}$	$CV(p_{h\|jt})$
10279	Bandages Elastic	1	0.0	11,237	23,216	481	81	3,176	3	0.226
11315	Dressings	1	0.2	44,105	42,514	581	151	10,492	10	0.238
11883	Gloves Surgical	1	0.6	107,187	51,660	579	49	48,532	2	0.158
12368	Drapes Surgical	1	0.2	47,462	76,910	565	171	7,235	7	0.191
12745	Needles Injection Hypodermic	1	0.1	22,351	66,170	618	61	100,657	0	0.156
13655	Containers Specimen	1	0.2	43,895	28,761	496	88	49,321	3	0.333
923840	Drill Bits	1	0.5	91,421	227,351	573	1,192	610	162	0.166
	Commodities	1	1.9	277,910	516,582	748	1,793	166,520	74	0.195
10688	Catheters Vascular Angiography	2	0.1	38,249	46,764	420	144	982	27	0.126
10894	Clip Appliers	2	0.5	118,993	26,179	499	18	629	118	0.170
11502	Forceps Electrosurgical	2	0.4	108,463	10,689	302	113	137	1,770	0.115
11910	Grafts Bone	2	1.3	377,912	34,324	471	290	136	2,638	0.124
11925	Guide Wires	2	1.1	192,401	189,814	592	859	1,528	82	0.158
12830	Orthopedic Cement	2	0.5	110,897	21,954	466	85	306	341	0.199
13050	Plates Bone	2	1.4	300,071	198,059	570	3,205	411	724	0.158
13909	Sutures Synthetic Nonabsorbable Polypropylene	2	0.2	42,126	98,186	572	300	3,950	15	0.140
13910	Sutures Natural Nonabsorbable Silk	2	0.1	19,211	54,640	502	73	4,532	4	0.262
14085	Tubes Tracheal	2	0.2	34,726	34,323	525	102	4,428	7	0.217
16071	Lenses Intraocular Posterior Chamber	2	0.1	76,046	15,621	268	184	495	161	0.092
16078	Nails Bone	2	0.6	133,181	51,900	482	457	96	1,424	0.154
16104	Wires Bone	2	0.1	22,144	41,367	535	376	214	110	0.189
16655	Catheters Cardiac Ablation	2	0.4	187,729	17,711	256	85	165	1,020	0.127
17184	Catheters Vascular Angioplasty Balloon	2	1.1	276,092	139,972	449	641	1,202	291	0.131
17471	Sutures Synthetic Absorbable Polyglactin	2	0.4	69,078	78,022	605	51	14,241	9	0.587
17600	Bone Matrix Implants	2	0.5	93,743	28,302	475	266	137	751	0.179
17846	Catheters Vascular Guiding	2	0.6	105,814	71,039	430	380	1,169	157	0.177
18609	Catheters Cardiac Mapping/Ablation	2	0.6	208,145	10,742	208	134	173	1,418	0.123
20317	Staplers Surgical	2	0.6	140,430	19,993	400	80	421	355	0.216
20318	Staplers Surgical Linear	2	0.4	110,985	21,776	423	47	635	150	0.245
20453	Stent/Grafts Vascular Aortic	2	1.0	303,258	11,376	290	130	54	5,581	0.050
22538	Suture Anchors	2	0.5	95,811	31,637	526	217	223	319	0.155
22908	Sutures Synthetic Nonabsorbable Nylon Monofilament	2	0.1	15,805	85,860	596	179	3,425	6	0.181
34196	Screws Bone Spinal	2	0.5	296,353	4,265	67	442	545	695	0.198
	Other Med/Surg	2	13.2	1,981,568	1,344,515	701	8,858	29,572	415	0.166
12324	Lenses Intraocular	3	0.1	44,309	11,372	279	121	301	365	0.138
12913	Pacemakers Cardiac Implantable	3	1.7	543,311	29,907	392	126	107	4,316	0.136
15766	Orthopedic Internal Fixation Systems Spinal	3	0.9	338,861	62,248	392	1,256	566	628	0.237
15870	Prostheses Cardiac Valve Biological	3	1.0	463,734	13,737	249	37	301	5,699	0.131
16040	Stents Ureteral	3	0.1	30,037	29,097	480	80	218	137	0.117
16084	Prostheses Joint Hip Acetabular Component	3	0.8	247,393	61,747	457	621	200	1,294	0.242
16095	Prostheses Joint Hip Femoral Component	3	1.3	375,090	97,347	467	998	236	1,568	0.244
16097	Prostheses Joint Knee Femoral Component	3	1.5	405,992	61,277	461	553	197	2,474	0.189
16098	Prostheses Joint Knee Tibial Component	3	1.3	364,878	110,185	469	1,304	307	1,259	0.201
16111	Prostheses Mammary Internal	3	0.6	166,975	37,651	373	223	211	815	0.111
16921	Prostheses Joint Shoulder Humeral Component	3	0.4	118,044	21,559	394	273	52	2,424	0.185
17165	Allografts	3	0.6	157,176	27,301	483	608	94	1,638	0.155
17241	Stimulators Electrical Spinal Cord Analgesic	3	0.8	310,908	5,693	284	19	23	15,249	0.104
18253	Grafts Skin Biological	3	0.4	135,230	6,857	342	65	201	2,916	0.087
18504	Defrib/Pacemakers Implantable	3	2.0	851,354	7,212	280	180	51	16,415	0.125
20376	Defrib/Pacemakers Implantable Resynchronization	3	2.3	1,007,055	5,331	264	109	45	21,199	0.122
20383	Stents Vascular Coronary Drug-Eluting	3	2.9	1,132,582	15,529	336	13	606	1,623	0.080
20422	Stents Vascular Coronary Balloon-Expandable	3	0.3	142,779	46,854	334	200	157	720	0.110
22543	Occluders Vascular Intravascular Embolization Coil	3	0.4	112,271	17,407	304	319	179	769	0.101
923150	Orthopedic Fixation Systems Implantable Spine	3	1.5	476,814	35,243	362	975	148	3,235	0.177
	PPIs	3	20.9	4,314,551	703,554	601	8,080	2,298	2,032	0.188

probability that the seller is the weak type ($\omega \geq \omega^*$), then the buyer offers:

$$p_w^{AI}(\omega \geq \omega^*) := c + \delta_w^S \frac{1 - \delta^{B^2}(1 - \omega) - \delta^B \omega}{1 - \delta^{B^2}(1 - \omega) - \delta^B \delta_w^S \omega} V, \tag{3}$$

which the weak seller type accepts. The strong seller type will reject this offer, and counteroffer with a price that would make a weak seller no better off than p_w^{AI}, but that the strong seller strictly prefers:

$$p_s^{AI}(\omega \geq \omega^*) := c + \frac{1 - \delta^{B^2}(1 - \omega) - \delta^B \omega}{1 - \delta^{B^2}(1 - \omega) - \delta^B \delta_w^S \omega} V, \tag{4}$$

which the buyer accepts.

B.1.1 Effects of transparency: complete information assumption

This equilibrium has direct implications for what we would expect to happen to prices in a move from this type of asymmetric information to complete information. First, note that $p_s^{CI} = p^{AI}(\omega < \omega^*) > p_s^{AI}(\omega \geq \omega^*) > p_w^{AI}(\omega \geq \omega^*) > p_w^{CI}$ (where p_s^{CI} is the equilibrium price for the strong supplier type, and p_w^{CI} the weak type, with complete information). Thus the weak type seller is strictly better off with asymmetric information. The strong type seller is weakly worse off (strictly when the buyer's prior is sufficiently optimistic). A sufficiently pessimistic buyer is also weakly worse off with asymmetric information. For more optimistic buyers, whether complete information would make them better off ex-ante depends on parameter values.

In our context, we are interested in when a buyer might benefit from benchmarking information that reveals the seller's type, and what would happen to price in such a case. The intuition for how this unfolds in practice is a scenario where a manufacturer sales representative says "This is the best price I can offer. Corporate won't let me go any lower." Benchmarking information allows the hospital negotiator to perform the due diligence of checking the prices at other hospitals in order to verify or refute this statement.

We begin by assuming that this information *fully* reveals a seller's type.

Prediction 1 (Direct Effect of Complete Information on High Prices) If information is costless, pessimistic buyers will always become informed. This information will cause a proportion of the highest prices p_s^{CI} to fall to p_w^{CI} for those cases where the supplier was in fact the weak type. Thus exposure to benchmarking information should lead to some of the highest prices falling.[41]

---

[41]The Rubinstein (1982, 1985) framework with bilateral negotiations and no externalities across buyers

Prediction 2 (Direct Effect of Complete Information on High Prices with High Quantity)

If information is costly to obtain (in the sense that searching and analyzing the data takes time that could be used on other productive activity), a pessimistic buyer will become informed whenever the expected benefit $\omega(p_s^{CI} - p_w^{CI})q$ exceeds the cost of information. This information will cause a proportion of the highest prices p_s^{CI} to fall to p_w^{CI} for those cases where the supplier was in fact the weak type. Thus exposure to benchmarking information should lead to some of the highest prices falling, among those brands with the highest quantity used.

B.1.2 Effects of transparency: supplier response to information externality across buyers

The above illustrative model and Propositions assumed that benchmarking information fully reveals the seller's type. However, the extent to which the seller's type is revealed by benchmarking depends on how information on *other buyers' prices* alters a given buyer's beliefs. Moreover, it is intuitive to think that the transparency introduced by price benchmarking could have a countervailing effect, in that sellers may be less willing to offer low prices if they know those prices will then be included in the database and potentially hurt them in subsequent negotiations with other buyers (Duggan and Scott Morton 2006; Grennan 2013). To show how this effect might arise, we extend the model to allow for a stochastic information revelation process in a multi-period, multi-lateral bargaining environment.

Consider a focal supplier negotiating with two buyers: the supplier negotiates the price p_1 with buyer 1 (valuing the product at V_1) in the first period, and p_2 with buyer 2 (valuing the product at V_2) in the second period. Buyer 2 has prior belief ω_2^{prior} that the supplier is of the weak type ($\delta^S = \delta_w^S$). If there is transparency in the form of benchmarking, there is an exogenous probability ϵ that p_1 will be revealed to the second buyer.[42] For simplicity

---

and sellers is helpful in that it motivates a reduced form empirical strategy such as the one employed in our empirical analyses. If we were to impose a model of bargaining under asymmetric information and estimate it structurally, an extension of a multilateral bargaining framework (e.g., Horn and Wolinsky (1988); Collard-Wexler et al. (2017)) would be more appropriate. Given that we explore multiple models via which transparency may affect prices, we do not proceed down this route, though we do propose a simple multilateral bargaining extension with several buyers in Appendix B.1.2 below. Finally, Appendix D.3 explores whether, in a multilateral bargaining setting with multiple imperfect substitute products, a fall in price of substitute brand j will lead to a decrease in the prices of other brands $-j$, all else equal. We did not find any substantial effect of this type, which is consistent with the findings in Grennan (2013) and Grennan and Swanson (2018) of low price sensitivity in demand for stents and other devices.

[42]In a reduced form sense, this model is intended to allow for uncertainty as to whether a seller's type is revealed, via his price, to the second buyer. This uncertainty could manifest in several ways: posterior uncertainty about the first buyer's bargaining parameter, posterior uncertainty about the first buyer's valuation, or other stochastic shocks. For example, a model that allows for posterior uncertainty in buyers' beliefs after receiving benchmarking information – with heterogeneity in buyers' valuations $V_j \sim F[0,\infty)$

48

of exposition, we suppose that the negotiation between the supplier and buyer 1 is under complete information.[43] The model parameters $(\delta^B, \delta_w^S, \delta_s^S, \omega_2^{prior}, V_1, V_2, \epsilon)$ are known to all parties, but the realization δ^S is known only to the seller and the first buyer.

Define the period-2 price expected by the type-z seller ($z \in \{w, s\}$), given buyer 1 price $p_1^* \in \{\emptyset, p_1\}$, as

$$
p_2^z(p_1^*) := \begin{cases} c + \delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_2 & \omega_2^{post}(p_1^*) < \omega_2^* \\ c + \delta_w^S \frac{1-\delta^{B^2}(1-\omega_2^{post}(p_1^*))-\delta^B \omega_2^{post}(p_1^*)}{1-\delta^{B^2}(1-\omega_2^{post}(p_1^*))-\delta^B \delta_w^S \omega_2^{post}(p_1^*)} V_2 & \omega_2^{post}(p_1^*) \geq \omega_2^*, z = w \\ c + \frac{1-\delta^{B^2}(1-\omega_2^{post}(p_1^*))-\delta^B \omega_2^{post}(p_1^*)}{1-\delta^{B^2}(1-\omega_2^{post}(p_1^*))-\delta^B \delta_w^S \omega_2^{post}(p_1^*)} V_2 & \omega_2^{post}(p_1^*) \geq \omega_2^*, z = s. \end{cases}
$$

As before, if buyer 2 is sufficiently pessimistic, the seller receives the strong-type complete-information price; if buyer 2 is optimistic, the seller receives her type-specific asymmetric-information price in a separating equilibrium given ω_2^{post}. Buyer 2's posterior belief may be a function of an observed price ($p_1^* = p_1$), or may update from ω^{prior} based on the fact that no p_1 was observed ($p_1^* = \emptyset$). If p_1 perfectly reveals the seller's type, then the seller receives her type-specific complete-information price in period 2.

To illustrate how an externality across buyers can generate distortions in buyer 1's price, we proceed constructively and present a simple numerical example. Let $c = 0$, $V_2 = 100$, $\omega_2^{prior} = 0.2$, $\delta^B = 0.8$, $\delta_s^S = 0.9$, and $\delta_w^S = 0.7$. In this model, we would observe period-2 pooling under the assumed prior and under any posterior belief ω_2^{post} below the cutoff $\omega_2^* = 0.43$. For simplicity, we consider a case here where in all equilibrium outcomes $\omega_2^{post} < \omega_2^*$, unless the seller's type is fully revealed (we assume this for the moment and verify later).

Observe that in this case, the strong-type seller will always receive the strong-type complete-information price $\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_2$ in period 2, either because period-1 prices fully reveal types, or because pooling occurs. Thus, there is no externality of the period-2 negotiation on the period-1 negotiation for the strong-type seller.

However, it is costly to the weak-type seller for her price to be revealed in the period-2 negotiation, and this presents a credible threat point for breaking off negotiations with buyer 1. The period-1 negotiation with the weak-type seller is equivalent to a Rubinstein alternating-offers bargaining game with the added feature that the seller has access to a valuable outside option. The weak-type period-1 game then has a unique subgame perfect

---

and Bayesian updating of buyer beliefs – generates similar predictions to our chosen model extension, but with much greater complexity.

[43]This would be the case, for example, if the seller's type had been revealed to buyer 1 via a separating equilibrium in a prior negotiation. Alternatively, one might think of this as a stylized model with V_1 representing a measure of informed buyers with access to benchmarking, V_2 a measure of buyers not subscribing to benchmarking, and ϵ the rate at which the second type of buyers are joining the benchmarking group.

equilibrium with immediate agreement:

$$
p_1^w := \begin{cases} \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} V_1 & \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} V_1 \geq \epsilon \left(\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} - \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} \right) V_2 \\ \epsilon \left(\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} - \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} \right) V_2 & \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} V_1 < \epsilon \left(\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} - \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} \right) V_2 < \delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_1 \\ \delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_1 & \text{otherwise.} \end{cases}
$$

$$(5)$$

This is essentially a modification of the "outside option principle" presented in Osborne and Rubinstein (1990), subsection 3.12.1.[44] The relevant outside option is that the weak seller may walk away from the period-1 negotiation. If she does so, the presence of the ϵ-shock implies that her type cannot be perfectly revealed, though the period-2 buyer will update his belief that she is of the weak type. Specifically, if buyer 2 does not observe a period-1 price, he will infer that either a weak-type seller strategically refrained from negotiating, or the stochastic transparency shock was such that prices were exogenously not revealed. By Bayes' rule, he will update his belief that the seller is a weak type from $\omega_2^{prior} = 0.2$ to $\omega_2^{post} = \frac{\omega_2^{prior}}{1-\epsilon+\epsilon\omega_2^{prior}} = 0.33$. At these parameter values, the equilibrium in period 2 will involve pooling under both ω^{prior} and ω^{post}.[45] Thus, walking away from the period-1 negotiation gives the weak seller the outside option payoff of $\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_2$ across both periods.

In the first case in the above equilibrium, the weak seller's type is fully revealed if p_1 is observed by buyer 2, but the outside option payoff of $\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_2$ (across both periods) if she walks away is nevertheless small relative to her equilibrium payoff across both periods if there were no outside option: $\delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} V_1 + (\epsilon \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} + (1-\epsilon) \delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S}) V_2$. Opting out is thus not a credible threat for the seller, and prices are as in the complete-information Rubinstein (1982) game.

In the second case, the outside option binds, and the weak seller must be compensated for the negative externality of the period-1 negotiation on the surplus she expects to receive in the period-2 negotiation: $\epsilon \left(\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} - \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} \right) V_2$. The period-1 price increases above Rubinstein (1982) levels to compensate the weak seller for reduced profits in the period-2 negotiation.

In the final case, buyer 1's value V_1 is sufficiently low that the strong seller's price drops below the weak seller's "outside option" price $\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_1 < \epsilon \left(\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} - \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} \right) V_2$. When that occurs, the weak seller prefers to accept the lower price, equal to the strong type

---

[44] Osborne and Rubinstein (1990) present a model in which the seller has the option of terminating the negotiation and receiving her outside option after she has rejected an offer. In this case, the game has a unique subgame perfect equilibrium as characterized above. This model was variously suggested by Binmore, Shaked, and Sutton; see Shaked and Sutton (1984); Binmore (1985); Binmore et al. (1989).

[45] This confirms that the strong-type seller has no incentive to deviate from the complete-information Rubinstein (1982) price in period 1, and in turn that $\omega_2^{post} = 0.33$ is consistent with Bayesian updating.

seller's price, as it allows her to pool with the strong seller if price information is revealed in period 2, and she receives a strong-type seller price in both periods. Buyer 1 also prefers to pay the lower price, supporting this equilibrium.

In Figure A6 below, we present equilibrium prices for the strong- and weak-type seller in the absence of the externality (i.e., if period-1 prices are never observed by period-2 buyers: $(p_1^{s,NoExt}, p_1^{w,NoExt})$) and with the information externality ($\epsilon = 0.5 : (p_1^{s,InfoExt}, p_1^{w,InfoExt})$). All model parameters are held fixed except V_1; the Figure shows equilibrium prices for each $V_1 \in [0, 100]$.

Figure A6: Asymmetric Information with Externality – Numerical Example

Absent the externality, strong sellers' prices are uniformly above weak sellers' prices at each value of V_1, and each increases less than one-for-one with buyer 1's valuations as a function of their complete-information bargaining splits. For the strong seller, the $NoExt$ and $InfoExt$ prices are equivalent ($p_1^{s,InfoExt} = p_1^{s,NoExt}$), as benchmarking does not impact the strong seller's prices. However, for the weak seller, the outside option binds for $V_1 \leq 51$, and prices in period 1 are higher with the information externality than without. For $V_1 \in [26, 51]$, $p_1^{s,NoExt} < p_1^{s,InfoExt} = 16.23$. And for $V_1 < 26$, the strong and weak type prices will be equivalent: $p_1^{w,InfoExt} = p_1^{s,InfoExt}$. That is, for sufficiently low V_1, the weak seller will accept a price below 16.23, as it increases her period-2 payoff by allowing her to pool with the strong seller.[46]

---

[46]There is a discontinuous drop in the weak seller's equilibrium payoff across both periods as buyer 1's value crosses 26. This is because buyer 2's knowledge of V_1 makes $p_1^{s,InfoExt}$ the unique price that allows pooling when period-1 prices are revealed. If the strategy space were expanded to allow hidden lump-sum transfers between the period-1 negotiating parties, this would break the informativeness of period-1 prices, and the weak seller would prefer to give buyer 1 hidden lump sum rebates in exchange for higher unit prices at higher values of V_1.

In this stylized model, the presence of the buyer-2 negotiation effectively puts a lower bound on p_1: whenever the standard Rubinstein (1982) allocation would make the seller prefer to break off negotiations with buyer 1, p_1 must at least be large enough to make the weak seller indifferent. Thus – via this externality created by the benchmarking information – some of a seller's lowest prices may increase. The likelihood and magnitude of this externality will depend on parameters, and in particular on the relative magnitudes of V_1 and V_2.[47]

The addition of an externality across negotiations in this extended model implies potential "convergence" of prices from both the top and the bottom – the highest prices will tend to decrease as information is revealed (the buyer 2 effect here), and the lowest prices will tend to rise as suppliers take into account their externalities on other buyers (the buyer 1 effect here).

Prediction 3 (Effect of Information Externality Across Buyers on Low Prices) Via the benchmarking process, a negotiated price with one buyer becomes information available to other buyers, potentially introducing an externality in future negotiations. If this externality is large enough, it can provide the seller a credible threat not to accept low prices. Thus some prices may rise with the introduction of benchmarking information, particularly among buyers with relatively low prices and a relatively low quantity under negotiation.

B.2   Negotiator Agency

The other candidate mechanism via which we propose benchmarking information could be valuable to buyers would be through providing aggregate information to help the buying firm solve a moral hazard problem with its purchasing agent. Here we provide a specific model of information in the bargaining context that generates predictions in our setting. Modifying Holmstrom (1982) to our context, let price p_h at hospital h be as in the full information Rubinstein (1982) bargaining game. However, instead of the hospital negotiator's bargaining parameter being exogenous, the price will be a function of the hospital agent's *choice* of discount factor δ_h^B and the discount factor of the supplier, which takes value $\delta_w^S \epsilon_h$ with probability ω and $\delta_s^S \epsilon_h$ with probability $1 - \omega$. As before, the discount factor of the strong supplier type is greater than that of the weak type $(1 > \delta_s^S > \delta_w^S > 0)$. ϵ_h is a random term distributed uniform on $[0, 1]$. Importantly, the realization of ϵ_h is independent across hospital buyers, but whether the seller is weak or strong is common to all buyers. The realizations of both of these random variables are observable to the negotiating agents, but not to the principals who manage them.

---

[47]Note that in our empirical context, the payoff to the seller in a given negotiation scales with the buyer's willingness-to-pay, as well as the quantity under negotiation.

A moral hazard problem arises in this setting because bargaining effort is costly and provides the agent disutility $v(\delta_h^B)$. The agent is compensated by some contract based on the price $m(p_h)$. The agent is risk averse in money, so the optimal solution to the agency problem involves risk sharing between the principal and the agent. Holmstrom (1982) shows that if agents face some common parameter which is uncertain from the principals' perspectives, then relative performance evaluation compared to some aggregate sufficient statistic can be used to write a better contract with each agent. The intuition in our real-world setting is one where with the benchmarking data, hospital administrators can make their negotiators' performance reviews contingent on the prices they negotiate relative to other hospitals for the same brand. This motivates the following Predictions:

Prediction 4 (Monitoring Effect on Prices) If buyer negotiators are imperfect agents of the buying firm, then benchmarking information (observing the distribution of price realizations across hospitals $\{p_h\}_{h=1}^H$) allows the principal to estimate whether the seller is the weak or strong type, and thus reduce the risk to which the agent is exposed and write a contract which induces more bargaining effort and a lower price than in the case where only p_h is observed.[48]

Prediction 5 (Monitoring Effect on Prices with High Quantity) If buyer negotiators are imperfect agents of the buying firm, but it is costly for hospital managers to search and analyze the data in a way that allows them to write better contracts, then managers will use benchmarking information (observing the distribution of price realizations across hospitals $\{p_h\}_{h=1}^H$) to write a contract which induces more bargaining effort by the agent and a lower price than in the case where only p_h is observed if $(p_h(m) - p_h(m(\{p_h\}_{h=1}^H)))q_h$ exceeds the cost of information use.

C   Details on the Contracting Environment

In this Appendix, we provide some additional background on how medical devices are used and purchased, and we provide additional evidence regarding the validity of our assumptions and potential mechanisms.

---

[48]The model as written has a strong prediction that this effect will be independent of price. However, in general the prediction of how the price distribution would move with information depends on where in the model the current heterogeneity is coming from. For example, if the heterogeneity were due to different levels of risk aversion among negotiators, then benchmarking information would tend to decrease the highest prices more than the lowest.

C.1 Hospital-Supplier Contracting

Hospitals purchase thousands of product categories, and prices for each of these product categories are determined in negotiation. Purchasers within a hospital must factor in clinical value, safety, cost, and other conditions of sale and service (e.g., for capital equipment, maintenance is an important consideration). Negotiation can take place directly between a hospital administrator and a representative of the brand's manufacturer, or hospitals may rely on group purchasing organizations or other contracting coalitions to negotiate their contracts. GPO prices are typically used as a starting point for direct hospital-manufacturer negotiations for some products, particularly physician preference items and capital equipment (Schneller 2009).

Our conversations with industry participants revealed substantial heterogeneity across organizations in the purchasing practices via which supply contracts are negotiated. Some hospitals have large materials management or purchasing departments with agents who specialize in negotiations. Sometimes a large business unit, such as a catheter lab in the case of stents, will coordinate its own purchasing separately from the rest of the hospital. Even absent access to benchmarking information, hospitals vary in access to information on the prices other hospitals pay via GPOs, hospital system membership, or informal networks of peers. While GPO membership is unknown, many of the facilities in our data are members of health systems at some point. The extent to which hospital systems bargain collectively over supplies will vary by context. Hospitals often argue that the ability to bargain collectively is a substantial advantage over and above that conferred by GPO membership (Noether and May 2017). However, we find in other work that hospital mergers lead to limited savings relative to the dispersion we document for merging parties (Craig et al. 2018); we also observe substantial price dispersion across hospitals in the same system in these data. For example, the Gini coefficient for the average physician preference item is 0.058 across all sample facilities that are members of health systems, whereas the average within-health system Gini coefficient is actually slightly higher at 0.061.[49] The same comparison is 0.103 vs. 0.104 for commodities, and 0.077 vs. 0.076 for other medical/surgical products. Thus, we observe little evidence of collective bargaining by hospitals under the same ownership in these data.

There is also significant variation in discounts obtained across product categories within a given facility. To investigate this, we have run some regressions analogous to those in Figure 1(a). We stack all product categories and run a single regression of $ln(price)$ on brand-time

---

[49]Facilities in the same system are identified as those under the same PriceGuide parent account. To aid in comparison, we estimate Gini coefficients across all facilities, and average Gini coefficients across facilities in the same system, using the same sample of facilities under multi-member parents. The Gini coefficient is one half of the mean absolute difference between any randomly selected pair of hospitals.

fixed effects, and several different versions of hospital fixed effects: hospital fixed effects, hospital-product class fixed effects, and hospital-product category fixed effects. First, we document that hospitals' performance in bargaining is positively, but imperfectly, correlated across product classes: the correlation between a given hospital's average discount on PPIs and that same hospital's average discount on commodities is only 0.228; the same statistic for PPIs and other medical/surgical products is 0.229. There is a stronger correlation (0.436) between hospital fixed effects for commodities and those for other medical/surgical products. Second, and relatedly, we find that a significant component of hospitals' performance in bargaining is idiosyncratic across categories. Conditional on fixed effects for each product category-brand-time, hospital fixed effects explain only 4.3 percent of the residual price variation in the stacked regression, hospital-product class fixed effects explain 7.2 percent of the residual price variation, and hospital-product category fixed effects explain 30.4 percent of the residual price variation.[50] This finding is consistent with our understanding that purchasing is somewhat siloed across departments within a hospital.

C.1.1 Stent Usage and Purchasing

The main product category we analyze is coronary stents; coronary stents are small metal tubes placed into narrowed coronary arteries to widen them and allow blood flow to the heart.[51] In the US, hospitals spend more than two billion dollars annually on stents used in over 700,000 procedures.[52]

For stents, as for other physician preference technologies, usage is driven by physicians choosing which brand to use to treat a given patient, while prices are determined in negotiation between a hospital administrator and a representative of the brand's manufacturer. There is no "search" in the conventional sense, as a given brand can only be purchased directly from its manufacturer. The manufacturer holds inventory on-site at the hospital, and the purchase is made when the physician pulls the product off the shelf and implants it into the patient.

---

[50]More precisely, the R^2 of the baseline pooled regression with only category-brand-time fixed effects is 0.9931. Adding a hospital fixed effect increases the R^2, but only to 0.9934 $((0.9934 - 0.9931) * 100/(1 - 0.9931) = 4.3$ percent; this indicates that some hospitals have systematically better prices, holding both time and product portfolio fixed. Adding hospital-class fixed effects increases the R^2 to 0.9936. Finally, adding hospital-product category fixed effects increases the R^2 to 0.9952.

[51]The original technology, the bare metal stent (BMS), was approved in the early 1990s; in the early 2000s, the drug-eluting stent (DES) was introduced as an improvement over the older technology with lower risk of restenosis, a condition that may arise when scar tissue builds up around the stent and restricts blood flow yet again.

[52]700,000 estimate from Waldman et al. (2013), referencing stent procedures in Medicare enrollee population. Two billion dollar figure based on authors' calculations using Boston Scientific's reported US revenue in 2012 (BSX 10-K 2012) and Boston Scientific's 2012 market share in purchase order data.

Stent contracts typically specify a linear price for the contract duration, often a year. In the short run, hospitals are reimbursed a fixed amount by private or public insurers based on the services they provide, and so device prices come directly from the hospital's bottom line.

C.2   Compensation of Purchasing Agents

In this Appendix, we use survey and interview data to explore the agency hypothesis. First, we have obtained access to management practices survey data for a large sample of hospital cardiac units from McConnell et al. (2013). Second, we have conducted interviews with three different types of hospitals regarding their purchasing practices, contracting with purchasers, and the interaction of these with using benchmarking information.

McConnell et al. (2013) developed a survey instrument, adapted from "lean manufacturing" surveys, to measure management practices in the cardiac inpatient setting. The survey included open-ended questions on standardization, performance monitoring, setting targets, and incentivizing employees and managers. Many of these questions shed light on how staff in cardiac units are tracked and rewarded. Those that are most relevant for the current setting are the following (in each case, we summarize the "best case scenario"):

- **Technology Adoption** Score 5: "There is a systematic approach in the adoption of all new technologies, including a review of best evidence on costs and effectiveness; unit may participate in learning collaborative to accelerate information gathering."

- **Performance Review** "Performance is continually reviewed, based on the indicators tracked; all aspects are followed up to ensure continuous improvement; results are communicated to all staff."

- **Target Stretch**: "Goals are genuinely demanding for all parts of the organization and developed in consultation with senior staff (eg, to adjust external benchmarks appropriately)."

- **Rewarding High Performers**: "There is an evaluation system for the awarding of performance related rewards, including personal financial rewards and shared group/team rewards."

- **Removing Poor Performers**: "We work hard to identify weaknesses and improve or remove poor performers."

Each survey answer was scored on a scale of 1 to 5, with 5 corresponding to the "best" answer as described above. Cardiac units were heterogeneous in their responses to these

questions, suggesting that the ex ante potential for our "moral hazard" mechanism would be limited to a small number of hospitals in our data. The raw scores are summarized in Table A6. First, they suggest that only 33 percent of units received a 4 or 5 on the technology adoption question – while not touching on incentives explicitly, this question characterizes the extent to which cost-effective input use is an *objective* of the unit. Second, more than half of units received a 4 or 5 on the performance review question. Together, these findings imply that only a small subset of cardiac units might both prioritize cost-effective purchasing and review performance on that metric. Finally, the percent of respondents scoring a 4 or 5 on the target stretch (i.e., external benchmarks), rewarding high performers, and removing poor performers questions ranged from 20-27 percent. In our view, this paints a relatively pessimistic picture of the potential role of performance pay based on benchmarking to mitigate moral hazard in our setting.

Table A6: Proportion of answers to five practice-related questions in %

	1-worst	2	3	4	5-best
Technology adoption	8	15	44	26	8
Performance review	1	12	33	40	14
Target stretch	8	28	37	23	4
Rewarding high performers	20	19	40	16	4
Removing poor performers	9	30	36	21	3

Raw interview score tabulations shared by K. J. McConnell; summary data reported in McConnell et al. (2013).

We also interviewed several representatives of hospital purchasing departments. Each interviewee worked within the central purchasing unit of a hospital or hospital system and was asked about use of external benchmarking data in employee compensation. One interviewee worked within a large regional hospital system, one within a large urban academic medical center, and one within a small community hospital. Our interview with the representative of the large regional hospital chain indicated that the central purchasing manager regularly makes individual business units aware of opportunities for savings based on external benchmarking, and that annual performance evaluation involves discussion of the prior year's accomplishments, including savings on medical supplies. Staff are also eligible to receive small bonuses (around $300) for particularly strong performance, akin to employee-of-the-month recognition. Our interview with the representative of the academic medical center described performance pay for purchasing staff in a similarly holistic way: each contract manager is evaluated periodically based on general performance on his/her overall portfolio of contracts, with dimensions including timely execution of contracts, management of con-

tracts, and "savings they drive." In each of these interviews with representatives of large purchasers, we learned that benchmarking data provide information on savings achieved by contract managers and that annual performance is evaluated in part based on savings, but there is no formal or specific link between compensation and performance relative to metrics from benchmarking resources. The interview with the small community hospital indicated that pricing information was primarily obtained from the hospital's GPO and a small regional purchasing collaborative; that hospital was in the process of transitioning to a subscription benchmarking service, but did not mention employee performance evaluation as a goal of that subscription.

In sum, our understanding is that benchmarking can be part of a purchasing agent's overall performance evaluation, consistent with the theoretical benefit of benchmarking to resolve moral hazard problems. However, both large-sample surveys of cardiac units and our in-depth interviews of purchasing units indicate that reliance on benchmarking to evaluate purchasing employees is heterogeneous and informal, consistent with our finding that asymmetric information can account for the overall effect of transparency on price realizations.

C.3 Evidence Regarding Linear Price Assumption

It is important for our analysis that the prices we observe are comparable across observations in the sense that there are not important contract dimensions that we do not observe (e.g., bundling, exclusivity, or market share based contracting). Our conversations with industry participants indicate that stents tend to have simple linear price contracts, so we assume that transactions data capture real prices. Here, we examine these assumptions in the data.

C.3.1 Extent of "exclusivity" and correlations with prices

In panels (a) and (c) of Figure A7, we show histograms of total unique manufacturers and stent products (brands) purchased over each quarter by each hospital in the sample. The vast majority of hospitals purchase multiple brands from multiple manufacturers, rather than purchasing a single most-preferred brand for the whole facility. Panels (b) and (d) show these histograms for only hospitals above the 25th percentile in total stent volume, and show even less evidence of "exclusivity" – fewer than 3 percent of hospital-quarters involve a single brand and fewer than 7 percent involve a single manufacturer. The fact that the majority of the already small amount of observed "exclusivity" occurs at hospitals with lower utilization is consistent with the anecdotal evidence that exclusivity does not play a systematic role in stent contracting – with true linear price contracts, "exclusive" purchasing patterns are more likely to be observed among hospitals with low utilization due to random variation and

Figure A7: Histograms – Number of Unique Brands/Manufacturers per Hospital-Quarter

(a) Brands

(b) Brands (excluding smallest hospitals)

(c) Manufacturers

(d) Manufacturers (excluding smallest hospitals)

Regressions of price on indicators for exclusivity:		
Specification	β^{Excl}	(s.e.)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \epsilon_{hjt}$	$25.3^{\dagger}$	(9.0)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \theta_h + \epsilon_{hjt}$	13.2	(9.5)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \theta_{hj} + \epsilon_{hjt}$	-2.9	(8.0)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \epsilon_{hjt}$	8.6	(9.4)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \theta_h + \epsilon_{hjt}$	-15.9^{**}	(6.5)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \theta_{hj} + \epsilon_{hjt}$	-4.2	(6.8)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,223$. Standard errors clustered at hospital level, $N_h = 507$ in first two specifications and hospital-brand level $N_{hj} = 2,227$ in brand-hospital fixed effects specification. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

costly contracting.

As a further check, we look at the pricing consequences of the observed sole sourcing in the usage data. For the minority of hospitals that do happen to use only a single brand or manufacturer in a given quarter, we create indicators for $\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}}$ and $\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}}$ and regress price on each indicator and brand-month fixed effects θ_{jt}. The only negative and statistically significant result is a \$15.90 lower price associated with manufacturer sole sourcing when hospital fixed effects are included. This is consistent with a small price savings from "standardization," holding the hospital fixed. However, the result goes away when including hospital-brand fixed effects, so the \$16 result appears to be either spurious or driven by composition effects. We conclude that the small amount of sole sourcing observed is most likely due to other factors besides contracting concerns.

Finally, we check for any evidence of near-exclusivity in the form of market share based contracts (which we are told are commonly used for many medical products, but not stents). Figure A8 plots the cumulative density of observations by brand market share at the hospital-quarter level. We do not observe the bunching that we would expect if contracts commonly specified market share thresholds in either the full sample (panel (a)) or restricting to the most used brand at each hospital (panel (b)). In the table below the Figure, we also rerun the "exclusivity" price regressions using cases where market shares exceed 90 percent, i.e. $\mathbf{1}_{\{s_{jht}>0.90\}}$, and find no economically or statistically meaningful relationship with prices.

C.3.2 Standard offers and multipliers/bundling

Many real-world bargaining settings have "list prices" as a starting point for negotiations. Sometimes these prices are in practice paid by almost no one, and so less relevant to the bargaining problem. Sometimes they are paid by a large mass of smaller customers. They may also act as a starting point for a negotiation of a single "multiplier" that would modify the list prices across a variety of products sold by the supplier to the buyer. Our understanding from talking to both buyers and suppliers in the coronary stent market is that these practices are not prevalent. Here we examine the data to verify this understanding and better understand the nature of price setting.

Figure A9 plots the price distributions across hospitals for each of four DES sold to at least 100 hospitals in January 2012. While the contracted prices do aggregate at the "round" \$50 increments, they do so in a way that appears to approximate single-peaked distributions that are smooth in the sense of no mass suggesting an obvious focal price standing out from the rest. In particular, they do not show evidence of a list price at the top paid by a large subset of customers. These figures are representative of the price distributions for DES across hospitals in other time periods.

60

Figure A8: Cumulative distributions by market share

(a) All

(b) Most Used Brands Only

Regressions of price on indicators for exclusivity (share-based contracts):		
Specification	β^{Excl}	(s.e.)
$p_{hjt} = \beta^{Excl} \mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \epsilon_{hjt}$	12.2	(11.5)
$p_{hjt} = \beta^{Excl} \mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \theta_h + \epsilon_{hjt}$	-9.4	(10.7)
$p_{hjt} = \beta^{Excl} \mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \theta_{hj} + \epsilon_{hjt}$	-2.9	(8.0)
$p_{hjt} = \beta^{Excl} \mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \epsilon_{hjt}$	18.3	(13.5)
$p_{hjt} = \beta^{Excl} \mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \theta_h + \epsilon_{hjt}$	-7.5	(9.9)
$p_{hjt} = \beta^{Excl} \mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \theta_{hj} + \epsilon_{hjt}$	-4.2	(6.8)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,223$. Standard errors clustered at hospital level, $N_h = 507$ in first two specifications and hospital-brand level $N_{hj} = 2,227$ in brand-hospital fixed effects specification. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A9: Histograms of prices across hospitals for each brand in January 2012

Table A7: Price Co-movement at the Manufacturer-Hospital Level

$1_{\{\Delta p \neq 0\}}$ (DES)	(1)	(2)	(3)
$1_{\{\Delta p \neq 0\}}$ (BMS)	0.0621		
	(0.0556)		
$1_{\{\Delta p \neq 0\}}$ (Guiding Cath)		0.0180	
		(0.0383)	
$1_{\{\Delta p \neq 0\}}$ (Guide Wires)			0.0646
			(0.0828)
R^2	0.005	0.000	0.003
Δp (DES)	(1)	(2)	(3)
Δp (BMS)	0.133		
	(0.119)		
Δp (Guiding Cath)		0.0169	
		(0.0254)	
Δp (Guide Wires)			4.133
			(4.365)
R^2	0.011	0.000	0.008

$N = 973$ manufacturer-hospital-month observations. Standard errors in parentheses, clustered by hospital. $N_h = 85$. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Table A7 explores the extent to which prices move together for products sold by the same manufacturer to the same hospital. We flag price changes (using the same algorithm for identifying price changes for our renegotiation analysis) for DES, BMS, Guiding Catheters, and Guide Wires – all of these product categories are used together in interventional cardiology procedures and sold by the same manufacturers, and potentially even sold via the same sales force.[53] We then merge the four data sets at the manufacturer-hospital-month level, keeping the most purchased brand for each manufacturer-hospital pair to ensure at most one observation in a category. Before any further analysis, we note that only 85 hospitals in the sample purchase their modal brands from the same manufacturer across these categories – this fact alone reveals the limited extent to which pricing that references other categories could be feasible, given observed usage patterns. The top panel of the table then regresses the flag for DES price change on the flag for a price change in each of the three other categories (e.g. $1_{\{\Delta p_{mht}^{des} \neq 0\}} = \beta_0 + \beta_1 1_{\{\Delta p_{mht}^{bms} \neq 0\}} + \epsilon$). The interpretation of the coefficient β_1 is then the percentage of hospital-manufacturer-months in which price changes for the other cardiology device category are accompanied by price changes in DES, and R^2 is then the

---

[53]Guiding Catheters and Guide Wires are used in catheter-based interventional procedures outside of the coronary and even vascular settings, and so used by physicians outside of interventional cardiology. Depending on the hospital and manufacturer, these could be sold via separate channels.

percentage of price co-movement overall – in the case of perfect co-movement, $\beta_1 = 1$ and $R^2 = 1$. The results indicate that DES prices change 6 percent of the times BMS prices change, 2 percent of the times catheter prices change, and 6 percent of the times guide wire prices change (though none of these point estimates are statistically different from zero, all are statistically different from 1). R^2 are also very small, indicating DES prices often change in months when other category prices do not.

The bottom panel of Table A7 shows the output of similar regressions but using the actual magnitudes of price changes (e.g. $\Delta p_{mht}^{des} = \beta_0 + \beta_1 \Delta p_{mht}^{bms} + \epsilon$) to understand the direction and magnitude of co-movements. Here there are again no statistically significant relationships. Our interpretation of these results is that both the degree and magnitude of price co-movements are small. This seems inconsistent with a model of a multiplier on list prices negotiated across many product categories, where we would expect large degrees of positive correlation in the incidence and relative magnitude of price changes. It is also inconsistent with most models of bundled pricing, which would tend to predict positive coincidence of price changes and negative correlation in price change magnitudes.

In unreported analyses, available upon request, we conduct similar analyses for the three other groups of product categories meeting the criterion that the same vendors account for at least 70 percent of hospital spend: pacemakers and defibrillators; knee and hip prosthesis components; and several types of sutures. The patterns are similar to those shown here for stents, catheters, and guide wires. Within manufacturer-hospital pair, across categories: (1) the overall frequency of price co-movements is small; and (2) the correlations in price change magnitudes are also small, and positive, where they exist.

D Stent Analyses – Detail and Robustness

D.1 Stent Analyses in Text – Detail and Alternative Controls

The differences in differences estimate of the average treatment effect of access to information is small, statistically and economically, at -\$3 (s.e. 3). Here, we also use an event study specification that includes indicators for each month relative to the hospital's "info" date:

$$P_{hjt} \;=\; \sum_{mo=-12}^{+12} \beta^{Info,mo} * \mathbb{1}_{\{mo=t-t_{info_{hj}}\}} + \theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j}) + \varepsilon_{hjt}$$

Figure A10 shows results for these estimated differences between treated and untreated prices. The plot shows evidence of a slight decline in prices prior to accessing information,

though the pre-trends in price in the six months leading up to the timing of information are essentially zero.[54] After the hospital accesses the benchmarking information, there is a steady downward trend in the coefficients. The downward trend in the post-period may be due to price stickiness – it may take newly-informed hospitals some time to arrive at the bargaining table.

In general, estimates for each relative month effect are insignificant and there is not strong evidence of a trend break. Moreover, estimated patterns are similar across the different specifications of controls, though standard errors are large in the richest specification (Version 3).[55]

Figures A11 and A12 show the same information as in Figure 3 in the main text, but allowing for alternative specifications of fixed effects. As with the overall average treatment effects, the estimates are significantly smaller when we control for hospital-by-brand (rather than hospital *and* brand) fixed effects. However, we see the same qualitative pattern regardless of controls: the treatment effects are statistically zero in all but the top quintile of the pre-information price distribution, the previously high-priced hospital-brands achieve significant price decreases in the presence of benchmarking information, and these price decreases manifest somewhat slowly after the initial information shock.

The following three Figures show how the specifications in Table 2, each of which explores the mechanisms underlying the core results for drug-eluting stents, vary with the included controls. Figure A13 shows how the distribution of treatment effects as a function of pre-information price vary by whether the hospital-brand involved high or low purchase quantities pre-information. Figure A14 explores whether our treatment effects can be attributed to the agency vs. asymmetric information hypotheses. Finally, Figure A15 shows the effects of information on renegotiation per se, as well as the effects of information on price decreases conditional on renegotiation taking place.

---

[54]It should be noted that there are fewer "pre-info" observations available 6-12 months prior to accessing information because of the presence of entering brands and because some hospitals do not submit retrospective data until a few months after joining the database. Accordingly, the earlier relative month effects are less precisely estimated.

[55]It was not possible to estimate the monthly event studies with hospital-brand and brand-month fixed effects. However, the quarterly event study with hospital-brand and brand-month fixed effects is essentially identical to the quarterly event study with hospital-brand and brand-specific linear trends.

Ver-sion		Month relative to info date ($\beta^{Info,mo}$)																								
		-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12
1		17	15	20*	20*	17*	3	-1	-5	3	0	-1	3	-8†	-10†	-14†	-16†	-20†	-21†	-22†	-24†	-22†	-20†	-22†	21†
		(14)	(14)	(11)	(11)	(9)	(7)	(6)	(5)	(5)	(4)	(3)	(3)	(3)	(3)	(4)	(4)	(5)	(5)	(6)	(6)	(7)	(7)	(7)	(8)
2		22	19	19*	22**	16*	7	5	1	9	8	4	6	-13†	-17†	-18†	-21†	-27†	-27†	-25†	-24†	-24†	-21**	-23†	-18*
		(14)	(14)	(11)	(10)	(9)	(8)	(8)	(7)	(6)	(6)	(5)	(4)	(4)	(4)	(5)	(6)	(7)	(7)	(7)	(8)	(9)	(8)	(9)	(9)
3		16	14	22	19	14	7	7	3	5	2	1	0	-4	-6	-10	-11	-15	-17	-19	-23	-19	-22	-23	-22
		(35)	(32)	(28)	(26)	(23)	(20)	(17)	(14)	(11)	(9)	(6)	(3)	(3)	(6)	(9)	(12)	(15)	(18)	(19)	(22)	(25)	(28)	(30)	(33)

Authors' calculations from PriceGuide data, 2009-2014. $N = 23,016$ member-brand-months. Includes 507 members, twelve months pre- and post-join only. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Version 3) level shown in parentheses. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A10: Event Studies of Treatment Effect of Access to Benchmarking Information

Version	Pre-info price quintiles ($\beta_{quintile}^{Info}=$)				
	1	2	3	4	5
1	4	$17^{\dagger}$	9	-2	$-55^{\dagger}$
	(6)	(6)	(6)	(6)	(9)
2	-1	11	4	-7	$-63^{\dagger}$
	(7)	(7)	(7)	(8)	(10)
3	-7	6	6	1	$-27^{\dagger}$
	(5)	(5)	(5)	(5)	(7)
4	-10	1	2	-3	$-34^{\dagger}$
	(6)	(6)	(6)	(6)	(8)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32{,}453$ member-brand-months. Includes 508 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A11: Treatment Effect Estimates Throughout the Price Distribution

Ver-sion	Month relative to join date ($\beta_5^{Join,mo}$)																							
	−12	−11	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	1	2	3	4	5	6	7	8	9	10	11	12
1	−24	−33	−20	−24	−14	−14	−21*	−31†	−27†	−22†	−14**	−8	−23†	−40†	−52†	−61†	−64†	−74†	−77†	−88†	−83†	−78†	−104†	−104†
	(20)	(25)	(22)	(19)	(16)	(17)	(12)	(10)	(8)	(7)	(7)	(6)	(7)	(8)	(9)	(9)	(11)	(12)	(10)	(13)	(13)	(11)	(13)	(13)
2	−17	−15	−10	−11	−6	−8	−14	−29†	−22**	−15*	−9	−6	−26†	−43†	−54†	−65†	−70†	−80†	−80†	−89†	−83†	−75†	−104†	−102†
	(19)	(24)	(18)	(18)	(15)	(17)	(13)	(10)	(9)	(8)	(7)	(7)	(7)	(9)	(9)	(10)	(11)	(13)	(11)	(13)	(14)	(12)	(13)	(13)
3	−18	−36	−27	−40	−34	−26	−8	−4	−6	−4	−5	−4	−16†	−28†	−43†	−50†	−56†	−62†	−65†	−82†	−76†	−75**	−93†	−96†
	(36)	(34)	(30)	(27)	(24)	(22)	(18)	(16)	(12)	(10)	(6)	(4)	(5)	(9)	(11)	(14)	(17)	(20)	(21)	(24)	(27)	(29)	(32)	(35)

Authors' calculations from PriceGuide data, 2009-2014. $N = 23{,}016$ member-brand-months. Includes 507 members, twelve months pre- and post-join only. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Version 3) level shown in parentheses. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A12: Event Studies of Treatment Effect of Access to Benchmarking Information, Top Quintile of Price Only

(a) Low Quantity

(b) High Quantity

Version	Pre-info price quintiles ($\beta^{Info}_{quintile,low^q}$)					Pre-info price quintiles ($\beta^{Info}_{quintile,low^q} + \beta^{Info}_{quintile,high^q}$)				
	1	2	3	4	5	1	2	3	4	5
1	5	$19^†$	16^{**}	1	$-51^†$	3	13	-4	-10	$-73^†$
	(7)	(6)	(7)	(7)	(9)	(10)	(9)	(9)	(10)	(15)
2	-1	12	11	-4	$-60^†$	0	7	-12	-14	$-79^†$
	(8)	(8)	(8)	(8)	(11)	(11)	(10)	(10)	(11)	(16)
3	-4	9	9	5	-17^{**}	-11	0	0	-9	$-71^†$
	(6)	(6)	(6)	(6)	(7)	(9)	(8)	(7)	(8)	(13)
4	-9	4	6	1	$-23^†$	-12	-4	-5	-12	$-78^†$
	(7)	(7)	(7)	(7)	(8)	(9)	(8)	(8)	(9)	(13)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,453$ member-brand-months. Includes 508 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A13: Treatment Effect Estimates Across the Price and Quantity Distributions

(a) Agency

(b) Asymmetric Info

Version	Pre-info price quintiles ($\beta_{quintile}^{Agency}$)					Pre-info price quintiles ($\beta_{quintile}^{Info}$)				
	1	2	3	4	5	1	2	3	4	5
1	8	14	−9	−21*	−60†	3	12**	15†	8	−33†
	(12)	(11)	(10)	(12)	(21)	(6)	(5)	(5)	(7)	(7)
2	19	31**	1	−11	−53**	−8	−2	4	−2	−46†
	(13)	(12)	(11)	(14)	(22)	(9)	(8)	(8)	(10)	(10)
3	−17	−3	2	7	13	−1	7	5	−1	−30†
	(11)	(12)	(10)	(12)	(18)	(6)	(5)	(5)	(5)	(7)
4	−3	16	16	18	25	−15*	−11	−11	−15**	−47†
	(12)	(12)	(11)	(12)	(19)	(8)	(7)	(7)	(7)	(9)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,453$ member-brand-months. Includes 508 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A14: Treatment Effect Estimates Across the Price Distribution, Separating Agency and Asymmetric Information Mechanisms

(a) Conditional on Renegotiation

(b) $\mathbb{1}_{\{reneg_{hjt}\}}$

Version	Pre-join price quintiles ($\beta_{quintile}^{Info}$)					Pre-join price quintiles ($\beta_{quintile}^{Info}$)				
	1	2	3	4	5	1	2	3	4	5
1	-7	13	11	-22	$-91^{\dagger}$	.008	.002	.018**	.018*	.034†
	(12)	(12)	(14)	(14)	(19)	(.008)	(.009)	(.008)	(.01)	(.008)
2	-11	5	1	-25	$-111^{\dagger}$	.017*	.01	.023**	.025**	.04†
	(16)	(15)	(17)	(17)	(23)	(.009)	(.01)	(.009)	(.011)	(.009)
3	-14	4	1	-13	$-76^{\dagger}$	.01	.013	.016*	.018	.023**
	(15)	(14)	(19)	(17)	(18)	(.01)	(.01)	(.009)	(.011)	(.009)
4	-8	4	-3	-11	$-80^{\dagger}$	.022**	.022*	.024**	.029**	.032†
	(19)	(19)	(22)	(20)	(22)	(.011)	(.012)	(.01)	(.012)	(.011)

Authors' calculations from PriceGuide data, 2009-2014. $N = 6,510$ member-brand-months in regressions conditional on renegotiation. $N = 32,453$ member-brand-months in renegotiation dummies regression. Includes 508 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A15: Treatment Effects Conditional on Renegotiation and on Occurrence of Renegotiation

D.2   Robustness

The following Figures show the results of our richest regression specification, allowing for different treatment effects for different parts of the price and quantity distributions, for specifications that (1) focus only on the twelve months before and after information (Figure A16); (2) limit the sample to those facilities registered with the database as "hospitals" as opposed to "health systems" or other facility types (Figure A17); and (3) examine heterogeneity in results as a function of hospital type (Table A8). Results are qualitatively and quantitatively similar in specifications (1) and (2).

(a) Low Quantity **(b)** High Quantity

Version	Pre-info price quintiles ($\beta_{quintile,low^q}^{Info}$)					Pre-info price quintiles $(\beta_{quintile,low^q}^{Info} + \beta_{quintile,high^q}^{Info})$				
	1	2	3	4	5	1	2	3	4	5
1	-7	15^{**}	14^{**}	8	$-30^{\dagger}$	-11	7	7	1	$-54^{\dagger}$
	(7)	(6)	(6)	(7)	(9)	(9)	(8)	(7)	(8)	(16)
2	-13	9	6	3	$-38^{\dagger}$	-16	0	-2	-3	$-61^{\dagger}$
	(8)	(8)	(7)	(8)	(9)	(10)	(8)	(8)	(9)	(16)
3	-7	10^{*}	9^{*}	14^{**}	-11^{*}	-7	2	5	1	$-60^{\dagger}$
	(6)	(5)	(5)	(6)	(6)	(8)	(7)	(6)	(7)	(15)
4	-11^{*}	7	5	11^{*}	-16^{**}	-10	-1	1	-2	$-66^{\dagger}$
	(7)	(6)	(6)	(6)	(7)	(8)	(8)	(6)	(7)	(15)

Authors' calculations from PriceGuide data, 2009-2014. $N = 23,016$ member-brand-months. Includes 507 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A16: Treatment Effect Estimates Across the Price and Quantity Distributions – Twelve Months Pre/Post

To further explore convergence within hospitals of differing "types", we also estimate

(a) Low Quantity   **(b)** High Quantity

Version	Pre-info price quintiles ($\beta^{Info}_{quintile,low^q}$)					Pre-info price quintiles ($\beta^{Info}_{quintile,low^q} + \beta^{Info}_{quintile,high^q}$)				
	1	2	3	4	5	1	2	3	4	5
1	8	$21^{\dagger}$	15^{**}	-2	$-53^{\dagger}$	-1	8	-10	-7	$-80^{\dagger}$
	(7)	(6)	(7)	(8)	(11)	(10)	(10)	(10)	(12)	(18)
2	2	13	11	-5	$-64^{\dagger}$	-5	2	-18	-13	$-84^{\dagger}$
	(8)	(8)	(8)	(9)	(12)	(11)	(10)	(11)	(12)	(18)
3	0	12^{**}	9	6	-20^{**}	-11	-2	-3	-7	$-72^{\dagger}$
	(6)	(5)	(7)	(7)	(8)	(9)	(8)	(8)	(9)	(16)
4	-3	7	7	2	$-26^{\dagger}$	-12	-5	-7	-10	$-77^{\dagger}$
	(7)	(6)	(8)	(7)	(9)	(9)	(9)	(9)	(9)	(16)

Authors' calculations from PriceGuide data, 2009-2014. $N = 27,698$ member-brand-months. Includes 436 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A17: Treatment Effect Estimates Across the Price and Quantity Distributions – Hospitals Only

73

our richest specification within hospital size bins, within regions, and within size-region.[56] We find several patterns of interest, now reported in Table A8.[57] In the second panel of Table A8, we see suggestive evidence that treatment effects vary across regions. Focusing on the high-price and high-quantity hospital-brands, the treatment effects are smallest in the northeast and south (-\$42 and -\$43, respectively), largest in the west (-\$104), and similar to our baseline, pooled-region estimates (-\$71) in the midwest (-\$70); none of these estimates is statistically significantly different from the pooled-region estimate at the 5% level. For high-price and relatively low-quantity hospital-brands, only the treatment effect for the western region is statistically significant. In the third panel of Table A8, we show that the distribution of treatment effect estimates in price and quantity space differs by hospital size. The treatment effect estimates for small- and medium-sized hospitals (1-199 and 200-399 beds, respectively) are statistically significant and of similar magnitude to our baseline estimate (-\$17) for high-price, relatively low-quantity hospital-brands: the treatment effects estimates are -\$27 and -\$24, respectively. The treatment effect estimates for medium-sized and large hospitals (400+ beds) are large and statistically significant for high-price, relatively high-quantity hospital-brands: the treatment effects estimates are -\$90 and -\$68, respectively. This pattern is driven by the fact that large hospitals tend to purchase brands in large quantities and small hospitals tend to purchase brands in low quantities. In unreported results, we see no clear evidence that the overall distribution of treatment effects varies with hospital size; the treatment effect for all hospital-brands that were high-price pre-join is largest (-\$31) among medium-sized hospitals, compared to (-\$23) for both small and large hospitals.

These within-type results are noisier than our baseline pooled results, as expected, but largely confirm our finding of large treatment effects of information among hospital-brands where there is room for improvement *and* a great deal of spending at stake, at least within the limited set of hospital characteristics we have. Furthermore, the final panel of Table A8 reports estimates that mirror our full-sample regressions in the first panel, but using *k*-nearest neighbors matching. The results are nearly identical to our baseline estimates. Thus, price convergence pertains across all hospital-brands and within sets of similar hospitals.

---

[56]The results split by size crossed with region are noisy, but largely confirm the marginal interactions with size and region separately.

[57]Note that, for the sake of comparison with the baseline results, the flags for high- vs. low-quantity hospital-brands and high- vs. low-price hospital-brands are relative to the full sample, not the sample for the relevant hospital type.

Table A8: Exploring the Results by Hospital Type

(1) Baseline Estimates – Treatment Effect Variation with Quantity Purchased:

	Low quantity $\beta^{Info}_{quintile,low^q}$					High quantity $\beta^{Info}_{quintile,high^q}$				
$p_{quintile}$	1	2	3	4	5	1	2	3	4	5
	-4	9	9	5	-17^{**}	-11	0	0	-9	$-71^{\dagger}$
	(6)	(6)	(6)	(6)	(7)	(9)	(8)	(7)	(8)	(13)

(2) Treatment Effect Variation by Region:

	Low quantity $\beta^{Info}_{quintile,low^q}$					High quantity $\beta^{Info}_{quintile,high^q}$				
$p_{quintile}$	1	2	3	4	5	1	2	3	4	5
Midwest	-7	18^{*}	-4	6	-12	-24^{*}	21^{*}	-1	-25	-70^{**}
	(9)	(10)	(12)	(13)	(13)	(14)	(12)	(10)	(21)	(30)
Northeast	-14	10	$42^{\dagger}$	24^{**}	-9	-11	9	20	-14	-42
	(24)	(23)	(14)	(11)	(11)	(27)	(23)	(14)	(14)	(31)
South	7	9	-1	6	11	16^{*}	-5	-6	9	$-43^{\dagger}$
	(13)	(7)	(8)	(10)	(16)	(9)	(13)	(12)	(14)	(11)
West	-4	-2	9	0	$-42^{\dagger}$	-61^{**}	-4	5	1	$-104^{\dagger}$
	(13)	(10)	(13)	(14)	(14)	(29)	(7)	(10)	(19)	(26)

(3) Treatment Effect Variation by Bed Count:

	Low quantity $\beta^{Info}_{quintile,low^q}$					High quantity $\beta^{Info}_{quintile,high^q}$				
$p_{quintile}$	1	2	3	4	5	1	2	3	4	5
0-199	-16	5	7	-10	-27^{*}	-19	13^{*}	-23	-16	29
	(11)	(7)	(12)	(12)	(14)	(22)	(7)	(18)	(23)	(18)
200-399	-4	-2	0	-8	-24^{*}	3	-5	-19	19	$-90^{\dagger}$
	(7)	(11)	(9)	(9)	(13)	(10)	(12)	(18)	(14)	(26)
400+	2	$26^{\dagger}$	19^{*}	$24^{\dagger}$	-3	-14	4	12	-13	$-68^{\dagger}$
	(16)	(9)	(10)	(9)	(9)	(13)	(11)	(8)	(10)	(15)

(4) Matching by p, q, Region, Bed Count:

	Low quantity $\beta^{Info}_{quintile,low^q}$					High quantity $\beta^{Info}_{quintile,high^q}$				
$p_{quintile}$	1	2	3	4	5	1	2	3	4	5
	-10	8	23^{**}	17^{**}	-19	-16	-25	8	-2	$-74^{\dagger}$
	(11)	(7)	(9)	(7)	(12)	(17)	(22)	(10)	(9)	(9)

Authors' calculations from PriceGuide data, 2009-2014. (1): Full sample $N = 32,453$ member-brand-months; 508 members. (2): $N_{Midwest} = 7,540$; $N_{Northeast} = 5,944$; $N_{South} = 10,451$; $N_{West} = 8,518$. (3): $N_{0-199} = 6,991$; $N_{200-399} = 12,884$; $N_{400+} = 12,578$. (4) based on k-Nearest Neighbors matching algorithm. Each treated hospital-brand hj matched to five similar hospitals $k \neq h$ based on region and bed count, as well as price and quantity *for brand j* in the three months prior to hj treatment. $N_{match} = 25,961$. Standard errors clustered at hospital-brand level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.
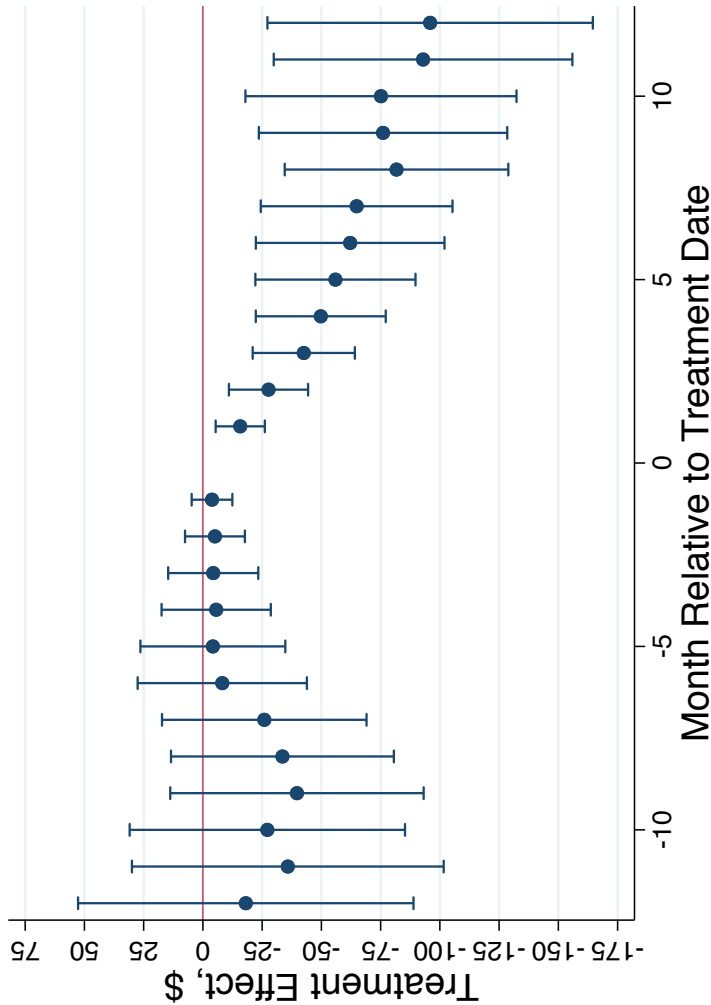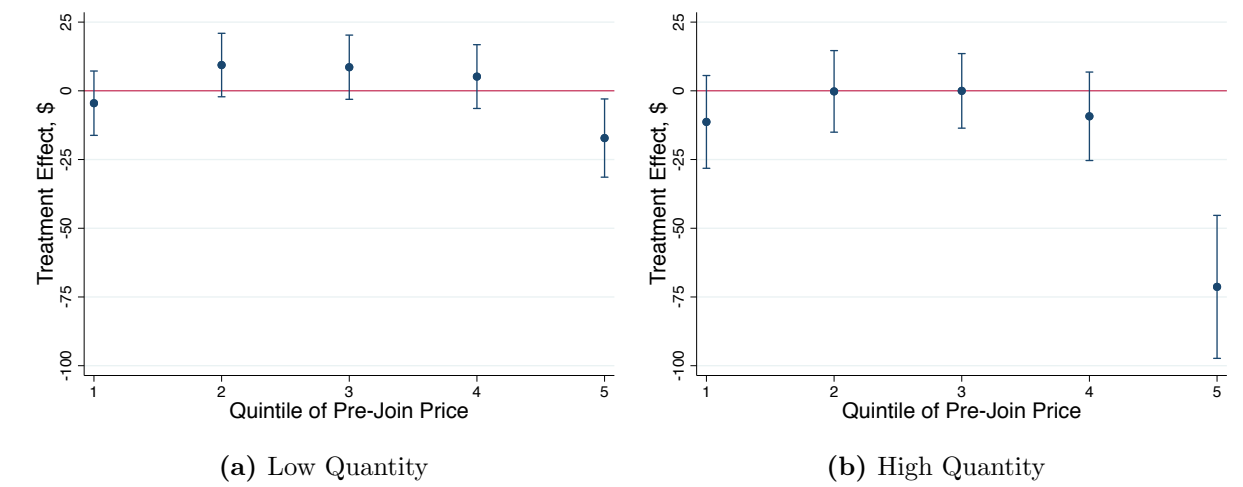
D.3 Cross-Supplier Externalities in Bargaining

As noted briefly in the main text, a full analysis of the information we examine should take into account the potential externality that information relevant for one supplier imposes on negotiations with other suppliers. For example, consider the delegated agent model in Collard-Wexler et al. (2017) (which would apply to our context since transfers are not lump sum). In such a model, information about supplier A's bargaining parameter would enter all other suppliers' negotiations via its expected impact on supplier A's price, which in turn enters the elasticity and marginal contribution terms in all other suppliers' negotiations. In

our framework, this consideration would yield an additional empirical prediction:

Prediction 7 (Indirect Information/Competition Effect on All Prices) With imperfect substitute products, under reasonable assumptions on how the negotiation for one brand affects the disagreement payoff of other brand negotiations, a fall in price of substitute brand j will decrease the surplus up for negotiation for other brands $-j$, leading to a decrease in the prices of other brands $-j$, all else equal.[58] Thus exposure to benchmarking information that leads to a fall in a high price for j should also lead to a fall in any price for other brands $-j$, and the size of this fall will be increasing to the extent the brands are good substitutes for j.

Table A9: Externalities: Effect of New Brand Information on Other Brands

Version of Controls	[1]	[2]	[3]	[4]
β^{Join}	-32†	-32†	-9	-9
	(10)	(10)	(8)	(8)
$\beta^{OtherInfo}$	19†	23†	-4	2
	(8)	(8)	(8)	(9)
$\beta^{Externality}$	-15	-17	-23	-25
	(19)	(19)	(18)	(19)
Hospital+Brand FEs	Y	Y	N	N
Hospital×Brand FEs	N	N	Y	Y
Linear Brand Trends	Y	N	Y	N
Brand×Month FEs	N	Y	N	Y

Authors' calculations from PriceGuide data, 2009-2014. $N = 11,975$ member-brand-months. Includes 508 members. Standard errors clustered at hospital level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.
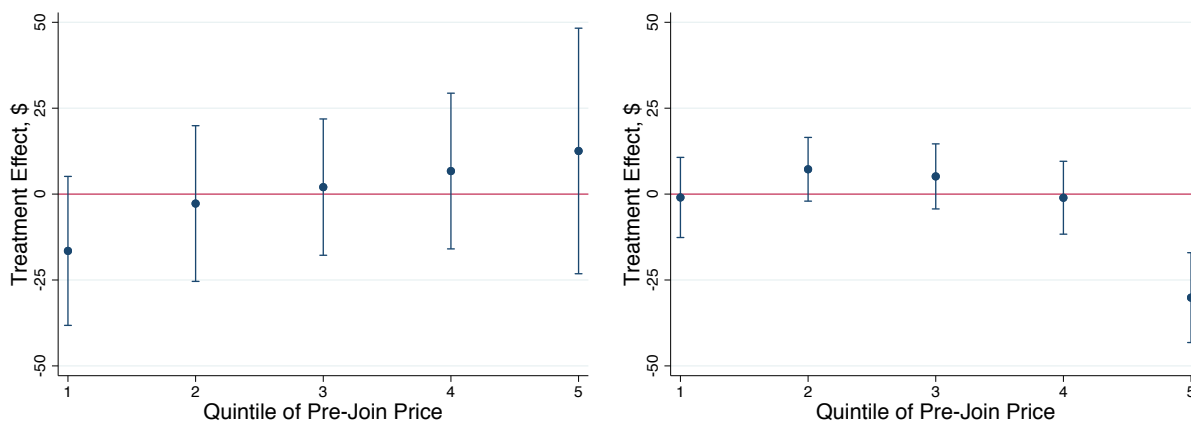
While our empirical approach in the main text does incorporate these types of externalities, it is difficult to disentangle them. To shed additional light, we run quadruple-difference regressions of the form:

$$
\begin{aligned}
P_{hjt} = \ & \beta^{Join} * \mathbb{1}_{\{post_{ht}^{Join_h}\}} + \beta^{OtherInfo} * \mathbb{1}_{\{post_{hjt}^{Info_{hj'\neq j}}\}} \\
& + \beta^{Externality} * \mathbb{1}_{\{post_{hjt}^{Info_{hj'\neq j}}\}} * \mathbb{1}_{\{P_{hj',pre_h} \in quintile(5, \{P_{h'j',pre_h}\}_{h'\in\mathcal{H}})\}} * \mathbb{1}_{\{high_{hj',pre}^q\}} \\
& + \theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j}) + \varepsilon_{hjt}.
\end{aligned} \tag{6}
$$

When a hospital joins the database, all brands receive information shocks, so any externalities between brands are captured in the treatment effect (β^{Join}). In cases of new

---

[58]This will be the case in any model where disagreement payoffs are a function of the prices agreed to with other manufacturers, which has been the case in the empirical bargaining literature thus far and much of the negotiation with externalities theory. It would not be the case in a model such as the Core, where disagreements are based on the primitives of willingness-to-pay and costs.

brand entry, there are information shocks to single brands, so in principle this provides an opportunity to identify externalities. The challenge is that, among hospitals receiving these information shocks, new brand entry occurs sometime after join, so the simple effect of the new brand information on existing brands ($\beta^{OtherInfo}$) may capture delayed treatment effects of joining – like picking the post month in an event study that corresponds with the new brand information date. To address this, we focus on the instances where the hospital is a high-price, high-quantity case when information is revealed about the entering brand. These are the instances where the information shock is strongest, and we difference them versus the rest of the sample for that same brand to isolate the externality effect on other brands in the same hospital-month ($\beta^{Externality}$). For our preferred specifications with hospital-brand fixed effects, the point estimates in Table A9 suggest externalities less than half the size of the point estimates for the focal entering brands, but the estimates are quite noisy, and none are statistically different from zero at standard confidence levels.

E   Expanded Analyses – Detail

The estimates of our main difference-in-differences regression of price on join timing for each "important" product category are shown in Table A10. In the results in Section 6 in the main text, we summarize these results across product category within each physician preference class. In the detailed results here, the left panel shows the results for the full sample described in Appendix A.3. The right panel shows the results for a restricted sample of brands with "significant" market presence (defined as those brands purchased by at least 100 hospitals per year on average), where we can be more confident that our price quintile estimates have reasonable statistical power. In each case, the regression coefficient has been normalized by the mean unit price of the product category to facilitate comparisons across products with dramatically different prices.

The columns in Table A10 are presented as follows: within each sample, the first column shows the average treatment effect; the second through sixth columns show the treatment effect by price quintile; and the seventh column shows the treatment effect for hospital-brands in the top quintile of the pre-join price distribution *and* in the top quartile of the pre-join quantity distribution. Each panel of rows shows results for each individual product category, as well as a spend-weighted average for each product class.

While there is quite a bit of variation in results for this large number of product categories and different specifications, several patterns emerge. The average treatment effects are, as with stents, small and not often statistically significant. Also echoing the findings for stents: the treatment effects for the top quintile of the price distribution are generally negative

and, particularly for PPIs, often statistically significant. The treatment effects are often, though not universally, larger in magnitude for high-price, high-quantity hospital-brands. In contrast to what we find for stents, the treatment effects for hospital-brands in the bottom quintiles of the pre-join price distribution are sometimes negative, but are also sometimes positive and significant. Focusing on PPIs, the average treatment effects in the bottom of the distribution are positive and significant (1.4 percent), though smaller than the price decreases documented in the top of the distribution (3.3 percent).

On balance, as we found with drug-eluting stents, these results demonstrate that the largest savings are achieved by hospitals and for PPI brands with the highest prices and quantities prior to the hospital joining the database. The treatment effects are often larger in magnitude for the restricted sample of brands purchased by an average of 100 or more hospitals; those brands may have received additional attention from hospitals searching the benchmarking database.

In the following analyses, we explore the sensitivity of our focal findings (significant savings among high-price, high-quantity hospital-brands) to alternative regression specifications. We then further explore the small positive price effects documented in the bottom of the distribution.

Table A10: Treatment Effects for Important Product Categories, with hj and jy FE

ID	Name	Class	Unrestricted Sample							$\mathcal{H}_{ij} \geq 100$ Sample						
			ATE	\multicolumn{5}{c}{TE by Price Quintile}	$TE_{high^\ast,high^q}$	ATE	\multicolumn{5}{c}{TE by Price Quintile}	$TE_{high^\ast,high^q}$
				1	2	3	4	5			1	2	3	4	5	
10279	Bandages Elastic	1	-0.010	0.007	-0.018	0.002	-0.038**	-0.009	-0.018	-0.014	0.002	0.025†	-0.002	-0.110*	-0.033	-0.046
11315	Dressings	1	0.000	0.021**	-0.001	0.020*	-0.024	-0.021	-0.022*	0.007*	0.021**	0.008	0.003	-0.001	0.001	0.005
11883	Gloves Surgical	1	-0.005	-0.018	0.028	0.008	-0.028	-0.016	-0.009	-0.008**	-0.005	-0.014**	0.000	-0.011	-0.009	-0.006
12368	Drapes Surgical	1	0.007	0.004	0.024**	0.009	0.015**	-0.024	0.001	-0.013	-0.019	-0.009	0.035**	-0.000	-0.082**	0.007
12745	Needles Injection Hypodermic	1	-0.000	0.014**	0.005	-0.000	0.004	-0.020**	-0.028	0.000	0.011**	0.007	0.003	0.008**	-0.026†	-0.000
13655	Containers Specimen	1	0.018	0.082	0.004	0.003	0.005	0.011	-0.045**	0.050	-0.046	0.040	0.100**	0.102**	-0.007	-0.310†
923840	Drill Bits	1	0.004	0.019†	0.008**	-0.000	0.004	-0.011	-0.012	0.011*	0.022**	0.019†	0.004	0.017*	-0.006	-0.019
	Commodities	1	0.002	0.013†	0.013†	0.006*	-0.009	-0.013†	-0.016**	0.007	-0.003	0.008	0.021**	0.013†	-0.019	-0.053†
10688	Catheters Vascular Angiography	2	0.004	0.025**	0.001	0.001	0.007	-0.013	-0.078**	-0.001	0.000	0.005	0.001	0.012	-0.017	-0.092**
10894	Clip Appliers	2	0.008**	0.035**	0.006	-0.001	0.012**	-0.008	-0.043**	0.008*	0.040**	0.013	-0.006	0.015*	-0.013	-0.055**
11502	Forceps Electrosurgical	2	-0.000	-0.002	0.004**	0.003	0.001	-0.001	-0.001	-0.000*	-0.002	0.006**	0.004	-0.000	-0.001**	-0.001*
11910	Grafts Bone	2	-0.001	-0.008	0.004	-0.005	0.001	-0.000	0.000	-0.000	-0.012	0.000†	0.000†	0.000†	0.000	0.000
11925	Guide Wires	2	0.004	0.007	0.004	0.008*	0.001	-0.001	0.005	0.003	-0.002	0.013	0.009	-0.003	-0.002	0.004
12830	Orthopedic Cement	2	-0.004	-0.007	0.001	0.019†	-0.011	-0.017	-0.005	-0.002	0.004	0.004	0.030*	-0.037*	-0.016	-0.023
13050	Plates Bone	2	-0.003	0.004	0.008	-0.001	-0.010	-0.017**	-0.026**	0.002	0.005	0.019**	0.018†	0.006	-0.034*	-0.076**
13909	Sutures Synthetic Nonabsorbable Polypropylene	2	-0.000	-0.019	0.033*	0.009	0.001	-0.017	0.008	0.011*	-0.012†	0.025**	0.009	0.021†	0.010	0.026
13910	Sutures Natural Nonabsorbable Silk	2	0.009	0.007	0.029	0.009	0.028	-0.022	-0.047	0.010	0.002	0.013*	0.012	0.044	-0.021	-0.075*
14085	Tubes Tracheal	2	0.005	0.015**	0.023†	0.008	0.011	-0.046†	-0.011	0.002	0.009	0.060**	0.016	-0.022	-0.086**	-0.056
16071	Lenses Intraocular Posterior Chamber	2	0.008*	0.009	0.011*	0.012**	0.007	0.005	0.003	0.008**	0.028**	0.014**	0.009†	0.012**	-0.002	0.003
16078	Nails Bone	2	0.009	0.002	0.024†	0.017†	0.027†	-0.024	-0.015	0.010	0.002	0.029†	0.012	0.033†	-0.028	-0.018
16104	Wires Bone	2	0.013	0.053**	0.017*	0.027†	0.013	-0.031	-0.040	0.021*	0.038	0.023	0.026**	0.013	-0.001	0.023
16655	Catheters Cardiac Ablation	2	0.000	0.011	-0.003	0.008	0.002	-0.017**	-0.032**	0.000	-0.001	-0.001	-0.005	0.006	0.001	-0.006
17184	Catheters Vascular Angioplasty Balloon	2	-0.005**	0.003	0.007*	0.004	-0.009**	-0.030†	-0.035†	-0.006**	0.002	0.008**	0.003	-0.007	-0.032†	-0.047†
17471	Sutures Synthetic Absorbable Polyglactin	2	-0.025	-0.056	-0.008	-0.016	0.041	-0.110	0.214	0.033	-0.004	-0.045	0.047	0.152**	0.002	0.037
17600	Bone Matrix Implants	2	-0.011**	-0.007	-0.027**	-0.007	-0.002	-0.007	-0.027**	-0.007	0.028	-0.048†	-0.001	-0.021*	0.012	-0.003
17846	Catheters Vascular Guiding	2	0.004	0.010	0.016**	0.003	-0.002	-0.004	0.003	0.006	0.009	0.021*	0.002	-0.005	0.003	0.009
18609	Catheters Cardiac Mapping/Ablation	2	0.005	0.042*	0.016	-0.035*	0.009	-0.003	0.004	-0.002	0.029	0.001	-0.053**	0.023*	-0.023	-0.055
20317	Staplers Surgical	2	0.001	-0.030	0.018	0.015	0.032	-0.034	0.007	-0.001	-0.045*	0.021	0.010	0.027	-0.032	0.043
20318	Staplers Surgical Linear	2	0.001	-0.005	-0.049	0.015	0.020*	-0.002	0.056	0.001	0.019**	0.001	-0.008	0.030**	-0.032*	0.004
20453	Stent/Grafts Vascular Aortic	2	-0.000	-0.005	-0.004	0.014	-0.001	0.001	-0.004	-0.003	-0.017	-0.011	0.002	0.002	0.004	-0.002
22538	Suture Anchors	2	-0.000	0.005	0.005	0.009	0.000	-0.008	-0.001	-0.001	0.007	0.009	0.009	-0.001	-0.014	-0.007
22908	Sutures Synthetic Nonabsorbable Nylon Monofilament	2	-0.004	0.046**	0.061	-0.003	-0.019	-0.068**	-0.047	-0.011	0.047	0.077	-0.026	-0.040**	-0.073†	-0.004
34196	Screws Bone Spinal	2	-0.034**	0.004	-0.037	-0.001	-0.083†	-0.061**	-0.016	-0.064**	-0.014	-0.068	-0.003	-0.117**	-0.145**	-0.231**
	Other Med/Surg	2	-0.003	0.003	0.001	0.003	-0.004	-0.017†	-0.006	-0.005	0.001	-0.002	0.001	-0.004	-0.025†	-0.038†
12324	Lenses Intraocular	3	0.001	0.051**	0.012	0.010*	0.012*	-0.009*	-0.008	0.002	0.105**	0.000†	0.002**	0.001†	0.000	-0.000
12913	Pacemakers Cardiac Implantable	3	0.009*	0.035†	0.012**	-0.003	0.009	-0.028**	-0.028	0.009*	0.038†	0.012*	-0.004	0.012	-0.031**	-0.036
15766	Orthopedic Internal Fixation Systems Spinal	3	-0.017*	0.003	-0.004	-0.012	-0.050**	-0.034**	-0.041*	-0.021	-0.011	0.016	-0.042**	-0.063**	-0.013	0.026
15870	Prostheses Cardiac Valve Biological	3	-0.002	0.008	0.008	0.004	-0.015*	-0.015	0.003	-0.006	-0.003	0.020	0.006	-0.016	-0.047	-0.025
16040	Stents Ureteral	3	-0.000	0.005	-0.007	0.008*	0.008	-0.020*	-0.015	0.003	0.003	0.002	0.004	0.017**	-0.022	-0.036
16084	Prostheses Joint Hip Acetabular Component	3	-0.006	0.044†	0.009	-0.013	-0.025†	-0.086†	-0.102†	-0.012	0.044†	0.004	0.014	-0.040**	-0.123**	-0.197†
16095	Prostheses Joint Hip Femoral Component	3	0.004	0.047†	0.020†	-0.002	-0.024**	-0.065†	-0.049**	0.003	0.056†	-0.003	-0.017	-0.003	-0.070†	-0.078*
16097	Prostheses Joint Knee Femoral Component	3	0.003	0.045†	0.014**	0.006	-0.012	-0.056†	-0.058†	-0.002	0.052†	0.015**	-0.008	-0.027	-0.074†	-0.107**
16098	Prostheses Joint Knee Tibial Component	3	-0.002	0.041†	0.017**	0.006	-0.020**	-0.079†	-0.048**	-0.009	0.019*	0.000	-0.011	-0.013	-0.062†	0.010
16111	Prostheses Mammary Internal	3	0.003	0.025†	0.005	0.016**	-0.002	-0.014	-0.016	0.004	0.030†	0.005	0.014**	-0.004	-0.012	-0.020
16921	Prostheses Joint Shoulder Humeral Component	3	-0.001	0.001	-0.002	0.003	-0.009	0.000	-0.028	-0.027	-0.046**	-0.045	0.004	0.004	-0.035	-0.107
17165	Allografts	3	0.008	0.043*	0.069*	0.028	-0.010	-0.033	-0.024	0.049	0.115	0.105	0.144**	0.008	-0.164	-0.690
17241	Stimulators Electrical Spinal Cord Analgesic	3	-0.006	-0.008	-0.017†	-0.012**	0.016**	-0.007	-0.027*	-0.002	-0.014	-0.014**	-0.001	-0.040**	-0.005	-0.026
18253	Grafts Skin Biological	3	-0.008*	-0.016**	-0.020	0.000	-0.016	0.003	0.011	-0.011*	-0.025**	-0.046**	0.001	0.021**	0.006	0.017
18504	Defibrillator/Cardioverter/Pacemakers Implantable	3	-0.013*	0.011	0.014†	-0.018	-0.038**	-0.042†	-0.028	-0.019**	0.009	0.038**	-0.037	-0.031	-0.051†	-0.065**
20376	Defibrillator/Cardioverter/Pacemakers Implantable Resynchronization	3	-0.002	0.013	0.010	0.006	-0.021	-0.017	-0.026	-0.012	-0.009	0.009	0.002	-0.038**	-0.011	-0.007
20383	Stents Vascular Coronary Balloon-Expandable Drug-Eluting	3	-0.013†	-0.013**	-0.018**	-0.008	-0.011	-0.019	-0.075†	-0.014†	-0.013**	-0.019**	-0.008	-0.011	-0.023*	-0.079†
20422	Stents Vascular Coronary Balloon-Expandable	3	-0.010*	-0.025**	0.002	0.009	0.005	-0.040†	-0.031**	-0.005	-0.018	0.006	0.014	0.009	-0.043†	-0.034*
22543	Occluders Vascular Intravascular Embolization Coil	3	0.004	0.008	-0.003	0.015*	-0.002	0.002	0.018	0.010	0.032†	-0.032	0.019*	0.017	0.002	0.005
923150	Orthopedic Fixation Systems Implantable Spine	3	-0.013**	0.007	0.016*	-0.014†	-0.023*	-0.056†	-0.047†	-0.037	-0.010	0.055	0.039	0.014	-0.194†	-0.298†
	PPIs	3	-0.005†	0.014†	0.006†	-0.003	-0.016†	-0.033†	-0.039†	-0.009†	0.010*	0.011**	-0.001	-0.015†	-0.025†	-0.081†

Authors' calculations from PriceGuide data, 2009-2014. Specifications include hospital-brand and brand-year fixed effects; alternative fixed effects in Table A11. Standard errors clustered at hospital-brand level. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Table A10 reported the results of our preferred specification, with controls for hospital-brand and brand-year fixed effects. This choice of specification is intended to accommodate trends in technology prices across a wide variety of product categories, some of which are purchased frequently and some infrequently. Table A11 below compares summary results with alternative fixed effects: hospital and brand fixed effects vs. hospital-brand fixed effects, and brand-specific linear trends vs. brand-year fixed effects vs. brand-month fixed effects. In each case, we show spend-weighted average treatment effects within each product class, for the unrestricted and restricted samples.

Table A11: Treatment Effects with Alternative Fixed Effects

| | Unrestricted Sample | | | | | | $|\mathcal{H}_j| \geq 100$ Sample | | | | | |
| | ATE | | TE_{high^p} | | $TE_{high^p,high^q}$ | | ATE | | TE_{high^p} | | $TE_{high^p,high^q}$ | |
	\$	%	\$	%	\$	%	\$	%	\$	%	\$	%
$\theta_h + \theta_j + \theta_t + \gamma_j(t - tmin_j)$												
Commodities	-0.10	-0.007	-3.48	-0.084	-3.75	-0.051	0.37*	0.004	-2.18†	-0.056†	-3.13†	-0.069†
	(0.23)	(0.014)	(993.66)	(5.768)	(333.35)	(2.256)	(0.22)	(0.007)	(0.46)	(0.016)	(0.81)	(0.019)
Other Med/Surg	0.91	-0.002	-32.69	-0.059	-39.58	-0.086	-2.73	-0.002	-25.82†	-0.049†	-43.12†	-0.075†
	(2.63)	(0.002)	(5023.61)	(2.844)	(1523.34)	(1.886)	(3.08)	(0.003)	(6.93)	(0.010)	(11.50)	(0.015)
PPIs	36.78	-0.001	-349.07	-0.072	-506.62	-0.087	-29.31	-0.005	-237.32†	-0.069†	-486.63†	-0.108†
	(27.44)	(0.002)	(2514.22)	(3.420)	(3389.52)	(4.044)	(34.88)	(0.004)	(52.02)	(0.008)	(105.63)	(0.018)
$\theta_{hj} + \theta_t + \gamma_j(t - tmin_j)$												
Commodities	0.24	0.001	-0.36	-0.009	-0.33	-0.009	0.50**	0.010	-0.09	-0.012	-0.50	-0.038†
	(0.19)	(0.004)	(0.41)	(0.006)	(0.60)	(0.011)	(0.20)	(0.006)	(0.45)	(0.012)	(0.83)	(0.011)
Other Med/Surg	0.16	-0.000	-6.36**	-0.014†	-0.16	-0.002	-3.30	-0.002	-12.10**	-0.021†	-25.15†	-0.033†
	(2.24)	(0.002)	(3.06)	(0.004)	(5.47)	(0.006)	(2.93)	(0.003)	(5.86)	(0.008)	(9.23)	(0.012)
PPIs	30.96	0.002	-123.25†	-0.025†	-92.65	-0.026†	-27.72	-0.003	-198.47†	-0.045†	-276.28†	-0.073†
	(21.80)	(0.002)	(42.74)	(0.004)	(57.87)	(0.005)	(31.97)	(0.003)	(51.68)	(0.008)	(89.28)	(0.017)
$\theta_h + \theta_{jy}$												
Commodities	-0.05	-0.001	-3.35†	-0.072†	-3.88†	-0.073*	0.28	0.002	-2.32†	-0.064†	-3.19†	-0.082†
	(0.21)	(0.013)	(0.43)	(0.027)	(0.66)	(0.040)	(0.21)	(0.006)	(0.44)	(0.016)	(0.75)	(0.023)
Other Med/Surg	-1.81	-0.005**	-35.43†	-0.059†	-43.53†	-0.078†	-5.06	-0.006*	-29.46†	-0.054†	-47.27†	-0.081†
	(2.58)	(0.002)	(4.03)	(0.006)	(5.43)	(0.016)	(3.10)	(0.003)	(6.56)	(0.009)	(15.05)	(0.015)
PPIs	-36.00	-0.008†	-417.52†	-0.081†	-601.89†	-0.100†	-79.45**	-0.011†	-295.05†	-0.077†	-568.09†	-0.117†
	(25.31)	(0.002)	(48.48)	(0.004)	(84.13)	(0.007)	(32.25)	(0.003)	(51.81)	(0.008)	(105.52)	(0.017)
$\theta_{hj} + \theta_{jy}$												
Commodities	0.17	0.002	-0.52	-0.013†	-0.58	-0.016**	0.39*	0.007	-0.28	-0.019	-0.72	-0.053†
	(0.19)	(0.004)	(0.39)	(0.005)	(0.56)	(0.008)	(0.20)	(0.006)	(0.42)	(0.012)	(0.76)	(0.017)
Other Med/Surg	-1.86	-0.003	-8.72†	-0.017†	-8.89**	-0.006	-5.44*	-0.005	-14.87**	-0.025†	-28.09†	-0.038†
	(2.22)	(0.002)	(3.10)	(0.004)	(3.90)	(0.005)	(3.03)	(0.003)	(5.80)	(0.008)	(10.01)	(0.012)
PPIs	-38.54*	-0.005†	-189.55†	-0.033†	-206.32†	-0.039†	-78.09†	-0.009†	-246.15†	-0.052†	-342.61†	-0.081†
	(23.09)	(0.002)	(46.82)	(0.004)	(74.41)	(0.005)	(30.10)	(0.003)	(51.66)	(0.008)	(90.21)	(0.016)
$\theta_h + \theta_{jm}$												
Commodities	0.04	-0.003	-3.11†	-0.058**	-3.91†	-0.061	0.39*	0.003	-2.22†	-0.059†	-3.17†	-0.080†
	(0.23)	(0.011)	(0.46)	(0.023)	(0.77)	(0.041)	(0.22)	(0.007)	(0.46)	(0.016)	(0.77)	(0.022)
Other Med/Surg	-1.42	-0.003	-33.37†	-0.057†	-46.54†	-0.081†	-2.53	-0.003	-25.73†	-0.049†	-47.17	-0.083†
	(3.04)	(0.003)	(6.37)	(0.009)	(8.30)	(0.016)	(3.42)	(0.004)	(7.21)	(0.010)	(62.15)	(0.021)
PPIs	46.24	-0.000	-351.19†	-0.077†	-561.28†	-0.095†	-28.86	-0.006	-220.50†	-0.069†	-466.85†	-0.105†
	(31.32)	(0.002)	(57.60)	(0.005)	(97.90)	(0.008)	(38.35)	(0.004)	(56.23)	(0.009)	(103.52)	(0.018)
$\theta_{hj} + \theta_{jm}$												
Commodities	0.28	0.002	-0.36	-0.013†	-0.68	-0.018**	0.49**	0.009	-0.15	-0.015	-0.68	-0.052†
	(0.20)	(0.003)	(0.42)	(0.005)	(0.66)	(0.007)	(0.21)	(0.006)	(0.44)	(0.013)	(0.80)	(0.017)
Other Med/Surg	-1.34	-0.002	-8.87*	-0.019†	-11.80	-0.013	-3.69	-0.003	-12.47**	-0.022†	-29.35**	-0.042**
	(2.57)	(0.003)	(4.99)	(0.007)	(8.13)	(0.011)	(3.05)	(0.003)	(5.90)	(0.008)	(12.77)	(0.018)
PPIs	29.25	0.002	-150.63†	-0.030†	-170.29*	-0.035†	-29.29	-0.003	-198.75†	-0.045†	-284.08†	-0.072†
	(28.57)	(0.002)	(55.41)	(0.005)	(101.30)	(0.007)	(35.43)	(0.004)	(54.51)	(0.008)	(94.47)	(0.017)

Authors' calculations from PriceGuide data, 2009-2014. $N_{hjt}^{(1)} = 516,582$; $N_{hjt}^{(2)} = 1,344,515$; $N_{hjt}^{(3)} = 703,544$. $N_h^{(1)} = 748$; $N_h^{(2)} = 701$; $N_h^{(3)} = 601$. Fixed effects for each panel indicated in first column. Standard errors in category-specific regressions clustered at hospital-brand level. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

The results indicate that inclusion of hospital-brand fixed effects has an important effect

on the estimates – in the first panel, which includes hospital and brand and month fixed effects, as well as brand-specific linear trends, the standard errors are quite large, capturing the fact that many product categories in the expanded analysis are more heterogeneous than drug-eluting stents and estimates can be sensitive to compositional changes. Within the set of results with hospital-brand fixed effects, the results with different controls for time trends are qualitatively similar: for example, among PPIs, the high-price, high-quantity hospital-brand results range from −0.025 to −0.039 in the unrestricted sample, and from −0.073 to −0.081 in the restricted sample. Similarly, the unrestricted sample treatment effects for high-price, high-quantity hospital-brands range from −0.009 to −0.018 for commodities and from −0.002 to −0.013 for other medical/surgical products.

Finally, Table A12 examines sensitivity of our preferred fixed effects specification to whether it is estimated as a pooled regression of all price quintiles together (left panel) vs. separate regressions of each price quintile (right panel). In each panel, we show results for the bottom and top quintiles of the price distribution, for all hospital-brands, low-quantity hospital-brands, and high-quantity hospital-brands.[59]

Several patterns emerge in this comparison. First, the results for the top quintile of the price distribution are both qualitatively and quantitatively similar regardless of whether the specification is estimated separately by quintile or pooled. The top-quintile average effects by class are −0.013 for commodities, −0.017 for other medical/surgical, and −0.033 for PPIs in the pooled specification; these estimates change to −0.012 for commodities, −0.012 for other medical/surgical, and −0.026 for PPIs in the separate-quintiles specifications. While the increased richness of the control set in the separate-quintiles specifications has the expected effect of decreasing the precision of the estimates and reduces the magnitudes slightly, the general findings are robust to specification: we document small, significant effects of transparency; larger effects for PPIs; and particularly large effects for relatively high-quantity purchasers of PPIs. We also note that the qualitative patterns across product categories are similar across specification: e.g., some of the largest effects are for high-price, high-quantity joint prostheses.

Second, the results for the bottom quintile of the price distribution are sensitive to specification. The bottom-quintile average effects by class are 0.013 for commodities, 0.003 for other medical/surgical, and 0.014 for PPIs in the pooled specification; however, these estimates change to −0.005 for commodities, −0.009 for other medical/surgical, and −0.007 for PPIs in the separate-quintiles specifications. The qualitative pattern across product

---

[59]The primary difference between these specifications is that the pooled regression imposes the same brand-year trends across hospitals in all quintiles of the price distribution, whereas the non-pooled version allows brand-year trends to vary based on hospitals' position in the price distribution.

81

categories also changes: many of the positive and significant bottom-quintile results for joint prostheses observed in the left panel shrink dramatically or disappear in the right panel. These differences in the bottom-quintile results may be due to differential brand-specific price trends for low-price hospitals. For example, if low-price hospitals are those that have relatively high bargaining parameters, it may be the case that those hospitals achieve high discounts regardless of changes in the market environment and accordingly, exhibit less steep downward trends than other hospitals.

Table A12: Treatment Effects for Important Product Categories, Alternative Quintiles Treatments

ID	Name	Class	Pooled Regressions							Separated Regressions by Price Quintile						
			ATE	All Q		Low Q		High Q		ATE	All Q		Low Q		High Q	
				1	5	1	5	1	5		1	5	1	5	1	5
10279	Bandages Elastic	1	-0.011	0.007	-0.01	-0.017	-0.005	0.038	-0.018	-0.011	-0.008	-0.012	-0.021*	-0.01	0.019	-0.015
11315	Dressings	1	0.000	0.021**	-0.021	0.027*	-0.02	0.014**	-0.022*	0.000	0.007	-0.017	0.017	-0.014	-0.002	-0.023
11883	Gloves Surgical	1	-0.005	-0.018	-0.016	-0.017	-0.019	-0.019	-0.009	-0.005	-0.008**	-0.015	-0.008**	0.026	-0.010*	-0.004
12368	Drapes Surgical	1	0.007	0.004	-0.024	-0.005	-0.033*	0.016	0.001	0.007	-0.013	-0.02	-0.018	-0.006	-0.001	0.006
12745	Needles Injection Hypodermic	1	0.000	0.014**	-0.020**	0.008	-0.018*	0.026*	-0.028	0.000	-0.003	-0.016*	-0.003	0.002	0.000	-0.028
13655	Containers Specimen	1	0.018	0.082	0.011	0.174	0.031*	-0.029	-0.045**	0.018	-0.014	0.01	-0.003	0.005	-0.017	-0.050**
923840	Drill Bits	1	0.004	0.019†	-0.011	0.009	-0.01	0.032†	-0.012	0.004	0.002	-0.01	-0.001	0.002	0.007	-0.016
	Commodities	1	0.002	0.013*	-0.013†	0.021	-0.012**	0.005	-0.016**	0.002	-0.005*	-0.012**	-0.004	0.006	-0.003	-0.015**
10688	Catheters Vascular Angiography	2	0.004	0.025**	-0.013	0.019	0.021	0.035*	-0.078**	0.004	0.017	-0.01	0.021**	0.015	0.018	-0.063**
10894	Clip Appliers	2	0.008**	0.035**	-0.008	0.026	0.004	0.052*	-0.043**	0.008**	0.011	-0.006	0.003	0.021**	0.033	-0.050**
11502	Forceps Electrosurgical	2	0.000	-0.002	-0.001	0.007	0.000	-0.006	-0.001	0.000	-0.004	0.000	-0.017	0.001	-0.001	0.001*
11910	Grafts Bone	2	-0.001	-0.008	0.000	-0.004	0.000	-0.011	0.000	-0.001	-0.020†	0.001	-0.038**	0.001	-0.012	0.002
11925	Guide Wires	2	0.004	0.007	-0.001	0.008	-0.004	0.006	0.005	-0.004	-0.009**	-0.019	-0.007	0.000	-0.012*	0.004
12830	Orthopedic Cement	2	-0.004	-0.007	-0.017	0.017	-0.022	-0.035	-0.005	-0.004	-0.006	-0.019	-0.009	-0.001	-0.003	-0.021
13050	Plates Bone	2	-0.003	0.004	-0.017**	0.001	-0.011	0.008	-0.026**	-0.003	-0.011**	-0.018**	-0.012*	0.007	-0.009	-0.031**
13909	Sutures Synthetic Nonabsorbable Polypropylene	2	0.000	-0.019	-0.017	-0.004	-0.027*	-0.049	0.008	0.000	-0.103*	-0.013	-0.101	-0.015	-0.076	-0.006
13910	Sutures Natural Nonabsorbable Silk	2	0.009	0.007	-0.022	0.002	-0.016	0.015	-0.047	0.009	-0.014*	-0.016	-0.025†	0.006	0.001	-0.072
14085	Tubes Tracheal	2	0.005	0.015**	-0.046†	0.01	-0.060†	0.020**	-0.011	0.005	-0.003	-0.026*	-0.005	-0.024	-0.004	0.021
16071	Lenses Intraocular Posterior Chamber	2	0.008*	0.009	0.005	0.016	0.006	0.003	0.003	0.008*	0.01	0.008	0.011	0.005	0.011	0.011
16078	Nails Bone	2	0.009	0.002	-0.024	-0.013	-0.031*	0.016*	-0.015	0.009	-0.01	-0.019	0.024**	-0.015	0.003	-0.007
16104	Wires Bone	2	0.013	0.053**	-0.031	0.055*	-0.026	0.050**	-0.04	0.013	-0.006	-0.014	-0.015	-0.03	0.018	-0.015
16655	Catheters Cardiac Ablation	2	0.000	0.011	-0.017**	-0.015	-0.01	0.036	-0.032**	0.000	0.011	-0.012	0.009	0.002	0.024	-0.028*
17184	Catheters Vascular Angioplasty Balloon	2	-0.005**	0.003	-0.030†	0.001	-0.028†	0.006	-0.035†	-0.005**	-0.015†	-0.024†	-0.018†	-0.015**	-0.011†	-0.029†
17471	Sutures Synthetic Absorbable Polyglactin	2	-0.025	-0.056	-0.11	0.037	-0.184**	-0.122	0.214	-0.025	-0.075	-0.091	-0.007	-0.107	-0.081	0.289
17600	Bone Matrix Implants	2	-0.011**	-0.007	-0.007	-0.017	0.002	0.003	-0.027**	-0.011**	-0.006	-0.007	-0.022	0.004	0.007	-0.025**
17846	Catheters Vascular Guiding	2	0.004	0.01	-0.004	0.019**	-0.006	-0.001	0.003	0.004	0.005	-0.005	0.011	-0.003	0.005	0.006
18609	Catheters Cardiac Mapping/Ablation	2	0.005	0.042*	-0.003	0.057**	-0.005	0.013	0.004	0.005	-0.004	0.003	-0.01	-0.01	-0.024	0.016
20317	Staplers Surgical	2	0.001	-0.03	-0.034	-0.032	-0.049	-0.029	0.007	0.001	-0.024	0.032	-0.036†	0.036	-0.014	0.002
20318	Staplers Surgical Linear	2	0.001	-0.005	-0.002	-0.019	-0.034**	0.012	0.056	0.001	-0.014	-0.003	-0.025	0.005	0.000	0.037
20453	Stent/Grafts Vascular Aortic	2	0.000	-0.005	0.001	-0.015	0.003	0.004	-0.004	0.000	-0.002	0.002	0.004	0.001	-0.003	0.004
22538	Suture Anchors	2	0.000	0.005	-0.008	0.006	-0.011	0.004	-0.001	0.000	0.000	-0.005	0.007	-0.009	-0.008	0.003
22908	Sutures Synthetic Nonabsorbable Nylon Monofilament	2	-0.004	0.046**	-0.068**	0.023*	-0.078**	0.078*	-0.047	-0.004	-0.039†	-0.034	-0.005	-0.021	-0.09	-0.028
34196	Screws Bone Spinal	2	-0.034*	0.004	-0.061**	0.019	-0.079**	-0.004	-0.016	-0.034**	-0.003	-0.062	0.01	-0.081	-0.006	-0.022
	Other Med/Surg	2	-0.003	0.003	-0.017†	0.004	-0.021†	0.002	-0.006	-0.003	-0.009†	-0.012†	-0.011	-0.009	-0.006*	-0.003
12324	Lenses Intraocular	3	0.001	0.051**	-0.009*	0.081**	-0.009	0.013	-0.008	0.001	0.04	-0.007	0.061	0.002	-0.002	-0.008
12913	Pacemakers Cardiac Implantable	3	0.009*	0.035†	-0.028**	0.038†	-0.028†	0.032†	-0.028	0.009*	-0.001	-0.009	-0.002	-0.005	-0.003	-0.005
15766	Orthopedic Internal Fixation Systems Spinal	3	-0.017*	0.003	-0.034**	0.015	-0.029	-0.005	-0.041*	-0.017*	-0.026**	-0.02	-0.031	-0.040**	-0.02	-0.027
15870	Prostheses Cardiac Valve Biological	3	-0.002	0.008	-0.015	0.011*	-0.021	0.005	0.003	-0.002	-0.003	-0.012	-0.005	0.004	-0.002	-0.014*
16040	Stents Ureteral	3	0.000	0.005	-0.020*	0.006	-0.023*	0.003	-0.015	0.000	-0.002	-0.016	0.005	-0.011	-0.039	-0.019
16084	Prostheses Joint Hip Acetabular Component	3	-0.006	0.044†	-0.086†	0.047†	-0.072†	0.042†	-0.102†	-0.006	0.001	-0.079†	0.006	-0.037**	-0.004	-0.108†
16095	Prostheses Joint Hip Femoral Component	3	0.004	0.047†	-0.065†	0.037†	-0.080†	0.053†	-0.049**	0.004	0.003	-0.057†	0.001	-0.038**	0.003	-0.048**
16097	Prostheses Joint Knee Femoral Component	3	0.003	0.045†	-0.056†	0.059†	-0.055†	0.037†	-0.058†	0.003	0.012*	-0.052†	0.015	-0.025*	0.011*	-0.048**
16098	Prostheses Joint Knee Tibial Component	3	-0.002	0.041†	-0.079†	0.052†	-0.099†	0.034†	-0.048**	-0.002	0.013**	-0.061†	0.007	-0.074†	0.01	-0.029*
16111	Prostheses Mammary Internal	3	0.003	0.025†	-0.014	0.033†	-0.013	0.015*	-0.016	-0.002	0.009	-0.008	0.021†	0.001	-0.005	-0.011
16921	Prostheses Joint Shoulder Humeral Component	3	-0.001	0.001	0.000	0.015	-0.013	-0.007	-0.028	-0.001	-0.001	-0.016	-0.003	-0.015	-0.003	-0.038
17165	Allografts	3	0.008	0.043*	-0.033	0.069*	-0.04	0.02	-0.024	0.008	0.03	-0.026	0.069	-0.015	-0.013	-0.003
17241	Stimulators Electrical Spinal Cord Analgesic	3	-0.006	-0.008	-0.007	-0.019	0.003	0.000	-0.027*	-0.006	-0.017†	0.006	-0.002	-0.028	-0.018†	-0.039**
18253	Grafts Skin Biological	3	-0.008†	-0.016**	0.003	-0.017*	-0.002	-0.008	0.011	-0.008**	-0.023**	0.007	-0.027**	0.005	0.001	0.019†
18504	Defibrillator/Cardioverter/Pacemakers Implantable	3	-0.013*	0.011	-0.042†	0.007	-0.045**	0.013*	-0.028	-0.013**	-0.004	-0.034**	-0.006	-0.046†	0.002	-0.018
20376	Defibrillator/Cardioverter/Pacemakers Implantable Resynchronization	3	-0.002	0.013	-0.017	-0.006	-0.014	0.031	-0.026	-0.002	-0.006	-0.021	-0.008	-0.033**	-0.001	-0.02
20383	Stents Vascular Coronary Balloon-Expandable Drug-Eluting	3	-0.013†	-0.013†	-0.019	-0.015**	-0.005	-0.009	-0.075†	-0.013†	-0.020†	-0.014	-0.022†	-0.009	-0.013*	-0.051†
20422	Stents Vascular Coronary Balloon-Expandable	3	-0.010†	-0.025**	-0.040†	-0.017*	-0.046†	-0.039**	-0.031**	-0.010**	-0.032†	-0.019	-0.032†	-0.015	-0.039†	-0.015
22543	Occluders Vascular Intravascular Embolization Coil	3	0.004	0.008	0.002	0.027†	-0.009	-0.018	0.018	0.004	-0.003	0.002	0.009	-0.007	-0.018	0.022
923150	Orthopedic Fixation Systems Implantable Spine Interbody Fusion	3	-0.013**	0.007	-0.056†	0.015	-0.062†	0.000	-0.047†	-0.013**	-0.025*	-0.043†	-0.031	-0.037†	-0.032*	-0.038†
	PPIs	3	-0.005†	0.014†	-0.024	0.015	-0.033†	0.014†	-0.039†	-0.005†	-0.007**	-0.026†	-0.006*	-0.024	-0.006*	-0.030†

Authors' calculations from PriceGuide data, 2009-2014. Specifications include hospital-brand and brand-year fixed effects; alternative fixed effects in Table A11. Left panel: estimates from pooled regression of all price quintiles together. Right panel: estimates from separate regressions of each price quintile. Standard errors clustered at hospital-brand level. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

83

In sum, we document a robust pattern of significant savings due to transparency among previously high-priced hospital-brands, which are particularly strong for relatively high-quantity hospital-brands. Given the stability of these focal estimates, in the main text, we report the results of our preferred pooled-quintiles specification with hospital-brand and brand-year fixed effects, which tend to have smaller standard errors due to the sparsity of monthly data for some brands in the expanded analysis. We also document some evidence consistent with an externality effect in which prices for low-price hospital-brands increase under transparency; these results are quite sensitive to specification, however.

F   Usage Pattern Changes and Demand

Stents and certain other expensive medical technologies are "physician preference items" where physician demand is based on strong preferences and is relatively insensitive to price – Grennan (2013) estimates own-elasticities centered around -0.32 for bare metal stents and -0.52 for drug-eluting stents. In general, however, the price benchmarking information treatment could influence quantities as well. Here, we perform a set of analyses to provide a check of this hypothesis and also provide proof of concept for how this analysis might be incorporated in the case of products where demand is more sensitive to price.

There are two primary ways in which quantities might adjust to benchmarking price information and subsequent renegotiations: (1) In a context where contracts specify quantities or market shares in addition to price, renegotiations to obtain better prices might also involve large quantity or share commitments—this effect was tested and ruled out in our analysis in Appendix C. (2) In a context where quantity is responsive to price, negotiation of better prices would lead to increased usage in the brands with the largest relative price decreases. We analyze this second case here.

We run the regression specifications allowing for heterogeneous treatment effects of information depending on pre-join prices and quantities, but here with quantity q_{hjt} as the dependent variable (results are qualitatively similar and so unreported for market shares and log transformations):

$$
\begin{aligned}
Q_{hjt} \;=\; & \beta^{Info}_{quintile} * \mathbb{1}_{\{post_{hjt}\}} * \mathbb{1}_{\left\{quintile_{hj,pre}\right\}} \\
& + \theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j}) + \varepsilon_{hjt}
\end{aligned}
$$

where $\beta^{Info}_{quintile}$ estimates the treatment effect, by price quintile. The results are shown in Table A13.

If quantity were responsive to price (with downward sloping demand), then we would expect quantity/share increases in exactly the areas we see relative price decreases. Because information leads to decreases in prices for brands in the high-price, high-quantity part of the pre-information distribution, we should expect potential quantity increases for those brands and decreases for other brands (whose prices haven't changed, but have become higher relative to the brands with price decreases). This is not the case in Table A13, where some specifications indicate that quantities may be changing slightly post join, but no specification shows that these changes vary significantly (economically or statistically) *as one moves across the pre-join price distribution.*

Version	Pre-info price quintiles ($\beta_{quintile}^{Info}$)				
	1	2	3	4	5
$\theta_h + \theta_j + \theta_t + \gamma_j * (t - t_{min_j})$	0.2	1.4	1.3	0.7	2.4**
	(1.5)	(1.9)	(1.7)	(1.7)	(1.1)
$\theta_h + \theta_{jt}$	−0.1	2.0	1.6	1.0	2.9**
	(1.6)	(2.0)	(1.9)	(1.8)	(1.3)
$\theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j})$	3.2†	2.7	3.8†	2.9**	3.0†
	(1.3)	(1.7)	(1.3)	(1.3)	(1.1)
$\theta_{hj} + \theta_{jt}$	4.0†	3.9*	5.2†	4.2†	5.0†
	(1.5)	(2.0)	(1.5)	(1.5)	(1.3)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,453$ member-brand-months. Includes 508 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Table A13: Treatment Effect *on Quantity* Estimates Across the Price Distribution

F.1 Censoring

The data used in this study is based on purchase orders issued by hospitals, so the data are "censored" in the sense that we only observe a price when a transaction takes place for a given brand-hospital-month. In our final analyses, we investigate whether this censoring could be a potential source of bias – e.g., if is correlated with the pricing process in a way that affects the interpretation of our results.

There are three distinct cases of these censored ($q_{hjt} = 0$, $p_{hjt} = .$) observations: (1) brands the hospital never contracts with; (2) brands a hospital contracts with at some point, but not the full time the brand is available; (3) brands that are currently contracted with, but not purchased in a given month due to sampling variation.

Cases (1) and (2) relate to the issue of whether hospitals contract with the "full support" of manufacturers, which is analyzed in Appendix C as a part of our analysis of potential standardization and share-based contracts. There, we document that almost half of hospitals contract for the "full support" set of manufacturers/brands (among the set of larger hospitals where sampling variation is less likely a problem). Of course, one cannot observe prices for the censored cases, so there is no way to compare prices for censored vs. uncensored observations. However, we can test for patterns in prices across hospitals that contract more vs. less with a given manufacturer. There we find no economically or statistically significant evidence that prices vary when hospitals contract exclusively with or devote more of their market share to a manufacturer (or brand).This holds in specifications across hospitals and within hospitals (and hospital-brand pairs). That we find no relationship on this margin we

can observe provides us some comfort regarding the margin we cannot.

Our quantity analyses thus far provide suggestive evidence on cases (2) and (3): if quantity were responsive to price (with downward sloping demand), then we would expect quantity/share increases in exactly the areas we see relative price decreases. This was not the case in Table A13. However, we find it instructive to dig deeper on points (2) and (3) by examining censoring explicitly: we examine the reduced form association between each hospital's position in the pre-join price and quantity distributions and the degree of censoring before and after join. That is, for each hospital-brand for which we were able to determine relative pre-join price and quantity position, we generated a panel of all months in which that hospital and brand existed in the database, with a dummy indicating whether we observed nonzero purchase for that hospital-brand.

The results are shown in Table A14. First, we note that, although the quintiles of the price distribution are mutually exclusive, this analysis only includes hospital-brands that were observed pre-join – thus, we generally observe fewer nonzero purchases (negative coefficients) for all hospital-brands in the post-join period as hospitals switch to newly entering brands. Second, the largest reduced form effects are in the middle of the pre-join price distribution: data are not differentially censored post-join for brands that were found to have relatively high prices. To put these numbers in perspective, consider the largest coefficient: -0.15 for the middle quintile hospital-brands in the preferred specification. Post-join, hospital-brands in the middle of the pre-join price distribution were 15 percentage points more likely to be censored; overall, observations for that quintile have nonzero purchase 49 percent of the time, and full discontinuation of all usage of those brands in the post-join period would decrease the nonzero purchase rate to 18 percent. That is, full discontinuation would more than double the censoring rate we observe. More to the point, we do not read too much into the average rate of censoring/discontinuation these results reveal, but rather focus on the fact that high-priced hospital-brands exhibit slower censoring post-join than mid-range hospital-brands.

Version	Pre-info price quintiles ($\beta_{quintile}^{Info}$)				
	1	2	3	4	5
$\theta_h + \theta_j + \theta_t + \gamma_j * (t - t_{min_j})$	-0.06	-0.01	$-0.11^\dagger$	$-0.11^\dagger$	-0.07^*
	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
$\theta_h + \theta_{jt}$	-0.07^*	-0.00	-0.07^{**}	$-0.12^\dagger$	-0.04
	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)
$\theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j})$	-0.03	0.01	$-0.15^\dagger$	$-0.10^\dagger$	-0.03
	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
$\theta_{hj} + \theta_{jt}$	-0.04	0.00	$-0.10^\dagger$	$-0.10^\dagger$	0.00
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)

Authors' calculations from PriceGuide data, 2009-2014. $N = 30,183$ member-brand-months. Identification focuses on timing of join; sample contains 331 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Table A14: Reduced Form Effect of Join on Censoring

88

ELECTRONIC APPENDICES – NOT FOR PRINT PUBLICATION

A Data Appendix

The raw transactions data contain 82.5 million observations for 2,111 members across 3,375 product categories and 1.9 million SKUs. Our main analyses focus on coronary stents; our expanded analysis includes 52 other important product categories. For stents, we assign product IDs to each branded product by examining each manufacturer's catalog and searching for brand and model names within the free-text item description field in the transaction database. The procedure for assigning brand IDs in the expanded sample is detailed in Appendix A.3 below.

Table A1 presents details for our main analysis sample of facilities purchasing coronary stents. The top panel summarizes facilities purchasing coronary stents: the average sample facility submitted transactions in 41 months, spending $3.4 million per month on 1,143 product categories. The bottom panel examines stent purchasing patterns by type: the average sample facility spent $80,000 on 59 stents per month. 82 percent of stents purchased were drug-eluting, as opposed to the older, bare metal technology. Stents comprised approximately 2 percent of overall monthly spending.

Table A1: Summary Statistics of Stent Purchase Orders

	All Stents [N=508]		Drug-Eluting [N=508]		Bare Metal [N= 504]	
	Mean	SD	Mean	SD	Mean	SD
Months of Data	41.4	21.4	41.4	21.4	41.6	21.4
Product Categories	1143.1	313.4	1143.1	313.4	1147.2	308.2
Total Spend/Month ($m)	3.4	3.2	3.4	3.2	3.4	3.2
% Hospital & Health System	0.97	0.17	0.97	0.17	0.97	0.18
Stent Spend/Month ($k)	80.4	73.8	73.1	69.2	7.4	7.0
Stent Qty/Month	58.7	52.7	48.4	46.1	10.4	9.8
Stent Brands/Month	3.8	1.4	2.3	0.9	1.5	0.7
Stent Mfrs/Month	2.1	0.7	1.9	0.7	1.4	0.6

Authors' calculations from PriceGuide data.

Table A2 shows each statistic separately by hospital bed count; larger hospitals generally submitted more months' data and, as logic would indicate, purchased more stents per month for a greater total monthly expenditure. Hospitals with ≥ 500 beds spent more than double the amount that the smallest hospitals did on stents per month. The vast majority of transactions in our data are for drug-eluting (as opposed to bare metal) stents; in the remainder of our description of the data and in our results, we focus on drug-eluting stents.

34

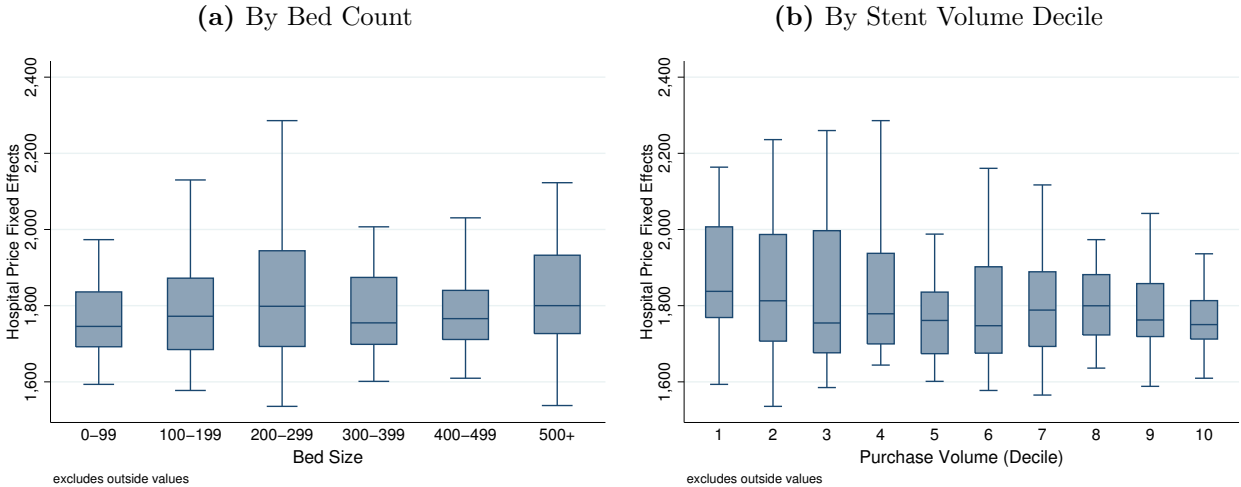**Table A2:** Summary Statistics – Stent Hospitals Only

Bed Size	Members	Months of Data	Monthly Exp. ($ k)	Monthly Quantity	% DES
0-99	52	31.4	59.0	45.0	82.0
		(19.4)	(56.6)	(44.4)	(10.8)
100-199	102	40.0	45.5	33.5	81.6
		(20.3)	(43.3)	(31.5)	(12.1)
200-299	117	43.4	55.6	40.7	77.4
		(22.0)	(45.9)	(33.5)	(14.3)
300-399	83	41.0	73.5	53.6	79.7
		(20.7)	(46.9)	(33.1)	(11.6)
400-499	47	41.4	128.9	93.5	79.6
		(21.3)	(92.1)	(65.5)	(12.2)
500+	107	45.9	135.2	97.7	81.1
		(22.0)	(94.2)	(65.6)	(9.5)

Authors' calculations from PriceGuide data.

The heterogeneity in prices observed in our sample is not well-explained by hospital characteristics that might seem a priori to be important for negotiation. For example, we observe no clear relationship between hospital size and stent prices. See the left panel of Figure A1, in which we display a box plot of drug-eluting stent prices for each category of bed count.[34] Mean prices are, if anything, increasing in bed count, though the differences are not statistically significant. Part of this (lack of) relationship could be due to the heterogeneity in purchasing behavior across hospitals with similar bed counts – e.g., small cardiac specialty hospitals may purchase stents in greater quantities than similarly-sized acute care hospitals. In the right panel of Figure A1, we also show box plots of stent prices for each decile of monthly stent purchasing volume. Here, we do see a relationship between "size" and price – the price distribution for hospitals with the smallest purchasing volumes is spread slightly upward relative to that of the hospitals with the largest volumes, so that low-volume hospitals' prices have larger means and variances than high-volume hospitals. For drug-eluting stents, 10th decile hospitals' prices are 7 percent lower than those obtained by 1st decile hospitals. These differences are economically and statistically significant; however, the price distributions for the high-volume and low-volume hospitals overlap substantially, so that there is a great deal of unexplained hospital price heterogeneity conditional on purchasing volume.

---

[34] "Prices" are hospital fixed effects obtained from a regression of price on hospital and brand-month fixed effects.

Figure A1: Distribution of Prices Across Hospitals

(a) By Bed Count

(b) By Stent Volume Decile

excludes outside values

excludes outside values

Drug-Eluting Stent Prices by Size Category (Regression Results)															
$\beta\mathbb{1}\{h\in \text{bed size } x\}$						$\beta\mathbb{1}\{h\in \text{decile } x\}$									
$0-99$	$100-199$	$200-299$	$300-399$	$400-499$	$500+$	1	2	3	4	5	6	7	8	9	10
1,785	1,801	1,852	1,778	1,805	1,829	1,916	1,864	1,830	1,832	1,766	1,784	1,801	1,807	1,787	1,774
(38)	(27)	(30)	(25)	(32)	(22)	(54)	(40)	(44)	(36)	(26)	(31)	(29)	(25)	(25)	(32)

Authors' calculations from PriceGuide data. Estimated mean hospital fixed effects within bed size categories and decile of monthly purchase volume. Hospital fixed effects obtained from regression of price on hospital and brand-month fixed effects, pre-join data only. Mean estimates from regression of fixed effects on indicators for size. Standard errors from nonparametric bootstrap of entire procedure, resampling at hospital level.

A.1 Representativeness of the Benchmarking Database Sample

The facilities in the purchase order data voluntarily joined a subscription service that allows them to benchmark their own prices and quantities to those of other members in the database and thus may not be a random sample of US hospitals. In particular, subscription is costly, so we expect hospitals with greater concerns about supply costs to be overrepresented in the database – for example, in a survey of database members, "cost reduction on PPIs" and "cost reduction on commodities" were the first and second (and nearly tied) most commonly cited reasons for joining. This is in accord with our own conversations with purchasing managers who cite a broad array of reasons and product areas as motivations for benchmarking. Here, we discuss representativeness of sample facilities.

The left panel of Figure A2 compares the distribution of sample hospitals across US census regions to American Hospital Association (AHA) member hospitals with cardiac catheterization labs. The Figure also compares our sample to that in the Millennium Research Group's (MRG) Markettrack™ survey of catheter labs. The MRG survey is explicitly intended to

provide an accurate picture of market shares and prices by US region. The Figure shows that, relative to both comparison samples, the west region is overrepresented in the benchmarking database sample, while the south is underrepresented. The average sample hospital is also larger than the average US hospital with cardiac catheterization capabilities – the right panel of Figure A2 shows that the sample contains disproportionately fewer hospitals in the < 200 beds range and disproportionately more hospitals in the ≥ 500 beds range, relative to AHA hospitals that would purchase stents. We do not have access to bed size for the MRG sample, but we do find that the member facilities in our estimation sample purchased in significantly higher volumes than MRG sample hospitals (60 vs. 33 stents per month).

Figure A2: Distribution of Benchmarking Database vs. Comparison Hospitals

(a) Across Census Regions **(b)** By Bed Size

Database sample computations from PriceGuide data, 2009-2014. AHA sample computations from American Hospital Association Annual Survey of Hospitals, 2012; hospitals with catheterization labs defined as those listed as having in-hospital adult or pediatric interventional or diagnostic catheterization services. MRG sample computations from Millennium Research Group survey of catheter labs, January 2010-June 2013.

For January 2010 through June 2013, we also have access to stent price data from the MRG survey. This allows us to further check the representativeness of the sample of hospitals joining the price benchmarking database. Figure A3 provides further details on the full distribution of prices across hospitals in the two samples of 143 pre-join database sample hospitals vs. 107 MRG hospitals. The prices paid in the two samples are statistically close to one another, with the average prices paid (controlling for brand-time trends) in the MRG sample slightly higher (mean \$1,666, s.d. \$149) than those paid by hospitals in the estimation sample (mean \$1,631, s.d. \$120) during the period before they joined the benchmarking service. These pre-join hospitals do have a slightly larger upper tail of high prices, with an 80th percentile of \$1,743 versus \$1,730 in the MRG sample, but this

difference is not statistically significant. The representation of larger facilities with slightly better negotiation outcomes ex ante in our sample may be due to small hospitals' limited ability to afford access to the database, though we would expect a countervailing effect to come from large hospitals' greater ability to purchase custom benchmarking services from consulting firms.

Figure A3: Pre-Join Distribution of Prices Across Hospitals: Comparison to MRG Sample

	mean	sd	cv= $\frac{\mu}{\sigma}$	$20^{th}_{\%ile}$	$50^{th}_{\%ile}$	$80^{th}_{\%ile}$	N_H
Pre-join $\bar{p}_h$ (within brand-month)	1631	120	0.07	1539	1607	1743	143
MRG $\bar{p}_h$ (within brand-month)	1666	149	0.09	1563	1651	1730	107

Authors' calculations from PriceGuide data and MRG survey. Sample restricted to Jan. 2010-Jun. 2013 when MRG data available. Hospital average prices with brand-month trends removed.

Despite potential selection into our sample, the estimation strategy we employ in this paper is internally valid in that it exploits the existence of pre/post data and staggered join dates within the sample of joiners (and uses no non-joiners) to estimate the effect of access to benchmarking information. These estimates are of direct interest, capturing the benefit of benchmarking for facilities that seek out such services.

A.2 Information Flags

In order to construct each member's information set upon joining the database and later upon new brands' entry, we linked the transactions data with several additional datasets relating each individual login ID with associated members and login activity. The first of these is a "clickstream" dataset containing timestamped observations of unique IDs' login

activity, to the minute.[35] The second is a membership "hierarchy" file linking individual database members with parent accounts for those cases where members are part of a set of health care organizations purchasing membership jointly. The third file associates each login ID with the given individual's direct-linked member organization and broader access level – i.e., the individual with ID X may work with member 1 but have access to the data for all members 1, 2, and 3 under the same parent organization. For individuals with higher-level access, data for all associated members is automatically reported to them when they log in to the database. Accordingly, when we observed a click for a particular login ID, we associated that click with all linked member organizations for which that login had access.

We used the above-described datasets to determine each date of login for each member. The goal of this exercise was to determine when benchmarking data on a given brand would enter each member's information set. For non-entering stent brands and all product categories in the expanded analyses, this is the date of the first observed login for each member. For entering stent brands, this is the date of the first observed login for each member after six months post-entry. This is to account for the lag between transactions occurring for new brands and transactions being submitted by member facilities, loaded into the database, and viewed by members logging in. The left panel of Figure A4 displays the steady increase over time in the count of members for which transactions for the average entering brand are observed, and demonstrates the lag with which members' transactions become available in the database for benchmarking purposes.[36] In any given month in the year following new brand entry, there are on average 56 more members we observe having transactions for new brands than there are members whose transactions data are actually loaded into the database. We also observe transactions data for members that have not yet joined the benchmarking database – in the year following brand entry, 9 percent of members observed in the average month are pre-join. To see this more concretely, the right panel of Figure A4 displays the trend in the number of hospitals purchasing the average new brand, overall and for pre-join hospitals in particular. For each new brand, we observe 10-15 hospitals in the pre-join state within a short window after brand entry, and the time period for our study contains many meaningful brand introductions. This is precisely what allows us to separately identify effects of *joining* the database per se, versus actually having access to information *on a particular brand*, on prices. In Figure 2 in the main text, we note the timing of entry of seven new drug-eluting stent brands between 2010 and 2014 (of the thirteen brands sold

---

[35]Each login ID corresponds to a unique individual's account within a member. For example, a given database member may have had login accounts for a number of different purchasing managers, administrators, and department clinicians.

[36]There may be an additional lag before joining hospitals become informed if they do not frequently log in to the database.

during this time period overall).

Figure A4: Transactions Observed After New Brand Entry

(a) Cumulative Members Purchasing

(b) Members Purchasing, by Join Status

Authors' calculations from PriceGuide data.

We use the linked login and transactions data to calculate each member-brand's position in the pre-information price and quantity distributions. All calculations are specific to the first *informed* login for the given member-brand, as defined above. Following the approach used in the database to aggregate data across all members' transactions and present summary statistics to members logging in at a point in time, we calculate percentiles of the price distribution using all members' *most recent* transactions for the same brand, in the past year, that would have been loaded into the database prior to the login. We calculate percentiles of the quantity distribution using all members' total quantity purchased per month for the same brand in the past year. Across all specifications, we consistently include only those observations that can be used to estimate the richest specification with interactions based on point in the price and quantity distributions – this requires that we observe pre-information data for the given member-brand.

This estimation sample for drug-eluting stents includes 395,271 transactions for 508 members and thirteen brands in 72 months between January 2009 and December 2014. Seven of the included drug-eluting stent brands entered the market during this time horizon. We collapse the transaction-level data to perform all analyses at the member-brand-month level (with mean price as the dependent variable). We do this in order to avoid overweighting member-brands that tend to have multiple transactions per month. The regression sample used in our main analyses focused on drug-eluting coronary stents contains 32,453 member-brand-month observations.

40

A.3 Other Important Products

For our work in Section 6, we generated analysis files for a large number of additional product categories, defined by their UMDNS codes. We restricted to the top 50 categories by spending or number of transactions, yielding 80 categories total. From these, we excluded categories that were too broad (e.g., "food item").[37] Next, creating the final analysis file required two key steps for each product category: (1) rationalizing the multiple units of measure in which different transactions' quantities were reported, in order to analyze price for a common quantity across transactions; and (2) generating brand IDs, in order to appropriately control for brand-specific price trends.

Regarding the first step, although many medical and surgical product categories are sold by the unit (a single coronary stent, e.g.), others are sold in pairs, boxes, cases, etc. The transactions data indicates this distinction in the "unit of measure" field, and further notes how many subunits are in each unit of measure using a "conversion factor" field. In order to perform our analyses on the cleanest and most internally-consistent transactions data possible, we transformed all transactions into price per single unit and quantity of single units purchased. We also excluded UMDNS codes with inconsistent or missing quantity data.[38]

Regarding the second step, the absence of a brand ID in the database creates a problem of sparsity, in which many SKUs are purchased by only a small number of hospitals, or in only a small number of months. The method we employed to identify brands within the coronary stents data involved examining manufacturer catalogs, finding likely brand names, searching for similar strings within the item description field, and validating SKUs for those brands against the catalog numbers. This was infeasible in the expanded analysis for several reasons: the average category has 30 manufacturers and more than 2,000 SKUs; many manufacturers' websites were found to be difficult to navigate, particularly once we extended the analysis beyond high-dollar physician preference items; and the item description field was often uninformative as to brand. Hence, we used an algorithmic approach to assign brand IDs for the other product categories.

The algorithm proceeded as follows: within each manufacturer, for each $k = 0, ..., K$, where K is the maximum SKU length for the manufacturer under consideration, we generated

---

[37]We did this based on "reasonableness" of the observed price variation – categories for which the coefficient of variation in price exceeded 10 were excluded – and selected categories by hand that seemed excessively broad based on their UMDNS names (e.g., "office supplies"). The list of product categories excluded by hand is 8889, 99936, 88885, 88884, 88883, 88695, 88539, 88311, 88073, and 16101.

[38]Specifically, we excluded UMDNS codes for which the conversion factor (e.g., ten units per box) was missing more than 1/3 of the time, or for which the modal unit of measure (e.g., "box" vs. "case") accounted for less than 50 percent of the data.

a candidate brand ID as the left-most k digits of each SKU. We then regressed price on dummies for each unique candidate brand ID, interacted with each hospital ID. We recorded the resulting R^2 for each k. The goal of this exercise was to identify brands that generate price variation in the data: looking within a given hospital, price variation proxies for contract variation. The R^2 is always increasing in k, so the algorithm is meant to choose k^{opt} that maximizes R^2 while keeping the number of identified "brands" relatively manageable. Table A3 below compares the algorithmic approach to the time-intensive hand-coding approach for coronary DES stents and for one manufacturer of surgical staplers whose catalog was available online. The first three rows show the number of brands implied by unique SKUs, vs. the hand-coding of manufacturer catalogs, vs. the algorithm. We also show the R^2 of a regression of price on hospital-brand dummies. The remaining rows show the detailed brand count and R^2 for each SKU digit count k. For the 52 products in our expanded analysis, we generally found that R^2 would climb steeply, then level off after a few SKU digits. For example, within stent manufacturer 1, R^2 increases from 0.42 at one SKU digit to 0.6 at three SKU digits, but only increases to 0.65 thereafter. Similarly, within linear surgical staplers, R^2 climbs from 0.61 to 0.95 as we move from one to three digits, eventually reaching a maximum of 0.97. These patterns are consistent with our intuition that the first few SKU digits would indicate a particular product line, while the last few SKU digits would indicate finer (and likely unpriced) distinctions such as color or size.

The machine learning literature has many methods for explicitly penalizing excessive complexity in classification and other modeling exercises. We chose $k^{opt} = min\{k|R^2(k) > 95\% * R^2(K)\}$; that is, we erred on the side of allowing for more sparsity and less classification error. For stents, this implies about ten times as many algorithmically-identified "brands" than our hand-coding would indicate; for staplers, the algorithm identifies 47 products while hand-coding indicates 53. Across all product category-manufacturer combinations in Table A3, the algorithmic brands explain as much price variation as the hand-coded brands, and mitigates the sparsity problem significantly. For stents (staplers), the number of algorithmic brands is 88 percent (49 percent) smaller than the number of unique SKUs.

The goal of our expanded analyses is to determine whether the patterns documented for stents are similar to those we observe for other important purchase categories. To that end, we employed the same regression specifications as in the stents analyses: (1) We estimate the average treatment effect of joining the database, for which we find no significant result in stents. (2) We estimate the average treatment effect of joining the database among all high-price hospital-brands. (3) We estimate the average treatment effect of joining the database specifically among high-price, high-quantity hospital-brands. For the latter two analyses, we find economically and statistically large effects in stents. We estimate treatment effects of

42

Table A3: Performance of Brand-Assignment Algorithm

| | Stents [$N_m = 4$] | | | | | | | | Staplers [$N_m = 1$] | | | |
| | Mfg 1 | | Mfg 2 | | Mfg 3 | | Mfg 4 | | Surgical | | Surgical Linear | |
	$\lvert J \rvert$	R^2	$\lvert J \rvert$	R^2	$\lvert J \rvert$	R^2	$\lvert J \rvert$	R^2	$\lvert J \rvert$	R^2	$\lvert J \rvert$	R^2
SKU	261	.65	510	.87	47	.6	350	.73	53	.99	40	.97
Hand-coded	5	.66	5	.82	1	.54	2	.69	29	.99	24	.96
Algorithm	47	.62	54	.84	23	.58	11	.69	20	.95	27	.95
$k = 1$	1	.42	1	.42	1	.54	4	.69	8	.84	7	.61
$k = 2$	2	.44	2	.73	2	.54	6	.69	20	.95	15	.7
$k = 3$	5	.6	5	.82	2	.54	8	.69	31	.98	27	.95
$k = 4$	5	.6	8	.82	8	.56	9	.69	40	.99	31	.95
$k = 5$	13	.6	14	.83	12	.56	11	.69	51	.99	38	.97
$k = 6$	29	.61	54	.84	23	.58	23	.69	52	.99	40	.97
$k = 7$	47	.62	112	.85	46	.6	43	.69	53	.99	40	.97
$k = 8$	137	.63	271	.86	47	.6	91	.7	53	.99	40	.97
$k = 9$	261	.65	510	.87	47	.6	129	.71	53	.99	40	.97

join only: due to the lack of precise brand data across all product categories, we do not use brand entry as an additional information shock.

A.3.1 Estimation Sample for Other Important Products

Table A4 shows the effect of each sampling restriction described above on the composition of facilities in our data. The first column summarizes the full set of 2,110 facilities in our data, the second shows the set of facilities purchasing the 52 product categories of interest, the third removes facilities missing conversion factor data or purchasing in non-standard units of measure, and the fourth removes facilities missing "brand" data and/or login data. The fifth column summarizes the final regression sample of facilities for which we were able to flag hospital-brands' positions in the pre-join price and quantity distribution. At each restriction, we tend to remove small facilities and non-hospitals. The final regression sample contains 775 facilities, 73 percent of which were hospitals or health systems, spending $1.8 million per month on 774 product categories.[39]

Figure A5 below compares the distribution of the expanded sample of database members to the general acute care hospitals in the AHA survey data. The patterns presented above

---

[39]One might expect "Restriction 1" to remove only a small number of facilities, as all health care facilities would be expected to purchase commodity products. The removal of a large number of facilities in that cut is due to our reliance on the UMDNS code to identify product categories – if some transactions in a facility's materials management database were recorded in such a way that UMDNS codes could not readily be assigned upon database upload, then it will likely not appear in our analytic sample.
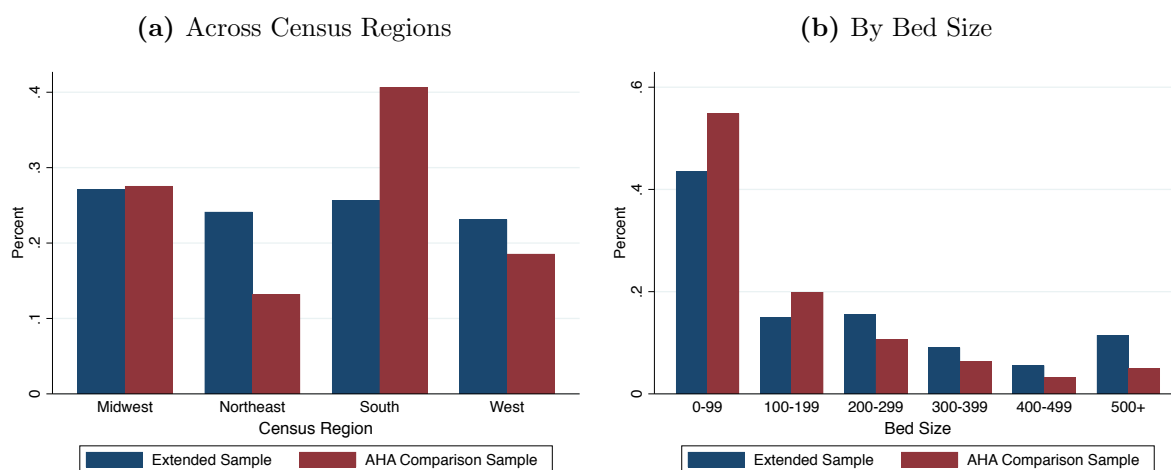
Table A4: Summary Statistics from Purchase Order Database

	Full Sample [N=2,110]		Restriction 1 [N= 1,605]		Restriction 2 [N=1,491]		Restriction 3 [N= 1,388]		Regression Sample [N= 775]	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Months of Data	31.2	21.2	34.5	21.1	35.0	21.2	36.7	20.9	41.4	21.4
Product Categories	462.1	502.7	605.5	496.2	649.3	487.8	670.1	488.3	774.0	487.9
Total Spend/Month ($m)	1.1	2.7	1.5	3.1	1.6	3.1	1.6	2.6	1.8	2.7
% Hospital & Health System	0.48	0.50	0.61	0.49	0.65	0.48	0.67	0.47	0.73	0.44

Authors' calculations from PriceGuide data, 2009-2014.

for stent purchasers are largely present here as well: database members in our sample are underrepresented in the south and somewhat larger than the average US general acute care hospital.

Figure A5: Distribution of Benchmarking Database vs. Comparison Hospitals – All Facilities Purchasing "Important" Products

(a) Across Census Regions **(b)** By Bed Size

Database sample computations from PriceGuide data, 2009-2014. AHA sample computations from American Hospital Association Annual Survey of Hospitals, 2012; includes all general acute care hospitals.

The hospital-brand-month level regression sample data are summarized for each product category in Table A5. The Table also shows aggregates at the product class level.[40] All

---

[40]As discussed in Section 6, Class 1 is "commodities": relatively inexpensive products that are unlikely to be chosen primarily by physicians; e.g., surgical gloves. Class 3 is "physician preference items": high-tech medical devices, mainly coronary and orthopedic products that are the primary implanted device in their corresponding surgical procedures; e.g., coronary stents. Class 2 is intermediate: other medical/surgical products used during invasive procedures, but explicitly excluding PPIs. Note that each row's data is within the set of hospitals with nonzero purchase in that category or class; i.e., average spending on bandages is within 481 hospitals whose bandage transactions we observe, whereas average spending on all commodities will include 267 hospitals whose bandage transactions we do not observe. As noted previously, the fact that we do not observe all facilities purchasing every commodity product is due to our reliance on the UMDNS code to identify product categories – if some transactions in a facility's materials management database were

together, top commodities account for 2 percent of hospital spend, vs. 21 percent for top PPIs, and 13 percent for other medical/surgical products. This is consistent with the large discrepancies in their pricing: the average unit price of a commodity is \$74, vs. \$2,032 for PPIs. Interestingly, we see similar degrees of price dispersion for commodities as we do for other medical/surgical products and PPIs – coefficients of variation across hospitals are 0.195, 0.166, and 0.188, respectively – hinting that physician preference alone cannot account for observed dispersion. Finally, note that the drug-eluting stent sample in Table A5 differs from the sample used in the main analyses in the draft because of the focus on join timing for identification of information effects: hospital-brands such that the brand enters after the hospital joins the database have no pre-information data and thus fall out of the analysis. However, average monthly spend and quantity for this subset of stent-purchasing hospitals is similar to our main stent sample.

B Mapping of Bargaining Setting into Models of Asymmetric Information and Agency

Here, we elaborate on the concise theory sections of the main text.

B.1 Asymmetric Information about Supplier Bargaining Parameters

We follow Rubinstein (1985) to model uncertainty among hospital buyers about the bargaining parameter of a given supplier. The model departs from the complete information model in Rubinstein (1982) in that the supplier is either of weak type with discount factor δ_w^S or strong type with discount factor δ_s^S ($1 > \delta_s^S > \delta_w^S > 0$). The supplier knows his own type, but the buyer has only a subjective prior ω of the probability that the supplier is the weak type.

The equilibrium split of this surplus depends on both the type of the supplier and the prior of the buyer as follows: Rubinstein (1985) shows that there exists a cutoff ω^* such that if the buyer is sufficiently pessimistic about the seller being the weak type ($\omega < \omega^*$), then the buyer simply offers what she would offer the strong type in the complete information game of Rubinstein (1982):

$$p^{AI}(\omega < \omega^*) := c + \delta_s^S \frac{1 - \delta^B}{1 - \delta^B \delta_s^S} V, \tag{2}$$

and both seller types accept this offer. However, if the buyer is more optimistic about the

recorded in such a way that UMDNS codes could not readily be assigned upon database upload, then it will likely not appear in our analytic sample.

Table A5: Summary of important product categories

ID	Name	Class	$\overline{Spend}_{hy}$ %	$\overline{Spend}_{hy}$ $	N_{jht}	N_h	N_j	$\overline{Q}_{hy}$	$\overline{p}_{hjt}$	$CV(p_{h|jt})$
10279	Bandages Elastic	1	0.0	11,237	23,216	481	81	3,176	3	0.226
11315	Dressings	1	0.2	44,105	42,514	581	151	10,492	10	0.238
11883	Gloves Surgical	1	0.6	107,187	51,660	579	49	48,532	2	0.158
12368	Drapes Surgical	1	0.2	47,462	76,910	565	171	7,235	7	0.191
12745	Needles Injection Hypodermic	1	0.1	22,351	66,170	618	61	100,657	0	0.156
13655	Containers Specimen	1	0.2	43,895	28,761	496	88	49,321	3	0.333
923840	Drill Bits	1	0.5	91,421	227,351	573	1,192	610	162	0.166
	Commodities	1	1.9	277,910	516,582	748	1,793	166,520	74	0.195
10688	Catheters Vascular Angiography	2	0.1	38,249	46,764	420	144	982	27	0.126
10894	Clip Appliers	2	0.5	118,993	26,179	499	18	629	118	0.170
11502	Forceps Electrosurgical	2	0.4	108,463	10,689	302	113	137	1,770	0.115
11910	Grafts Bone	2	1.3	377,912	34,324	471	290	136	2,638	0.124
11925	Guide Wires	2	1.1	192,401	189,814	592	859	1,528	82	0.158
12830	Orthopedic Cement	2	0.5	110,897	21,954	466	85	306	341	0.199
13050	Plates Bone	2	1.4	300,071	198,059	570	3,205	411	724	0.158
13909	Sutures Synthetic Nonabsorbable Polypropylene	2	0.2	42,126	98,186	572	300	3,950	15	0.140
13910	Sutures Natural Nonabsorbable Silk	2	0.1	19,211	54,640	502	73	4,532	4	0.262
14085	Tubes Tracheal	2	0.2	34,726	34,323	525	102	4,428	7	0.217
16071	Lenses Intraocular Posterior Chamber	2	0.1	76,046	15,621	268	184	495	161	0.092
16078	Nails Bone	2	0.6	133,181	51,900	482	457	96	1,424	0.154
16104	Wires Bone	2	0.1	22,144	41,367	535	376	214	110	0.189
16655	Catheters Cardiac Ablation	2	0.4	187,729	17,711	256	85	165	1,020	0.127
17184	Catheters Vascular Angioplasty Balloon	2	1.1	276,092	139,972	449	641	1,202	291	0.131
17471	Sutures Synthetic Absorbable Polyglactin	2	0.4	69,078	78,022	605	51	14,241	9	0.587
17600	Bone Matrix Implants	2	0.5	93,743	28,302	475	266	137	751	0.179
17846	Catheters Vascular Guiding	2	0.6	105,814	71,039	430	380	1,169	157	0.177
18609	Catheters Cardiac Mapping/Ablation	2	0.6	208,145	10,742	208	134	173	1,418	0.123
20317	Staplers Surgical	2	0.6	140,430	19,993	400	80	421	355	0.216
20318	Staplers Surgical Linear	2	0.4	110,985	21,776	423	47	635	150	0.245
20453	Stent/Grafts Vascular Aortic	2	1.0	303,258	11,376	290	130	54	5,581	0.050
22538	Suture Anchors	2	0.5	95,811	31,637	526	217	223	319	0.155
22908	Sutures Synthetic Nonabsorbable Nylon Monofilament	2	0.1	15,805	85,860	596	179	3,425	6	0.181
34196	Screws Bone Spinal	2	0.5	296,353	4,265	67	442	545	695	0.198
	Other Med/Surg	2	13.2	1,981,568	1,344,515	701	8,858	29,572	415	0.166
12324	Lenses Intraocular	3	0.1	44,309	11,372	279	121	301	365	0.138
12913	Pacemakers Cardiac Implantable	3	1.7	543,311	29,907	392	126	107	4,316	0.136
15766	Orthopedic Internal Fixation Systems Spinal	3	0.9	338,861	62,248	392	1,256	566	628	0.237
15870	Prostheses Cardiac Valve Biological	3	1.0	463,734	13,737	249	37	301	5,699	0.131
16040	Stents Ureteral	3	0.1	30,037	29,097	480	80	218	137	0.117
16084	Prostheses Joint Hip Acetabular Component	3	0.8	247,393	61,747	457	621	200	1,294	0.242
16095	Prostheses Joint Hip Femoral Component	3	1.3	375,090	97,347	467	998	236	1,568	0.244
16097	Prostheses Joint Knee Femoral Component	3	1.5	405,992	61,277	461	553	197	2,474	0.189
16098	Prostheses Joint Knee Tibial Component	3	1.3	364,878	110,185	469	1,304	307	1,259	0.201
16111	Prostheses Mammary Internal	3	0.6	166,975	37,651	373	223	211	815	0.111
16921	Prostheses Joint Shoulder Humeral Component	3	0.4	118,044	21,559	394	273	52	2,424	0.185
17165	Allografts	3	0.6	157,176	27,301	483	608	94	1,638	0.155
17241	Stimulators Electrical Spinal Cord Analgesic	3	0.8	310,908	5,693	284	19	23	15,249	0.104
18253	Grafts Skin Biological	3	0.4	135,230	6,857	342	65	201	2,916	0.087
18504	Defrib/Pacemakers Implantable	3	2.0	851,354	7,212	280	180	51	16,415	0.125
20376	Defrib/Pacemakers Implantable Resynchronization	3	2.3	1,007,055	5,331	264	109	45	21,199	0.122
20383	Stents Vascular Coronary Drug-Eluting	3	2.9	1,132,582	15,529	336	13	606	1,623	0.080
20422	Stents Vascular Coronary Balloon-Expandable	3	0.3	142,779	46,854	334	200	157	720	0.110
22543	Occluders Vascular Intravascular Embolization Coil	3	0.4	112,271	17,407	304	319	179	769	0.101
923150	Orthopedic Fixation Systems Implantable Spine	3	1.5	476,814	35,243	362	975	148	3,235	0.177
	PPIs	3	20.9	4,314,551	703,554	601	8,080	2,298	2,032	0.188

probability that the seller is the weak type ($\omega \geq \omega^*$), then the buyer offers:

$$p_w^{AI}(\omega \geq \omega^*) := c + \delta_w^S \frac{1 - \delta^{B^2}(1-\omega) - \delta^B \omega}{1 - \delta^{B^2}(1-\omega) - \delta^B \delta_w^S \omega} V, \tag{3}$$

which the weak seller type accepts. The strong seller type will reject this offer, and counteroffer with a price that would make a weak seller no better off than p_w^{AI}, but that the strong seller strictly prefers:

$$p_s^{AI}(\omega \geq \omega^*) := c + \frac{1 - \delta^{B^2}(1-\omega) - \delta^B \omega}{1 - \delta^{B^2}(1-\omega) - \delta^B \delta_w^S \omega} V, \tag{4}$$

which the buyer accepts.

B.1.1 Effects of transparency: complete information assumption

This equilibrium has direct implications for what we would expect to happen to prices in a move from this type of asymmetric information to complete information. First, note that $p_s^{CI} = p^{AI}(\omega < \omega^*) > p_s^{AI}(\omega \geq \omega^*) > p_w^{AI}(\omega \geq \omega^*) > p_w^{CI}$ (where p_s^{CI} is the equilibrium price for the strong supplier type, and p_w^{CI} the weak type, with complete information). Thus the weak type seller is strictly better off with asymmetric information. The strong type seller is weakly worse off (strictly when the buyer's prior is sufficiently optimistic). A sufficiently pessimistic buyer is also weakly worse off with asymmetric information. For more optimistic buyers, whether complete information would make them better off ex-ante depends on parameter values.

In our context, we are interested in when a buyer might benefit from benchmarking information that reveals the seller's type, and what would happen to price in such a case. The intuition for how this unfolds in practice is a scenario where a manufacturer sales representative says "This is the best price I can offer. Corporate won't let me go any lower." Benchmarking information allows the hospital negotiator to perform the due diligence of checking the prices at other hospitals in order to verify or refute this statement.

We begin by assuming that this information *fully* reveals a seller's type.

Prediction 1 (Direct Effect of Complete Information on High Prices) If information is costless, pessimistic buyers will always become informed. This information will cause a proportion of the highest prices p_s^{CI} to fall to p_w^{CI} for those cases where the supplier was in fact the weak type. Thus exposure to benchmarking information should lead to some of the highest prices falling.[41]

---

[41]The Rubinstein (1982, 1985) framework with bilateral negotiations and no externalities across buyers

Prediction 2 (Direct Effect of Complete Information on High Prices with High Quantity)

If information is costly to obtain (in the sense that searching and analyzing the data takes time that could be used on other productive activity), a pessimistic buyer will become informed whenever the expected benefit $\omega(p_s^{CI} - p_w^{CI})q$ exceeds the cost of information. This information will cause a proportion of the highest prices p_s^{CI} to fall to p_w^{CI} for those cases where the supplier was in fact the weak type. Thus exposure to benchmarking information should lead to some of the highest prices falling, among those brands with the highest quantity used.

B.1.2 Effects of transparency: supplier response to information externality across buyers

The above illustrative model and Propositions assumed that benchmarking information fully reveals the seller's type. However, the extent to which the seller's type is revealed by benchmarking depends on how information on *other buyers' prices* alters a given buyer's beliefs. Moreover, it is intuitive to think that the transparency introduced by price benchmarking could have a countervailing effect, in that sellers may be less willing to offer low prices if they know those prices will then be included in the database and potentially hurt them in subsequent negotiations with other buyers (Duggan and Scott Morton 2006; Grennan 2013). To show how this effect might arise, we extend the model to allow for a stochastic information revelation process in a multi-period, multi-lateral bargaining environment.

Consider a focal supplier negotiating with two buyers: the supplier negotiates the price p_1 with buyer 1 (valuing the product at V_1) in the first period, and p_2 with buyer 2 (valuing the product at V_2) in the second period. Buyer 2 has prior belief ω_2^{prior} that the supplier is of the weak type ($\delta^S = \delta_w^S$). If there is transparency in the form of benchmarking, there is an exogenous probability ϵ that p_1 will be revealed to the second buyer.[42] For simplicity

_____

and sellers is helpful in that it motivates a reduced form empirical strategy such as the one employed in our empirical analyses. If we were to impose a model of bargaining under asymmetric information and estimate it structurally, an extension of a multilateral bargaining framework (e.g., Horn and Wolinsky (1988); Collard-Wexler et al. (2017)) would be more appropriate. Given that we explore multiple models via which transparency may affect prices, we do not proceed down this route, though we do propose a simple multilateral bargaining extension with several buyers in Appendix B.1.2 below. Finally, Appendix D.3 explores whether, in a multilateral bargaining setting with multiple imperfect substitute products, a fall in price of substitute brand j will lead to a decrease in the prices of other brands $-j$, all else equal. We did not find any substantial effect of this type, which is consistent with the findings in Grennan (2013) and Grennan and Swanson (2018) of low price sensitivity in demand for stents and other devices.

[42]In a reduced form sense, this model is intended to allow for uncertainty as to whether a seller's type is revealed, via his price, to the second buyer. This uncertainty could manifest in several ways: posterior uncertainty about the first buyer's bargaining parameter, posterior uncertainty about the first buyer's valuation, or other stochastic shocks. For example, a model that allows for posterior uncertainty in buyers' beliefs after receiving benchmarking information – with heterogeneity in buyers' valuations $V_j \sim F[0, \infty)$

of exposition, we suppose that the negotiation between the supplier and buyer 1 is under complete information.[43] The model parameters $(\delta^B, \delta^S_w, \delta^S_s, \omega^{prior}_2, V_1, V_2, \epsilon)$ are known to all parties, but the realization δ^S is known only to the seller and the first buyer.

Define the period-2 price expected by the type-z seller ($z \in \{w, s\}$), given buyer 1 price $p^*_1 \in \{\emptyset, p_1\}$, as

$$
p^z_2(p^*_1) := \begin{cases} c + \delta^S_s \frac{1-\delta^B}{1-\delta^B \delta^S_s} V_2 & \omega^{post}_2(p^*_1) < \omega^*_2 \\ c + \delta^S_w \frac{1-\delta^{B^2}(1-\omega^{post}_2(p^*_1))-\delta^B \omega^{post}_2(p^*_1)}{1-\delta^{B^2}(1-\omega^{post}_2(p^*_1))-\delta^B \delta^S_w \omega^{post}_2(p^*_1)} V_2 & \omega^{post}_2(p^*_1) \geq \omega^*_2, z = w \\ c + \frac{1-\delta^{B^2}(1-\omega^{post}_2(p^*_1))-\delta^B \omega^{post}_2(p^*_1)}{1-\delta^{B^2}(1-\omega^{post}_2(p^*_1))-\delta^B \delta^S_w \omega^{post}_2(p^*_1)} V_2 & \omega^{post}_2(p^*_1) \geq \omega^*_2, z = s. \end{cases}
$$

As before, if buyer 2 is sufficiently pessimistic, the seller receives the strong-type complete-information price; if buyer 2 is optimistic, the seller receives her type-specific asymmetric-information price in a separating equilibrium given ω^{post}_2. Buyer 2's posterior belief may be a function of an observed price ($p^*_1 = p_1$), or may update from ω^{prior} based on the fact that no p_1 was observed ($p^*_1 = \emptyset$). If p_1 perfectly reveals the seller's type, then the seller receives her type-specific complete-information price in period 2.

To illustrate how an externality across buyers can generate distortions in buyer 1's price, we proceed constructively and present a simple numerical example. Let $c = 0$, $V_2 = 100$, $\omega^{prior}_2 = 0.2$, $\delta^B = 0.8$, $\delta^S_s = 0.9$, and $\delta^S_w = 0.7$. In this model, we would observe period-2 pooling under the assumed prior and under any posterior belief ω^{post}_2 below the cutoff $\omega^*_2 = 0.43$. For simplicity, we consider a case here where in all equilibrium outcomes $\omega^{post}_2 < \omega^*_2$, unless the seller's type is fully revealed (we assume this for the moment and verify later).

Observe that in this case, the strong-type seller will always receive the strong-type complete-information price $\delta^S_s \frac{1-\delta^B}{1-\delta^B \delta^S_s} V_2$ in period 2, either because period-1 prices fully reveal types, or because pooling occurs. Thus, there is no externality of the period-2 negotiation on the period-1 negotiation for the strong-type seller.

However, it is costly to the weak-type seller for her price to be revealed in the period-2 negotiation, and this presents a credible threat point for breaking off negotiations with buyer 1. The period-1 negotiation with the weak-type seller is equivalent to a Rubinstein alternating-offers bargaining game with the added feature that the seller has access to a valuable outside option. The weak-type period-1 game then has a unique subgame perfect

---

and Bayesian updating of buyer beliefs – generates similar predictions to our chosen model extension, but with much greater complexity.

[43]This would be the case, for example, if the seller's type had been revealed to buyer 1 via a separating equilibrium in a prior negotiation. Alternatively, one might think of this as a stylized model with V_1 representing a measure of informed buyers with access to benchmarking, V_2 a measure of buyers not subscribing to benchmarking, and ϵ the rate at which the second type of buyers are joining the benchmarking group.

equilibrium with immediate agreement:

$$
p_1^w := \begin{cases} \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} V_1 & \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} V_1 \geq \epsilon \left(\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} - \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} \right) V_2 \\ \epsilon \left(\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} - \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} \right) V_2 & \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} V_1 < \epsilon \left(\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} - \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} \right) V_2 < \delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_1 \\ \delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_1 & \text{otherwise.} \end{cases}
$$

$$(5)$$

This is essentially a modification of the "outside option principle" presented in Osborne and Rubinstein (1990), subsection 3.12.1.[44] The relevant outside option is that the weak seller may walk away from the period-1 negotiation. If she does so, the presence of the ϵ-shock implies that her type cannot be perfectly revealed, though the period-2 buyer will update his belief that she is of the weak type. Specifically, if buyer 2 does not observe a period-1 price, he will infer that either a weak-type seller strategically refrained from negotiating, or the stochastic transparency shock was such that prices were exogenously not revealed. By Bayes' rule, he will update his belief that the seller is a weak type from $\omega_2^{prior} = 0.2$ to $\omega_2^{post} = \frac{\omega_2^{prior}}{1-\epsilon+\epsilon\omega_2^{prior}} = 0.33$. At these parameter values, the equilibrium in period 2 will involve pooling under both ω^{prior} and ω^{post}.[45] Thus, walking away from the period-1 negotiation gives the weak seller the outside option payoff of $\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_2$ across both periods.

In the first case in the above equilibrium, the weak seller's type is fully revealed if p_1 is observed by buyer 2, but the outside option payoff of $\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_2$ (across both periods) if she walks away is nevertheless small relative to her equilibrium payoff across both periods if there were no outside option: $\delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} V_1 + (\epsilon \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} + (1-\epsilon)\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S})V_2$. Opting out is thus not a credible threat for the seller, and prices are as in the complete-information Rubinstein (1982) game.

In the second case, the outside option binds, and the weak seller must be compensated for the negative externality of the period-1 negotiation on the surplus she expects to receive in the period-2 negotiation: $\epsilon \left(\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} - \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} \right) V_2$. The period-1 price increases above Rubinstein (1982) levels to compensate the weak seller for reduced profits in the period-2 negotiation.

In the final case, buyer 1's value V_1 is sufficiently low that the strong seller's price drops below the weak seller's "outside option" price $\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} V_1 < \epsilon \left(\delta_s^S \frac{1-\delta^B}{1-\delta^B \delta_s^S} - \delta_w^S \frac{1-\delta^B}{1-\delta^B \delta_w^S} \right) V_2$. When that occurs, the weak seller prefers to accept the lower price, equal to the strong type

---

[44]Osborne and Rubinstein (1990) present a model in which the seller has the option of terminating the negotiation and receiving her outside option after she has rejected an offer. In this case, the game has a unique subgame perfect equilibrium as characterized above. This model was variously suggested by Binmore, Shaked, and Sutton; see Shaked and Sutton (1984); Binmore (1985); Binmore et al. (1989).

[45]This confirms that the strong-type seller has no incentive to deviate from the complete-information Rubinstein (1982) price in period 1, and in turn that $\omega_2^{post} = 0.33$ is consistent with Bayesian updating.

50

seller's price, as it allows her to pool with the strong seller if price information is revealed in period 2, and she receives a strong-type seller price in both periods. Buyer 1 also prefers to pay the lower price, supporting this equilibrium.

In Figure A6 below, we present equilibrium prices for the strong- and weak-type seller in the absence of the externality (i.e., if period-1 prices are never observed by period-2 buyers: $(p_1^{s,NoExt}, p_1^{w,NoExt})$) and with the information externality ($\epsilon = 0.5 : (p_1^{s,InfoExt}, p_1^{w,InfoExt})$). All model parameters are held fixed except V_1; the Figure shows equilibrium prices for each $V_1 \in [0, 100]$.

Figure A6: Asymmetric Information with Externality – Numerical Example

Absent the externality, strong sellers' prices are uniformly above weak sellers' prices at each value of V_1, and each increases less than one-for-one with buyer 1's valuations as a function of their complete-information bargaining splits. For the strong seller, the $NoExt$ and $InfoExt$ prices are equivalent ($p_1^{s,InfoExt} = p_1^{s,NoExt}$), as benchmarking does not impact the strong seller's prices. However, for the weak seller, the outside option binds for $V_1 \leq 51$, and prices in period 1 are higher with the information externality than without. For $V_1 \in [26, 51]$, $p_1^{s,NoExt} < p_1^{s,InfoExt} = 16.23$. And for $V_1 < 26$, the strong and weak type prices will be equivalent: $p_1^{w,InfoExt} = p_1^{s,InfoExt}$. That is, for sufficiently low V_1, the weak seller will accept a price below 16.23, as it increases her period-2 payoff by allowing her to pool with the strong seller.[46]

---

[46]There is a discontinuous drop in the weak seller's equilibrium payoff across both periods as buyer 1's value crosses 26. This is because buyer 2's knowledge of V_1 makes $p_1^{s,InfoExt}$ the unique price that allows pooling when period-1 prices are revealed. If the strategy space were expanded to allow hidden lump-sum transfers between the period-1 negotiating parties, this would break the informativeness of period-1 prices, and the weak seller would prefer to give buyer 1 hidden lump sum rebates in exchange for higher unit prices at higher values of V_1.

In this stylized model, the presence of the buyer-2 negotiation effectively puts a lower bound on p_1: whenever the standard Rubinstein (1982) allocation would make the seller prefer to break off negotiations with buyer 1, p_1 must at least be large enough to make the weak seller indifferent. Thus – via this externality created by the benchmarking information – some of a seller's lowest prices may increase. The likelihood and magnitude of this externality will depend on parameters, and in particular on the relative magnitudes of V_1 and V_2.[47]

The addition of an externality across negotiations in this extended model implies potential "convergence" of prices from both the top and the bottom – the highest prices will tend to decrease as information is revealed (the buyer 2 effect here), and the lowest prices will tend to rise as suppliers take into account their externalities on other buyers (the buyer 1 effect here).

Prediction 3 (Effect of Information Externality Across Buyers on Low Prices) Via the benchmarking process, a negotiated price with one buyer becomes information available to other buyers, potentially introducing an externality in future negotiations. If this externality is large enough, it can provide the seller a credible threat not to accept low prices. Thus some prices may rise with the introduction of benchmarking information, particularly among buyers with relatively low prices and a relatively low quantity under negotiation.

B.2   Negotiator Agency

The other candidate mechanism via which we propose benchmarking information could be valuable to buyers would be through providing aggregate information to help the buying firm solve a moral hazard problem with its purchasing agent. Here we provide a specific model of information in the bargaining context that generates predictions in our setting. Modifying Holmstrom (1982) to our context, let price p_h at hospital h be as in the full information Rubinstein (1982) bargaining game. However, instead of the hospital negotiator's bargaining parameter being exogenous, the price will be a function of the hospital agent's *choice* of discount factor δ_h^B and the discount factor of the supplier, which takes value $\delta_w^S \epsilon_h$ with probability ω and $\delta_s^S \epsilon_h$ with probability $1 - \omega$. As before, the discount factor of the strong supplier type is greater than that of the weak type ($1 > \delta_s^S > \delta_w^S > 0$). ϵ_h is a random term distributed uniform on $[0, 1]$. Importantly, the realization of ϵ_h is independent across hospital buyers, but whether the seller is weak or strong is common to all buyers. The realizations of both of these random variables are observable to the negotiating agents, but not to the principals who manage them.

---

[47]Note that in our empirical context, the payoff to the seller in a given negotiation scales with the buyer's willingness-to-pay, as well as the quantity under negotiation.

A moral hazard problem arises in this setting because bargaining effort is costly and provides the agent disutility $v(\delta_h^B)$. The agent is compensated by some contract based on the price $m(p_h)$. The agent is risk averse in money, so the optimal solution to the agency problem involves risk sharing between the principal and the agent. Holmstrom (1982) shows that if agents face some common parameter which is uncertain from the principals' perspectives, then relative performance evaluation compared to some aggregate sufficient statistic can be used to write a better contract with each agent. The intuition in our real-world setting is one where with the benchmarking data, hospital administrators can make their negotiators' performance reviews contingent on the prices they negotiate relative to other hospitals for the same brand. This motivates the following Predictions:

Prediction 4 (Monitoring Effect on Prices) If buyer negotiators are imperfect agents of the buying firm, then benchmarking information (observing the distribution of price realizations across hospitals $\{p_h\}_{h=1}^H$) allows the principal to estimate whether the seller is the weak or strong type, and thus reduce the risk to which the agent is exposed and write a contract which induces more bargaining effort and a lower price than in the case where only p_h is observed.[48]

Prediction 5 (Monitoring Effect on Prices with High Quantity) If buyer negotiators are imperfect agents of the buying firm, but it is costly for hospital managers to search and analyze the data in a way that allows them to write better contracts, then managers will use benchmarking information (observing the distribution of price realizations across hospitals $\{p_h\}_{h=1}^H$) to write a contract which induces more bargaining effort by the agent and a lower price than in the case where only p_h is observed if $(p_h(m) - p_h(m(\{p_h\}_{h=1}^H)))q_h$ exceeds the cost of information use.

C   Details on the Contracting Environment

In this Appendix, we provide some additional background on how medical devices are used and purchased, and we provide additional evidence regarding the validity of our assumptions and potential mechanisms.

---

[48]The model as written has a strong prediction that this effect will be independent of price. However, in general the prediction of how the price distribution would move with information depends on where in the model the current heterogeneity is coming from. For example, if the heterogeneity were due to different levels of risk aversion among negotiators, then benchmarking information would tend to decrease the highest prices more than the lowest.

C.1 Hospital-Supplier Contracting

Hospitals purchase thousands of product categories, and prices for each of these product categories are determined in negotiation. Purchasers within a hospital must factor in clinical value, safety, cost, and other conditions of sale and service (e.g., for capital equipment, maintenance is an important consideration). Negotiation can take place directly between a hospital administrator and a representative of the brand's manufacturer, or hospitals may rely on group purchasing organizations or other contracting coalitions to negotiate their contracts. GPO prices are typically used as a starting point for direct hospital-manufacturer negotiations for some products, particularly physician preference items and capital equipment (Schneller 2009).

Our conversations with industry participants revealed substantial heterogeneity across organizations in the purchasing practices via which supply contracts are negotiated. Some hospitals have large materials management or purchasing departments with agents who specialize in negotiations. Sometimes a large business unit, such as a catheter lab in the case of stents, will coordinate its own purchasing separately from the rest of the hospital. Even absent access to benchmarking information, hospitals vary in access to information on the prices other hospitals pay via GPOs, hospital system membership, or informal networks of peers. While GPO membership is unknown, many of the facilities in our data are members of health systems at some point. The extent to which hospital systems bargain collectively over supplies will vary by context. Hospitals often argue that the ability to bargain collectively is a substantial advantage over and above that conferred by GPO membership (Noether and May 2017). However, we find in other work that hospital mergers lead to limited savings relative to the dispersion we document for merging parties (Craig et al. 2018); we also observe substantial price dispersion across hospitals in the same system in these data. For example, the Gini coefficient for the average physician preference item is 0.058 across all sample facilities that are members of health systems, whereas the average within-health system Gini coefficient is actually slightly higher at 0.061.[49] The same comparison is 0.103 vs. 0.104 for commodities, and 0.077 vs. 0.076 for other medical/surgical products. Thus, we observe little evidence of collective bargaining by hospitals under the same ownership in these data.

There is also significant variation in discounts obtained across product categories within a given facility. To investigate this, we have run some regressions analogous to those in Figure 1(a). We stack all product categories and run a single regression of $ln(price)$ on brand-time

---

[49]Facilities in the same system are identified as those under the same PriceGuide parent account. To aid in comparison, we estimate Gini coefficients across all facilities, and average Gini coefficients across facilities in the same system, using the same sample of facilities under multi-member parents. The Gini coefficient is one half of the mean absolute difference between any randomly selected pair of hospitals.

fixed effects, and several different versions of hospital fixed effects: hospital fixed effects, hospital-product class fixed effects, and hospital-product category fixed effects. First, we document that hospitals' performance in bargaining is positively, but imperfectly, correlated across product classes: the correlation between a given hospital's average discount on PPIs and that same hospital's average discount on commodities is only 0.228; the same statistic for PPIs and other medical/surgical products is 0.229. There is a stronger correlation (0.436) between hospital fixed effects for commodities and those for other medical/surgical products. Second, and relatedly, we find that a significant component of hospitals' performance in bargaining is idiosyncratic across categories. Conditional on fixed effects for each product category-brand-time, hospital fixed effects explain only 4.3 percent of the residual price variation in the stacked regression, hospital-product class fixed effects explain 7.2 percent of the residual price variation, and hospital-product category fixed effects explain 30.4 percent of the residual price variation.[50] This finding is consistent with our understanding that purchasing is somewhat siloed across departments within a hospital.

C.1.1 Stent Usage and Purchasing

The main product category we analyze is coronary stents; coronary stents are small metal tubes placed into narrowed coronary arteries to widen them and allow blood flow to the heart.[51] In the US, hospitals spend more than two billion dollars annually on stents used in over 700,000 procedures.[52]

For stents, as for other physician preference technologies, usage is driven by physicians choosing which brand to use to treat a given patient, while prices are determined in negotiation between a hospital administrator and a representative of the brand's manufacturer. There is no "search" in the conventional sense, as a given brand can only be purchased directly from its manufacturer. The manufacturer holds inventory on-site at the hospital, and the purchase is made when the physician pulls the product off the shelf and implants it into the patient.

---

[50]More precisely, the R^2 of the baseline pooled regression with only category-brand-time fixed effects is 0.9931. Adding a hospital fixed effect increases the R^2, but only to 0.9934 $((0.9934 - 0.9931) * 100/(1 - 0.9931) = 4.3$ percent; this indicates that some hospitals have systematically better prices, holding both time and product portfolio fixed. Adding hospital-class fixed effects increases the R^2 to 0.9936. Finally, adding hospital-product category fixed effects increases the R^2 to 0.9952.

[51]The original technology, the bare metal stent (BMS), was approved in the early 1990s; in the early 2000s, the drug-eluting stent (DES) was introduced as an improvement over the older technology with lower risk of restenosis, a condition that may arise when scar tissue builds up around the stent and restricts blood flow yet again.

[52]700,000 estimate from Waldman et al. (2013), referencing stent procedures in Medicare enrollee population. Two billion dollar figure based on authors' calculations using Boston Scientific's reported US revenue in 2012 (BSX 10-K 2012) and Boston Scientific's 2012 market share in purchase order data.

Stent contracts typically specify a linear price for the contract duration, often a year. In the short run, hospitals are reimbursed a fixed amount by private or public insurers based on the services they provide, and so device prices come directly from the hospital's bottom line.

C.2   Compensation of Purchasing Agents

In this Appendix, we use survey and interview data to explore the agency hypothesis. First, we have obtained access to management practices survey data for a large sample of hospital cardiac units from McConnell et al. (2013). Second, we have conducted interviews with three different types of hospitals regarding their purchasing practices, contracting with purchasers, and the interaction of these with using benchmarking information.

McConnell et al. (2013) developed a survey instrument, adapted from "lean manufacturing" surveys, to measure management practices in the cardiac inpatient setting. The survey included open-ended questions on standardization, performance monitoring, setting targets, and incentivizing employees and managers. Many of these questions shed light on how staff in cardiac units are tracked and rewarded. Those that are most relevant for the current setting are the following (in each case, we summarize the "best case scenario"):

- **Technology Adoption** Score 5: "There is a systematic approach in the adoption of all new technologies, including a review of best evidence on costs and effectiveness; unit may participate in learning collaborative to accelerate information gathering."

- **Performance Review** "Performance is continually reviewed, based on the indicators tracked; all aspects are followed up to ensure continuous improvement; results are communicated to all staff."

- **Target Stretch**: "Goals are genuinely demanding for all parts of the organization and developed in consultation with senior staff (eg, to adjust external benchmarks appropriately)."

- **Rewarding High Performers**: "There is an evaluation system for the awarding of performance related rewards, including personal financial rewards and shared group/team rewards."

- **Removing Poor Performers**: "We work hard to identify weaknesses and improve or remove poor performers."

Each survey answer was scored on a scale of 1 to 5, with 5 corresponding to the "best" answer as described above. Cardiac units were heterogeneous in their responses to these

questions, suggesting that the ex ante potential for our "moral hazard" mechanism would be limited to a small number of hospitals in our data. The raw scores are summarized in Table A6. First, they suggest that only 33 percent of units received a 4 or 5 on the technology adoption question – while not touching on incentives explicitly, this question characterizes the extent to which cost-effective input use is an *objective* of the unit. Second, more than half of units received a 4 or 5 on the performance review question. Together, these findings imply that only a small subset of cardiac units might both prioritize cost-effective purchasing and review performance on that metric. Finally, the percent of respondents scoring a 4 or 5 on the target stretch (i.e., external benchmarks), rewarding high performers, and removing poor performers questions ranged from 20-27 percent. In our view, this paints a relatively pessimistic picture of the potential role of performance pay based on benchmarking to mitigate moral hazard in our setting.

Table A6: Proportion of answers to five practice-related questions in %

	1-worst	2	3	4	5-best
Technology adoption	8	15	44	26	8
Performance review	1	12	33	40	14
Target stretch	8	28	37	23	4
Rewarding high performers	20	19	40	16	4
Removing poor performers	9	30	36	21	3

Raw interview score tabulations shared by K. J. McConnell; summary data reported in McConnell et al. (2013).

We also interviewed several representatives of hospital purchasing departments. Each interviewee worked within the central purchasing unit of a hospital or hospital system and was asked about use of external benchmarking data in employee compensation. One interviewee worked within a large regional hospital system, one within a large urban academic medical center, and one within a small community hospital. Our interview with the representative of the large regional hospital chain indicated that the central purchasing manager regularly makes individual business units aware of opportunities for savings based on external benchmarking, and that annual performance evaluation involves discussion of the prior year's accomplishments, including savings on medical supplies. Staff are also eligible to receive small bonuses (around $300) for particularly strong performance, akin to employee-of-the-month recognition. Our interview with the representative of the academic medical center described performance pay for purchasing staff in a similarly holistic way: each contract manager is evaluated periodically based on general performance on his/her overall portfolio of contracts, with dimensions including timely execution of contracts, management of con-

tracts, and "savings they drive." In each of these interviews with representatives of large purchasers, we learned that benchmarking data provide information on savings achieved by contract managers and that annual performance is evaluated in part based on savings, but there is no formal or specific link between compensation and performance relative to metrics from benchmarking resources. The interview with the small community hospital indicated that pricing information was primarily obtained from the hospital's GPO and a small regional purchasing collaborative; that hospital was in the process of transitioning to a subscription benchmarking service, but did not mention employee performance evaluation as a goal of that subscription.

In sum, our understanding is that benchmarking can be part of a purchasing agent's overall performance evaluation, consistent with the theoretical benefit of benchmarking to resolve moral hazard problems. However, both large-sample surveys of cardiac units and our in-depth interviews of purchasing units indicate that reliance on benchmarking to evaluate purchasing employees is heterogeneous and informal, consistent with our finding that asymmetric information can account for the overall effect of transparency on price realizations.

C.3 Evidence Regarding Linear Price Assumption

It is important for our analysis that the prices we observe are comparable across observations in the sense that there are not important contract dimensions that we do not observe (e.g., bundling, exclusivity, or market share based contracting). Our conversations with industry participants indicate that stents tend to have simple linear price contracts, so we assume that transactions data capture real prices. Here, we examine these assumptions in the data.

C.3.1 Extent of "exclusivity" and correlations with prices

In panels (a) and (c) of Figure A7, we show histograms of total unique manufacturers and stent products (brands) purchased over each quarter by each hospital in the sample. The vast majority of hospitals purchase multiple brands from multiple manufacturers, rather than purchasing a single most-preferred brand for the whole facility. Panels (b) and (d) show these histograms for only hospitals above the 25th percentile in total stent volume, and show even less evidence of "exclusivity" – fewer than 3 percent of hospital-quarters involve a single brand and fewer than 7 percent involve a single manufacturer. The fact that the majority of the already small amount of observed "exclusivity" occurs at hospitals with lower utilization is consistent with the anecdotal evidence that exclusivity does not play a systematic role in stent contracting – with true linear price contracts, "exclusive" purchasing patterns are more likely to be observed among hospitals with low utilization due to random variation and

Figure A7: Histograms – Number of Unique Brands/Manufacturers per Hospital-Quarter

(a) Brands

(b) Brands (excluding smallest hospitals)

(c) Manufacturers

(d) Manufacturers (excluding smallest hospitals)

Regressions of price on indicators for exclusivity:		
Specification	β^{Excl}	(s.e.)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \epsilon_{hjt}$	$25.3^{\dagger}$	(9.0)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \theta_h + \epsilon_{hjt}$	13.2	(9.5)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \theta_{hj} + \epsilon_{hjt}$	-2.9	(8.0)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \epsilon_{hjt}$	8.6	(9.4)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \theta_h + \epsilon_{hjt}$	-15.9^{**}	(6.5)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \theta_{hj} + \epsilon_{hjt}$	-4.2	(6.8)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,223$. Standard errors clustered at hospital level, $N_h = 507$ in first two specifications and hospital-brand level $N_{hj} = 2,227$ in brand-hospital fixed effects specification. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

59

costly contracting.

As a further check, we look at the pricing consequences of the observed sole sourcing in the usage data. For the minority of hospitals that do happen to use only a single brand or manufacturer in a given quarter, we create indicators for $\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}}$ and $\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}}$ and regress price on each indicator and brand-month fixed effects θ_{jt}. The only negative and statistically significant result is a \$15.90 lower price associated with manufacturer sole sourcing when hospital fixed effects are included. This is consistent with a small price savings from "standardization," holding the hospital fixed. However, the result goes away when including hospital-brand fixed effects, so the \$16 result appears to be either spurious or driven by composition effects. We conclude that the small amount of sole sourcing observed is most likely due to other factors besides contracting concerns.

Finally, we check for any evidence of near-exclusivity in the form of market share based contracts (which we are told are commonly used for many medical products, but not stents). Figure A8 plots the cumulative density of observations by brand market share at the hospital-quarter level. We do not observe the bunching that we would expect if contracts commonly specified market share thresholds in either the full sample (panel (a)) or restricting to the most used brand at each hospital (panel (b)). In the table below the Figure, we also rerun the "exclusivity" price regressions using cases where market shares exceed 90 percent, i.e. $\mathbf{1}_{\{s_{jht}>0.90\}}$, and find no economically or statistically meaningful relationship with prices.

C.3.2 Standard offers and multipliers/bundling

Many real-world bargaining settings have "list prices" as a starting point for negotiations. Sometimes these prices are in practice paid by almost no one, and so less relevant to the bargaining problem. Sometimes they are paid by a large mass of smaller customers. They may also act as a starting point for a negotiation of a single "multiplier" that would modify the list prices across a variety of products sold by the supplier to the buyer. Our understanding from talking to both buyers and suppliers in the coronary stent market is that these practices are not prevalent. Here we examine the data to verify this understanding and better understand the nature of price setting.

Figure A9 plots the price distributions across hospitals for each of four DES sold to at least 100 hospitals in January 2012. While the contracted prices do aggregate at the "round" \$50 increments, they do so in a way that appears to approximate single-peaked distributions that are smooth in the sense of no mass suggesting an obvious focal price standing out from the rest. In particular, they do not show evidence of a list price at the top paid by a large subset of customers. These figures are representative of the price distributions for DES across hospitals in other time periods.

Figure A8: Cumulative distributions by market share

(a) All   **(b)** Most Used Brands Only

Regressions of price on indicators for exclusivity (share-based contracts):		
Specification	β^{Excl}	(s.e.)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \epsilon_{hjt}$	12.2	(11.5)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \theta_h + \epsilon_{hjt}$	-9.4	(10.7)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{J}_{ht}|=1\}} + \theta_{jt} + \theta_{hj} + \epsilon_{hjt}$	-2.9	(8.0)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \epsilon_{hjt}$	18.3	(13.5)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \theta_h + \epsilon_{hjt}$	-7.5	(9.9)
$p_{hjt} = \beta^{Excl}\mathbf{1}_{\{|\mathcal{M}_{ht}|=1\}} + \theta_{jt} + \theta_{hj} + \epsilon_{hjt}$	-4.2	(6.8)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,223$. Standard errors clustered at hospital level, $N_h = 507$ in first two specifications and hospital-brand level $N_{hj} = 2,227$ in brand-hospital fixed effects specification. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A9: Histograms of prices across hospitals for each brand in January 2012

Table A7: Price Co-movement at the Manufacturer-Hospital Level

$1_{\{\Delta p \neq 0\}}$ (DES)	(1)	(2)	(3)
$1_{\{\Delta p \neq 0\}}$ (BMS)	0.0621		
	(0.0556)		
$1_{\{\Delta p \neq 0\}}$ (Guiding Cath)		0.0180	
		(0.0383)	
$1_{\{\Delta p \neq 0\}}$ (Guide Wires)			0.0646
			(0.0828)
R^2	0.005	0.000	0.003
Δp (DES)	(1)	(2)	(3)
Δp (BMS)	0.133		
	(0.119)		
Δp (Guiding Cath)		0.0169	
		(0.0254)	
Δp (Guide Wires)			4.133
			(4.365)
R^2	0.011	0.000	0.008

$N = 973$ manufacturer-hospital-month observations. Standard errors in parentheses, clustered by hospital. $N_h = 85$. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Table A7 explores the extent to which prices move together for products sold by the same manufacturer to the same hospital. We flag price changes (using the same algorithm for identifying price changes for our renegotiation analysis) for DES, BMS, Guiding Catheters, and Guide Wires – all of these product categories are used together in interventional cardiology procedures and sold by the same manufacturers, and potentially even sold via the same sales force.[53] We then merge the four data sets at the manufacturer-hospital-month level, keeping the most purchased brand for each manufacturer-hospital pair to ensure at most one observation in a category. Before any further analysis, we note that only 85 hospitals in the sample purchase their modal brands from the same manufacturer across these categories – this fact alone reveals the limited extent to which pricing that references other categories could be feasible, given observed usage patterns. The top panel of the table then regresses the flag for DES price change on the flag for a price change in each of the three other categories (e.g. $1_{\{\Delta p_{mht}^{des} \neq 0\}} = \beta_0 + \beta_1 1_{\{\Delta p_{mht}^{bms} \neq 0\}} + \epsilon$). The interpretation of the coefficient β_1 is then the percentage of hospital-manufacturer-months in which price changes for the other cardiology device category are accompanied by price changes in DES, and R^2 is then the

---

[53]Guiding Catheters and Guide Wires are used in catheter-based interventional procedures outside of the coronary and even vascular settings, and so used by physicians outside of interventional cardiology. Depending on the hospital and manufacturer, these could be sold via separate channels.

percentage of price co-movement overall – in the case of perfect co-movement, $\beta_1 = 1$ and $R^2 = 1$. The results indicate that DES prices change 6 percent of the times BMS prices change, 2 percent of the times catheter prices change, and 6 percent of the times guide wire prices change (though none of these point estimates are statistically different from zero, all are statistically different from 1). R^2 are also very small, indicating DES prices often change in months when other category prices do not.

The bottom panel of Table A7 shows the output of similar regressions but using the actual magnitudes of price changes (e.g. $\Delta p_{mht}^{des} = \beta_0 + \beta_1 \Delta p_{mht}^{bms} + \epsilon$) to understand the direction and magnitude of co-movements. Here there are again no statistically significant relationships. Our interpretation of these results is that both the degree and magnitude of price co-movements are small. This seems inconsistent with a model of a multiplier on list prices negotiated across many product categories, where we would expect large degrees of positive correlation in the incidence and relative magnitude of price changes. It is also inconsistent with most models of bundled pricing, which would tend to predict positive coincidence of price changes and negative correlation in price change magnitudes.

In unreported analyses, available upon request, we conduct similar analyses for the three other groups of product categories meeting the criterion that the same vendors account for at least 70 percent of hospital spend: pacemakers and defibrillators; knee and hip prosthesis components; and several types of sutures. The patterns are similar to those shown here for stents, catheters, and guide wires. Within manufacturer-hospital pair, across categories: (1) the overall frequency of price co-movements is small; and (2) the correlations in price change magnitudes are also small, and positive, where they exist.

D   Stent Analyses – Detail and Robustness

D.1   Stent Analyses in Text – Detail and Alternative Controls

The differences in differences estimate of the average treatment effect of access to information is small, statistically and economically, at -\$3 (s.e. 3). Here, we also use an event study specification that includes indicators for each month relative to the hospital's "info" date:

$$P_{hjt} = \sum_{mo=-12}^{+12} \beta^{Info,mo} * \mathbb{1}_{\{mo=t-t_{info_{hj}}\}} + \theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j}) + \varepsilon_{hjt}$$

Figure A10 shows results for these estimated differences between treated and untreated prices. The plot shows evidence of a slight decline in prices prior to accessing information,

though the pre-trends in price in the six months leading up to the timing of information are essentially zero.[54] After the hospital accesses the benchmarking information, there is a steady downward trend in the coefficients. The downward trend in the post-period may be due to price stickiness – it may take newly-informed hospitals some time to arrive at the bargaining table.

In general, estimates for each relative month effect are insignificant and there is not strong evidence of a trend break. Moreover, estimated patterns are similar across the different specifications of controls, though standard errors are large in the richest specification (Version 3).[55]

Figures A11 and A12 show the same information as in Figure 3 in the main text, but allowing for alternative specifications of fixed effects. As with the overall average treatment effects, the estimates are significantly smaller when we control for hospital-by-brand (rather than hospital *and* brand) fixed effects. However, we see the same qualitative pattern regardless of controls: the treatment effects are statistically zero in all but the top quintile of the pre-information price distribution, the previously high-priced hospital-brands achieve significant price decreases in the presence of benchmarking information, and these price decreases manifest somewhat slowly after the initial information shock.

The following three Figures show how the specifications in Table 2, each of which explores the mechanisms underlying the core results for drug-eluting stents, vary with the included controls. Figure A13 shows how the distribution of treatment effects as a function of pre-information price vary by whether the hospital-brand involved high or low purchase quantities pre-information. Figure A14 explores whether our treatment effects can be attributed to the agency vs. asymmetric information hypotheses. Finally, Figure A15 shows the effects of information on renegotiation per se, as well as the effects of information on price decreases conditional on renegotiation taking place.

---

[54]It should be noted that there are fewer "pre-info" observations available 6-12 months prior to accessing information because of the presence of entering brands and because some hospitals do not submit retrospective data until a few months after joining the database. Accordingly, the earlier relative month effects are less precisely estimated.

[55]It was not possible to estimate the monthly event studies with hospital-brand and brand-month fixed effects. However, the quarterly event study with hospital-brand and brand-month fixed effects is essentially identical to the quarterly event study with hospital-brand and brand-specific linear trends.

Version	Month relative to info date ($\beta^{Info,mo}$)																							
	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12
1	17	15	20*	20*	17*	3	-1	-5	3	0	-1	3	-8†	-10†	-14†	-16†	-20†	-21†	-22†	-24†	-22†	-20†	-22†	-21†
	(14)	(14)	(11)	(11)	(9)	(7)	(6)	(5)	(5)	(4)	(3)	(3)	(3)	(3)	(4)	(4)	(5)	(5)	(6)	(6)	(7)	(7)	(7)	(8)
2	22	19	19*	22**	16*	7	5	1	9	8	4	6	-13†	-17†	-18†	-21†	-27†	-27†	-25†	-24†	-24†	-21**	-23†	-18*
	(14)	(14)	(11)	(10)	(9)	(8)	(8)	(7)	(6)	(6)	(5)	(5)	(4)	(4)	(5)	(6)	(7)	(7)	(7)	(8)	(9)	(8)	(9)	(9)
3	16	14	22	19	14	7	7	3	5	2	1	0	-4	-6	-10	-11	-15	-17	-19	-23	-19	-22	-23	-22
	(35)	(32)	(28)	(26)	(23)	(20)	(17)	(14)	(11)	(9)	(6)	(3)	(3)	(6)	(9)	(12)	(15)	(18)	(19)	(22)	(25)	(28)	(30)	(33)

Authors' calculations from PriceGuide data, 2009-2014. $N = 23,016$ member-brand-months. Includes 507 members, twelve months pre- and post-join. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Version 3) level shown in parentheses. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A10: Event Studies of Treatment Effect of Access to Benchmarking Information

Version	Pre-info price quintiles ($\beta_{quintile}^{Info}=$)				
	1	2	3	4	5
1	4	17†	9	-2	$-55^{\dagger}$
	(6)	(6)	(6)	(6)	(9)
2	-1	11	4	-7	$-63^{\dagger}$
	(7)	(7)	(7)	(8)	(10)
3	-7	6	6	1	$-27^{\dagger}$
	(5)	(5)	(5)	(5)	(7)
4	-10	1	2	-3	$-34^{\dagger}$
	(6)	(6)	(6)	(6)	(8)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,453$ member-brand-months. Includes 508 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A11: Treatment Effect Estimates Throughout the Price Distribution

67

Ver-sion	Month relative to join date ($\beta_5^{Join,mo}$)																							
	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12
1	-24	-33	-20	-24	-14	-14	-21*	-31†	-27†	-22†	-14**	-8	-23†	-40†	-52†	-61†	-64†	-74†	-77†	-88†	-83†	-78†	-104†	-104†
	(20)	(25)	(22)	(19)	(16)	(17)	(12)	(10)	(8)	(7)	(7)	(6)	(7)	(8)	(9)	(9)	(11)	(12)	(10)	(13)	(13)	(11)	(13)	(13)
2	-17	-15	-10	-11	-6	-8	-14	-29†	-22**	-15*	-9	-6	-26†	-43†	-54†	-65†	-70†	-80†	-80†	-89†	-83†	-75†	-104†	-102†
	(19)	(24)	(22)	(18)	(15)	(17)	(13)	(10)	(9)	(8)	(7)	(7)	(7)	(9)	(9)	(10)	(11)	(13)	(11)	(13)	(14)	(12)	(13)	(13)
3	-18	-36	-27	-40	-34	-26	-8	-4	-6	-4	-5	-4	-16†	-28†	-43†	-50†	-56†	-62†	-65†	-82†	-76†	-75**	-93†	-96†
	(36)	(34)	(30)	(27)	(24)	(22)	(18)	(16)	(12)	(10)	(6)	(4)	(5)	(9)	(11)	(14)	(17)	(20)	(21)	(24)	(27)	(29)	(32)	(35)

Authors' calculations from PriceGuide data, 2009–2014. $N = 23{,}016$ member-brand-months. Includes 507 members, twelve months pre- and post-join. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Version 3) level shown in parentheses. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A12: Event Studies of Treatment Effect of Access to Benchmarking Information, Top Quintile of Price Only

(a) Low Quantity **(b)** High Quantity

Version	Pre-info price quintiles ($\beta^{Info}_{quintile,low^q}$)					Pre-info price quintiles ($\beta^{Info}_{quintile,low^q} + \beta^{Info}_{quintile,high^q}$)				
	1	2	3	4	5	1	2	3	4	5
1	5	19†	16**	1	−51†	3	13	−4	−10	−73†
	(7)	(6)	(7)	(7)	(9)	(10)	(9)	(9)	(10)	(15)
2	−1	12	11	−4	−60†	0	7	−12	−14	−79†
	(8)	(8)	(8)	(8)	(11)	(11)	(10)	(10)	(11)	(16)
3	−4	9	9	5	−17**	−11	0	0	−9	−71†
	(6)	(6)	(6)	(6)	(7)	(9)	(8)	(7)	(8)	(13)
4	−9	4	6	1	−23†	−12	−4	−5	−12	−78†
	(7)	(7)	(7)	(7)	(8)	(9)	(8)	(8)	(9)	(13)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,453$ member-brand-months. Includes 508 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.
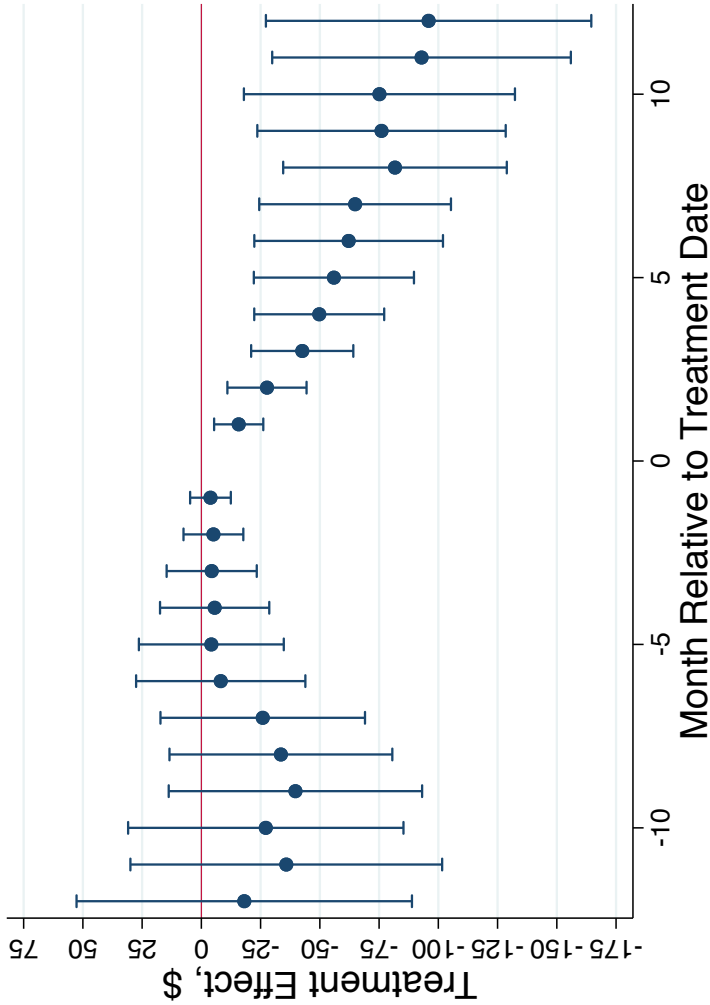
Figure A13: Treatment Effect Estimates Across the Price and Quantity Distributions

(a) Agency

(b) Asymmetric Info

Version	Pre-info price quintiles ($\beta_{quintile}^{Agency}$)					Pre-info price quintiles ($\beta_{quintile}^{Info}$)				
	1	2	3	4	5	1	2	3	4	5
1	8	14	−9	−21*	−60†	3	12**	15†	8	−33†
	(12)	(11)	(10)	(12)	(21)	(6)	(5)	(5)	(7)	(7)
2	19	31**	1	−11	−53**	−8	−2	4	−2	−46†
	(13)	(12)	(11)	(14)	(22)	(9)	(8)	(8)	(10)	(10)
3	−17	−3	2	7	13	−1	7	5	−1	−30†
	(11)	(12)	(10)	(12)	(18)	(6)	(5)	(5)	(5)	(7)
4	−3	16	16	18	25	−15*	−11	−11	−15**	−47†
	(12)	(12)	(11)	(12)	(19)	(8)	(7)	(7)	(7)	(9)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,453$ member-brand-months. Includes 508 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A14: Treatment Effect Estimates Across the Price Distribution, Separating Agency and Asymmetric Information Mechanisms

(a) Conditional on Renegotiation

(b) $\mathbb{1}_{\{reneg_{hjt}\}}$

Version	Pre-join price quintiles ($\beta^{Info}_{quintile}$)					Pre-join price quintiles ($\beta^{Info}_{quintile}$)				
	1	2	3	4	5	1	2	3	4	5
1	-7	13	11	-22	$-91^{\dagger}$	.008	.002	.018**	.018*	.034†
	(12)	(12)	(14)	(14)	(19)	(.008)	(.009)	(.008)	(.01)	(.008)
2	-11	5	1	-25	$-111^{\dagger}$	.017*	.01	.023**	.025**	.04†
	(16)	(15)	(17)	(17)	(23)	(.009)	(.01)	(.009)	(.011)	(.009)
3	-14	4	1	-13	$-76^{\dagger}$	.01	.013	.016*	.018	.023**
	(15)	(14)	(19)	(17)	(18)	(.01)	(.01)	(.009)	(.011)	(.009)
4	-8	4	-3	-11	$-80^{\dagger}$	.022**	.022*	.024**	.029**	.032†
	(19)	(19)	(22)	(20)	(22)	(.011)	(.012)	(.01)	(.012)	(.011)

Authors' calculations from PriceGuide data, 2009-2014. $N = 6,510$ member-brand-months in regressions conditional on renegotiation. $N = 32,453$ member-brand-months in renegotiation dummies regression. Includes 508 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A15: Treatment Effects Conditional on Renegotiation and on Occurrence of Renegotiation

D.2 Robustness

The following Figures show the results of our richest regression specification, allowing for different treatment effects for different parts of the price and quantity distributions, for specifications that (1) focus only on the twelve months before and after information (Figure A16); (2) limit the sample to those facilities registered with the database as "hospitals" as opposed to "health systems" or other facility types (Figure A17); and (3) examine heterogeneity in results as a function of hospital type (Table A8). Results are qualitatively and quantitatively similar in specifications (1) and (2).

(a) Low Quantity   (b) High Quantity

Version	Pre-info price quintiles ($\beta^{Info}_{quintile,lowq}$)					Pre-info price quintiles ($\beta^{Info}_{quintile,lowq} + \beta^{Info}_{quintile,highq}$)				
	1	2	3	4	5	1	2	3	4	5
1	−7	15**	14**	8	−30†	−11	7	7	1	−54†
	(7)	(6)	(6)	(7)	(9)	(9)	(8)	(7)	(8)	(16)
2	−13	9	6	3	−38†	−16	0	−2	−3	−61†
	(8)	(8)	(7)	(8)	(9)	(10)	(8)	(8)	(9)	(16)
3	−7	10*	9*	14**	−11*	−7	2	5	1	−60†
	(6)	(5)	(5)	(6)	(6)	(8)	(7)	(6)	(7)	(15)
4	−11*	7	5	11*	−16**	−10	−1	1	−2	−66†
	(7)	(6)	(6)	(6)	(7)	(8)	(8)	(6)	(7)	(15)

Authors' calculations from PriceGuide data, 2009-2014. $N = 23{,}016$ member-brand-months. Includes 507 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

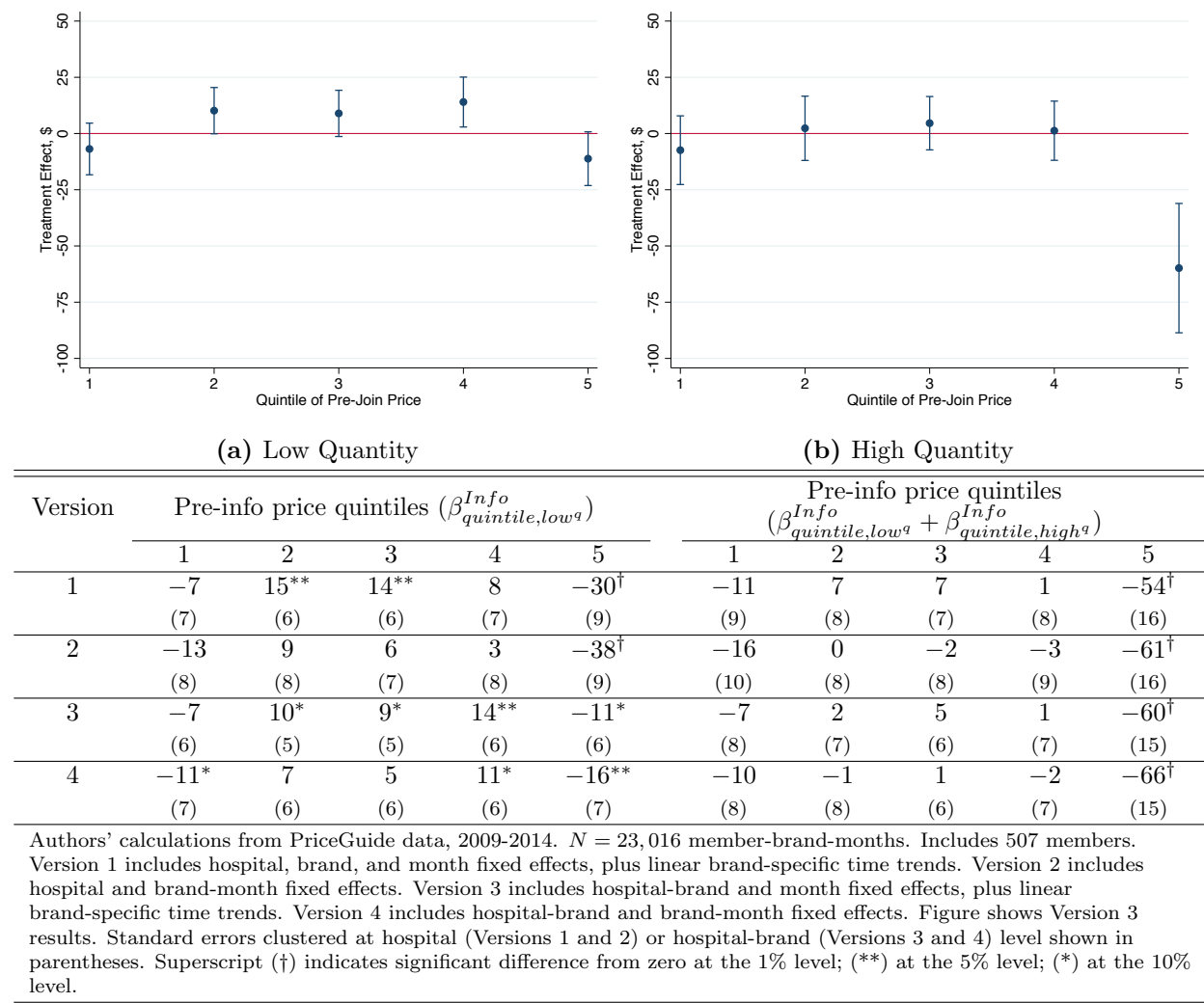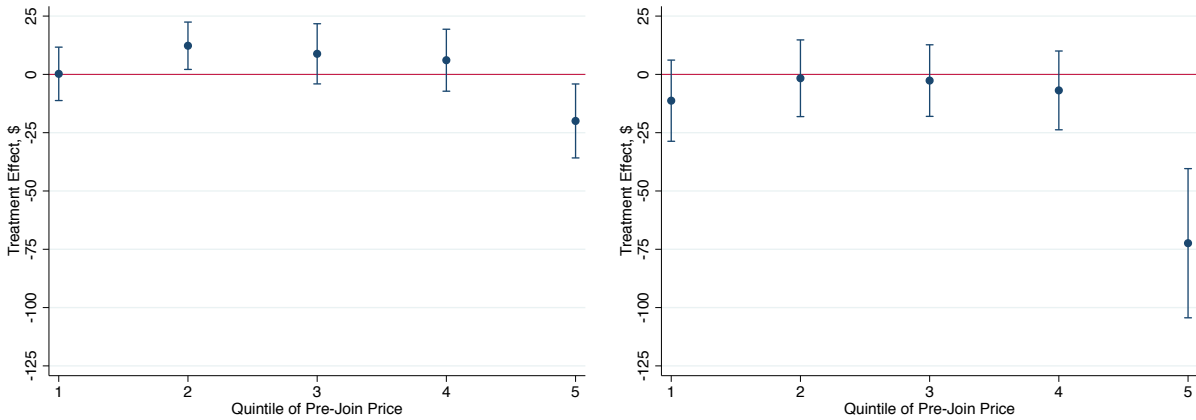**Figure A16:** Treatment Effect Estimates Across the Price and Quantity Distributions – Twelve Months Pre/Post

To further explore convergence within hospitals of differing "types", we also estimate

72

(a) Low Quantity

(b) High Quantity

Version	Pre-info price quintiles ($\beta^{Info}_{quintile,low^q}$)					Pre-info price quintiles ($\beta^{Info}_{quintile,low^q} + \beta^{Info}_{quintile,high^q}$)				
	1	2	3	4	5	1	2	3	4	5
1	8	21†	15**	−2	−53†	−1	8	−10	−7	−80†
	(7)	(6)	(7)	(8)	(11)	(10)	(10)	(10)	(12)	(18)
2	2	13	11	−5	−64†	−5	2	−18	−13	−84†
	(8)	(8)	(8)	(9)	(12)	(11)	(10)	(11)	(12)	(18)
3	0	12**	9	6	−20**	−11	−2	−3	−7	−72†
	(6)	(5)	(7)	(7)	(8)	(9)	(8)	(8)	(9)	(16)
4	−3	7	7	2	−26†	−12	−5	−7	−10	−77†
	(7)	(6)	(8)	(7)	(9)	(9)	(9)	(9)	(9)	(16)

Authors' calculations from PriceGuide data, 2009-2014. $N = 27,698$ member-brand-months. Includes 436 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Figure A17: Treatment Effect Estimates Across the Price and Quantity Distributions – Hospitals Only

73

our richest specification within hospital size bins, within regions, and within size-region.[56] We find several patterns of interest, now reported in Table A8.[57] In the second panel of Table A8, we see suggestive evidence that treatment effects vary across regions. Focusing on the high-price and high-quantity hospital-brands, the treatment effects are smallest in the northeast and south (-\$42 and -\$43, respectively), largest in the west (-\$104), and similar to our baseline, pooled-region estimates (-\$71) in the midwest (-\$70); none of these estimates is statistically significantly different from the pooled-region estimate at the 5% level. For high-price and relatively low-quantity hospital-brands, only the treatment effect for the western region is statistically significant. In the third panel of Table A8, we show that the distribution of treatment effect estimates in price and quantity space differs by hospital size. The treatment effect estimates for small- and medium-sized hospitals (1-199 and 200-399 beds, respectively) are statistically significant and of similar magnitude to our baseline estimate (-\$17) for high-price, relatively low-quantity hospital-brands: the treatment effects estimates are -\$27 and -\$24, respectively. The treatment effect estimates for medium-sized and large hospitals (400+ beds) are large and statistically significant for high-price, relatively high-quantity hospital-brands: the treatment effects estimates are -\$90 and -\$68, respectively. This pattern is driven by the fact that large hospitals tend to purchase brands in large quantities and small hospitals tend to purchase brands in low quantities. In unreported results, we see no clear evidence that the overall distribution of treatment effects varies with hospital size; the treatment effect for all hospital-brands that were high-price pre-join is largest (-\$31) among medium-sized hospitals, compared to (-\$23) for both small and large hospitals.

These within-type results are noisier than our baseline pooled results, as expected, but largely confirm our finding of large treatment effects of information among hospital-brands where there is room for improvement *and* a great deal of spending at stake, at least within the limited set of hospital characteristics we have. Furthermore, the final panel of Table A8 reports estimates that mirror our full-sample regressions in the first panel, but using *k*-nearest neighbors matching. The results are nearly identical to our baseline estimates. Thus, price convergence pertains across all hospital-brands and within sets of similar hospitals.

---

[56]The results split by size crossed with region are noisy, but largely confirm the marginal interactions with size and region separately.

[57]Note that, for the sake of comparison with the baseline results, the flags for high- vs. low-quantity hospital-brands and high- vs. low-price hospital-brands are relative to the full sample, not the sample for the relevant hospital type.

Table A8: Exploring the Results by Hospital Type

(1) Baseline Estimates – Treatment Effect Variation with Quantity Purchased:									
Low quantity $\beta_{quintile,low^q}^{Info}$					High quantity $\beta_{quintile,high^q}^{Info}$				
$p_{quintile}$ 1	2	3	4	5	1	2	3	4	5
-4	9	9	5	-17^{**}	-11	0	0	-9	$-71^{\dagger}$
(6)	(6)	(6)	(6)	(7)	(9)	(8)	(7)	(8)	(13)

(2) Treatment Effect Variation by Region:									
Low quantity $\beta_{quintile,low^q}^{Info}$					High quantity $\beta_{quintile,high^q}^{Info}$				
$p_{quintile}$ 1	2	3	4	5	1	2	3	4	5
Midwest -7	18^{*}	-4	6	-12	-24^{*}	21^{*}	-1	-25	-70^{**}
(9)	(10)	(12)	(13)	(13)	(14)	(12)	(10)	(21)	(30)
Northeast -14	10	$42^{\dagger}$	24^{**}	-9	-11	9	20	-14	-42
(24)	(23)	(14)	(11)	(11)	(27)	(23)	(14)	(14)	(31)
South 7	9	-1	6	11	16^{*}	-5	-6	9	$-43^{\dagger}$
(13)	(7)	(8)	(10)	(16)	(9)	(13)	(12)	(14)	(11)
West -4	-2	9	0	$-42^{\dagger}$	-61^{**}	-4	5	1	$-104^{\dagger}$
(13)	(10)	(13)	(14)	(14)	(29)	(7)	(10)	(19)	(26)

(3) Treatment Effect Variation by Bed Count:									
Low quantity $\beta_{quintile,low^q}^{Info}$					High quantity $\beta_{quintile,high^q}^{Info}$				
$p_{quintile}$ 1	2	3	4	5	1	2	3	4	5
0-199 -16	5	7	-10	-27^{*}	-19	13^{*}	-23	-16	29
(11)	(7)	(12)	(12)	(14)	(22)	(7)	(18)	(23)	(18)
200-399 -4	-2	0	-8	-24^{*}	3	-5	-19	19	$-90^{\dagger}$
(7)	(11)	(9)	(9)	(13)	(10)	(12)	(18)	(14)	(26)
400+ 2	$26^{\dagger}$	19^{*}	$24^{\dagger}$	-3	-14	4	12	-13	$-68^{\dagger}$
(16)	(9)	(10)	(9)	(9)	(13)	(11)	(8)	(10)	(15)

(4) Matching by p, q, Region, Bed Count:									
Low quantity $\beta_{quintile,low^q}^{Info}$					High quantity $\beta_{quintile,high^q}^{Info}$				
$p_{quintile}$ 1	2	3	4	5	1	2	3	4	5
-10	8	23^{**}	17^{**}	-19	-16	-25	8	-2	$-74^{\dagger}$
(11)	(7)	(9)	(7)	(12)	(17)	(22)	(10)	(9)	(9)

Authors' calculations from PriceGuide data, 2009-2014. (1): Full sample $N = 32,453$ member-brand-months; 508 members. (2): $N_{Midwest} = 7,540$; $N_{Northeast} = 5,944$; $N_{South} = 10,451$; $N_{West} = 8,518$. (3): $N_{0-199} = 6,991$; $N_{200-399} = 12,884$; $N_{400+} = 12,578$. (4) based on k-Nearest Neighbors matching algorithm. Each treated hospital-brand hj matched to five similar hospitals $k \neq h$ based on region and bed count, as well as price and quantity *for brand j* in the three months prior to hj treatment. $N_{match} = 25,961$. Standard errors clustered at hospital-brand level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

D.3 Cross-Supplier Externalities in Bargaining

As noted briefly in the main text, a full analysis of the information we examine should take into account the potential externality that information relevant for one supplier imposes on negotiations with other suppliers. For example, consider the delegated agent model in Collard-Wexler et al. (2017) (which would apply to our context since transfers are not lump sum). In such a model, information about supplier A's bargaining parameter would enter all other suppliers' negotiations via its expected impact on supplier A's price, which in turn

enters the elasticity and marginal contribution terms in all other suppliers' negotiations. In our framework, this consideration would yield an additional empirical prediction:

Prediction 7 (Indirect Information/Competition Effect on All Prices) With imperfect substitute products, under reasonable assumptions on how the negotiation for one brand affects the disagreement payoff of other brand negotiations, a fall in price of substitute brand j will decrease the surplus up for negotiation for other brands $-j$, leading to a decrease in the prices of other brands $-j$, all else equal.[58] Thus exposure to benchmarking information that leads to a fall in a high price for j should also lead to a fall in any price for other brands $-j$, and the size of this fall will be increasing to the extent the brands are good substitutes for j.

Table A9: Externalities: Effect of New Brand Information on Other Brands

Version of Controls	[1]	[2]	[3]	[4]
β^{Join}	$-32^{\dagger}$	$-32^{\dagger}$	-9	-9
	(10)	(10)	(8)	(8)
$\beta^{OtherInfo}$	$19^{\dagger}$	$23^{\dagger}$	-4	2
	(8)	(8)	(8)	(9)
$\beta^{Externality}$	-15	-17	-23	-25
	(19)	(19)	(18)	(19)
Hospital+Brand FEs	Y	Y	N	N
Hospital×Brand FEs	N	N	Y	Y
Linear Brand Trends	Y	N	Y	N
Brand×Month FEs	N	Y	N	Y

Authors' calculations from PriceGuide data, 2009-2014. $N = 11,975$ member-brand-months. Includes 508 members. Standard errors clustered at hospital level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

While our empirical approach in the main text does incorporate these types of externalities, it is difficult to disentangle them. To shed additional light, we run quadruple-difference regressions of the form:

$$
\begin{aligned}
P_{hjt} = {} & \beta^{Join} * \mathbb{1}_{\{post_{ht}^{Join_h}\}} + \beta^{OtherInfo} * \mathbb{1}_{\{post_{hjt}^{Info_{hj' \neq j}}\}} \\
& + \beta^{Externality} * \mathbb{1}_{\{post_{hjt}^{Info_{hj' \neq j}}\}} * \mathbb{1}_{\{P_{hj',pre_h} \in quintile(5, \{P_{h'j',pre_h}\}_{h' \in \mathcal{H}})\}} * \mathbb{1}_{\{high_{hj',pre}^q\}} \\
& + \theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j}) + \varepsilon_{hjt}.
\end{aligned}
\tag{6}
$$

---

[58]This will be the case in any model where disagreement payoffs are a function of the prices agreed to with other manufacturers, which has been the case in the empirical bargaining literature thus far and much of the negotiation with externalities theory. It would not be the case in a model such as the Core, where disagreements are based on the primitives of willingness-to-pay and costs.

When a hospital joins the database, all brands receive information shocks, so any externalities between brands are captured in the treatment effect (β^{Join}). In cases of new brand entry, there are information shocks to single brands, so in principle this provides an opportunity to identify externalities. The challenge is that, among hospitals receiving these information shocks, new brand entry occurs sometime after join, so the simple effect of the new brand information on existing brands ($\beta^{OtherInfo}$) may capture delayed treatment effects of joining – like picking the post month in an event study that corresponds with the new brand information date. To address this, we focus on the instances where the hospital is a high-price, high-quantity case when information is revealed about the entering brand. These are the instances where the information shock is strongest, and we difference them versus the rest of the sample for that same brand to isolate the externality effect on other brands in the same hospital-month ($\beta^{Externality}$). For our preferred specifications with hospital-brand fixed effects, the point estimates in Table A9 suggest externalities less than half the size of the point estimates for the focal entering brands, but the estimates are quite noisy, and none are statistically different from zero at standard confidence levels.

E   Expanded Analyses – Detail

The estimates of our main difference-in-differences regression of price on join timing for each "important" product category are shown in Table A10. In the results in Section 6 in the main text, we summarize these results across product category within each physician preference class. In the detailed results here, the left panel shows the results for the full sample described in Appendix A.3. The right panel shows the results for a restricted sample of brands with "significant" market presence (defined as those brands purchased by at least 100 hospitals per year on average), where we can be more confident that our price quintile estimates have reasonable statistical power. In each case, the regression coefficient has been normalized by the mean unit price of the product category to facilitate comparisons across products with dramatically different prices.

The columns in Table A10 are presented as follows: within each sample, the first column shows the average treatment effect; the second through sixth columns show the treatment effect by price quintile; and the seventh column shows the treatment effect for hospital-brands in the top quintile of the pre-join price distribution *and* in the top quartile of the pre-join quantity distribution. Each panel of rows shows results for each individual product category, as well as a spend-weighted average for each product class.

While there is quite a bit of variation in results for this large number of product categories and different specifications, several patterns emerge. The average treatment effects are, as

with stents, small and not often statistically significant. Also echoing the findings for stents: the treatment effects for the top quintile of the price distribution are generally negative and, particularly for PPIs, often statistically significant. The treatment effects are often, though not universally, larger in magnitude for high-price, high-quantity hospital-brands. In contrast to what we find for stents, the treatment effects for hospital-brands in the bottom quintiles of the pre-join price distribution are sometimes negative, but are also sometimes positive and significant. Focusing on PPIs, the average treatment effects in the bottom of the distribution are positive and significant (1.4 percent), though smaller than the price decreases documented in the top of the distribution (3.3 percent).

On balance, as we found with drug-eluting stents, these results demonstrate that the largest savings are achieved by hospitals and for PPI brands with the highest prices and quantities prior to the hospital joining the database. The treatment effects are often larger in magnitude for the restricted sample of brands purchased by an average of 100 or more hospitals; those brands may have received additional attention from hospitals searching the benchmarking database.

In the following analyses, we explore the sensitivity of our focal findings (significant savings among high-price, high-quantity hospital-brands) to alternative regression specifications. We then further explore the small positive price effects documented in the bottom of the distribution.

Table A10: Treatment Effects for Important Product Categories, with hj and jy FE

ID	Name	Class	Unrestricted Sample							$|\mathcal{H}_j| \geq 100$ Sample						
			ATE	TE by Price Quintile					$TE_{high,high}$	ATE	TE by Price Quintile					$TE_{high,high}$
				1	2	3	4	5			1	2	3	4	5	
10279	Bandages Elastic	1	-0.010	0.007	-0.018	0.002	-0.038**	-0.009	-0.018	-0.014	0.002	0.025*	-0.002	-0.110*	-0.033	-0.046
11315	Dressings	1	0.000	0.021**	-0.001	0.020*	-0.024	-0.021	-0.022*	0.007*	0.021**	0.008	0.003	-0.001	0.001	0.005
11883	Gloves Surgical	1	-0.005	-0.018	0.028	0.008	-0.028	-0.016	-0.009	-0.008**	-0.005	-0.014**	0.000	-0.011	-0.009	-0.006
12368	Drapes Surgical	1	0.007	0.004	0.024**	0.009	0.015**	-0.024	0.001	-0.013	-0.019	-0.009	0.035**	-0.000	-0.082**	0.007
12745	Needles Injection Hypodermic	1	-0.000	0.014**	0.005	-0.000	0.004	-0.020**	-0.028	0.000	0.011**	0.007	0.003	0.008**	-0.026†	-0.310†
13655	Containers Specimen	1	0.018	0.082	0.004	0.003	0.005	0.011	-0.045**	0.050	-0.046	0.040	0.100**	0.102**	-0.007	-0.019
923840	Drill Bits	1	0.004	0.019†	0.008**	-0.000	0.004	-0.011	-0.012	0.011*	0.022**	0.019†	0.004	0.017*	-0.006	-0.019
	Commodities	1	0.002	0.013*	0.013†	0.006*	-0.009	-0.013†	-0.016**	0.007	-0.003	0.008	0.021**	0.013*	-0.019	-0.053†
10688	Catheters Vascular Angiography	2	0.004	0.025**	0.001	0.001	0.007	-0.013	-0.078**	-0.001	0.000	0.005	0.001	0.012	-0.017	-0.092**
10894	Clip Appliers	2	0.008**	0.035**	0.006	-0.001	0.012**	-0.008	-0.043**	0.008*	0.040**	0.013	-0.006	0.015*	-0.013	-0.055**
11502	Forceps Electrosurgical	2	-0.000	-0.002	0.004**	0.003	0.001	-0.001	-0.001	-0.000†	-0.002	0.006**	0.004	-0.000	-0.001**	-0.001*
11910	Grafts Bone	2	-0.001	-0.008	0.004	-0.005	0.001	-0.000	0.000	-0.000	-0.012	0.000†	0.000†	0.000†	0.000	0.000
11925	Guide Wires	2	0.004	0.007	0.004	0.008*	0.001	-0.001	0.005	0.003	-0.002	0.013	0.009	-0.003	-0.002	0.004
12830	Orthopedic Cement	2	-0.004	-0.007	0.001	0.019†	-0.011	-0.017	-0.005	-0.002	0.004	0.004	0.030*	-0.037*	-0.016	-0.023
13050	Plates Bone	2	-0.003	0.004	0.008	-0.001	-0.010	-0.017**	-0.026**	0.002	0.005	0.019**	0.018†	0.006	-0.034*	-0.076*
13909	Sutures Synthetic Nonabsorbable Polypropylene	2	-0.000	-0.019	0.033*	0.009	0.001	-0.017	0.008	0.011*	-0.012†	0.025**	0.009	0.021†	0.010	0.026
13910	Sutures Natural Nonabsorbable Silk	2	0.009	0.007	0.029	0.009	0.028	-0.022	-0.047	0.010	0.002	0.013*	0.012	0.044	-0.021	-0.075*
14085	Tubes Tracheal	2	0.005	0.015**	0.023†	0.008	0.011	-0.046†	-0.011	0.002	0.009	0.060**	0.016	-0.022	-0.086**	-0.056
16071	Lenses Intraocular Posterior Chamber	2	0.008*	0.009	0.011*	0.012**	0.007	0.005	0.003	0.008**	0.028**	0.014**	0.009†	0.012**	-0.002	0.003
16078	Nails Bone	2	0.009	0.002	0.024†	0.017†	0.027†	-0.024	-0.015	0.010	0.002	0.029†	0.012	0.033†	-0.028	-0.018
16104	Wires Bone	2	0.013	0.053**	0.017*	0.027†	0.013	-0.031	-0.040	0.021*	0.038	0.023	0.026**	0.013	-0.001	0.023
16655	Catheters Cardiac Ablation	2	0.000	0.011	-0.003	0.008	0.002	-0.017**	-0.032**	0.000	-0.001	-0.001	-0.005	0.006	0.001	-0.006
17184	Catheters Vascular Angioplasty Balloon	2	-0.005**	0.003	0.007*	0.004	-0.009**	-0.030†	-0.035†	-0.006**	0.002	0.008**	0.003	-0.007	-0.032†	-0.047†
17471	Sutures Synthetic Absorbable Polyglactin	2	-0.025	-0.056	-0.008	-0.016	0.041	-0.110	0.214	0.033	-0.004	-0.045	0.047	0.152**	0.002	0.037
17600	Bone Matrix Implants	2	-0.011**	-0.007	-0.027**	-0.007	-0.002	-0.007	-0.027**	-0.007	0.028	-0.048†	-0.001	-0.021*	0.012	-0.003
17846	Catheters Vascular Guiding	2	0.004	0.010	0.016**	0.003	-0.002	-0.004	0.003	0.006	0.009	0.021*	0.002	-0.005	0.003	0.009
18609	Catheters Cardiac Mapping/Ablation	2	0.005	0.042*	0.016	-0.035*	0.009	-0.003	0.004	-0.002	0.029	0.001	-0.053**	0.023*	-0.023	-0.055
20317	Staplers Surgical	2	0.001	-0.030	0.018	0.015	0.032	-0.034	0.007	-0.001	-0.045*	0.021	0.010	0.027	-0.032	0.043
20318	Staplers Surgical Linear	2	0.001	-0.005	-0.049	0.015	0.020*	-0.002	0.056	0.001	0.019**	0.001	-0.008	0.030**	-0.032*	0.004
20453	Stent/Grafts Vascular Aortic	2	-0.000	-0.005	-0.004	0.014	-0.001	0.001	-0.004	-0.003	-0.017	-0.011	0.002	0.002	0.004	-0.002
22538	Suture Anchors	2	-0.000	0.005	0.005	0.009	0.000	-0.008	-0.001	-0.001	0.007	0.009	0.009	-0.001	-0.014	-0.007
22908	Sutures Synthetic Nonabsorbable Nylon Monofilament	2	-0.004	0.046**	0.061	-0.003	-0.019	-0.068**	-0.047	-0.011	0.047	0.077	-0.026	-0.040**	-0.073†	-0.004
34196	Screws Bone Spinal	2	-0.034*	0.004	-0.037	-0.001	-0.083†	-0.061*	-0.016	-0.064**	-0.014	-0.068	-0.003	-0.117**	-0.145**	-0.231**
	Other Med/Surg	2	-0.003	0.003	0.001	0.003	-0.004	-0.017†	-0.006	-0.005	0.001	-0.002	0.001	-0.004	-0.025†	-0.038†
12324	Lenses Intraocular	3	0.001	0.051**	0.012	0.010*	0.012*	-0.009*	-0.008	0.002	0.105**	0.000†	0.002**	0.001†	0.000	-0.000
12913	Pacemakers Cardiac Implantable	3	0.009*	0.035†	0.012**	-0.003	0.009	-0.028**	-0.028	0.009**	0.038†	0.012*	-0.004	0.012	-0.031**	-0.036
15766	Orthopedic Internal Fixation Systems Spinal	3	-0.017*	0.003	-0.004	-0.012	-0.050**	-0.034**	-0.041*	-0.021	-0.011	0.016	-0.042**	-0.063**	-0.013	0.026
15870	Prostheses Cardiac Valve Biological	3	-0.002	0.008	0.008	0.004	-0.015*	-0.015	0.003	-0.006	-0.003	0.020	0.006	-0.016	-0.047	-0.025
16040	Stents Ureteral	3	-0.000	0.005	-0.007	0.008*	0.008	-0.020*	-0.015	0.003	0.003	0.002	0.004	0.017*	-0.022	-0.036
16084	Prostheses Joint Hip Acetabular Component	3	-0.006	0.044†	0.009	-0.013	-0.025†	-0.086†	-0.102†	-0.012	0.044†	0.004	0.014	-0.040**	-0.123†	-0.197†
16095	Prostheses Joint Hip Femoral Component	3	0.004	0.047†	0.020†	-0.002	-0.024**	-0.065†	-0.049**	0.003	0.056†	-0.003	-0.017	-0.003	-0.070†	-0.078*
16097	Prostheses Joint Knee Femoral Component	3	0.003	0.045†	0.014**	0.006	-0.012	-0.056†	-0.058†	-0.002	0.052†	0.015**	-0.008	-0.027	-0.074†	-0.107**
16098	Prostheses Joint Knee Tibial Component	3	-0.002	0.041†	0.017**	0.006	-0.020**	-0.079†	-0.048**	-0.009	0.019*	0.000	-0.011	-0.013	-0.062†	0.010
16111	Prostheses Mammary Internal	3	0.003	0.025†	0.005	0.016**	-0.002	-0.014	-0.016	0.004	0.030†	0.005	0.014**	-0.004	-0.012	-0.020
16921	Prostheses Joint Shoulder Humeral Component	3	-0.001	0.001	-0.002	0.003	-0.009	0.000	-0.028	-0.027	-0.046**	-0.045	-0.007	0.004	-0.035	-0.107
17165	Allografts	3	0.008	0.043*	0.069†	0.028	-0.010	-0.033	-0.024	0.049	0.115	0.105	0.144*	0.008	-0.164	-0.690
17241	Stimulators Electrical Spinal Cord Analgesic	3	-0.006	-0.008	-0.017†	-0.012**	0.016**	-0.007	-0.027*	-0.002	-0.014	-0.014**	-0.001	0.021**	-0.005	-0.026
18253	Grafts Skin Biological	3	-0.008*	-0.016**	-0.020	0.000	-0.016	0.003	0.011	-0.011**	-0.025**	-0.046**	0.001	-0.017	0.006	0.017
18504	Defibrillator/Cardioverter/Pacemakers Implantable	3	-0.013*	0.011	0.014*	-0.013	-0.038**	-0.042†	-0.028	-0.019**	0.009	0.038**	-0.037	-0.031	-0.051†	-0.065*
20376	Defibrillator/Cardioverter/Pacemakers Implantable Resynchronization	3	-0.002	0.013	0.010	0.006	-0.021	-0.017	-0.026	-0.012	-0.009	0.009	0.002	-0.038**	-0.011	-0.007
20383	Stents Vascular Coronary Balloon-Expandable Drug-Eluting	3	-0.013†	-0.013**	-0.018**	-0.008	-0.011	-0.019	-0.075†	-0.014†	-0.013**	-0.019**	-0.008	-0.011	-0.023*	-0.079†
20422	Stents Vascular Coronary Balloon-Expandable	3	-0.010*	-0.025**	0.002	0.009	0.005	-0.040†	-0.031**	-0.005	-0.018	0.006	0.014	0.009	-0.043†	-0.034*
22543	Occluders Vascular Intravascular Embolization Coil	3	0.004	0.008	-0.003	0.015*	-0.002	0.002	0.018	0.010	0.032†	-0.032	0.039	0.017	0.002	0.005
923150	Orthopedic Fixation Systems Implantable Spine	3	-0.013**	0.007	0.016†	-0.014†	-0.023**	-0.056†	-0.047†	-0.037	-0.010	0.055	0.039	0.014	-0.194†	-0.298†
	PPIs	3	-0.005†	0.014†	0.006**	-0.003	-0.016†	-0.033†	-0.039†	-0.009†	0.010*	0.011**	-0.001	-0.015†	-0.052†	-0.081†

Authors' calculations from PriceGuide data, 2009-2014. Specifications include hospital-brand and brand-year fixed effects; alternative fixed effects in Table A11. Standard errors clustered at hospital-brand level. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Table A10 reported the results of our preferred specification, with controls for hospital-brand and brand-year fixed effects. This choice of specification is intended to accommodate trends in technology prices across a wide variety of product categories, some of which are purchased frequently and some infrequently. Table A11 below compares summary results with alternative fixed effects: hospital and brand fixed effects vs. hospital-brand fixed effects, and brand-specific linear trends vs. brand-year fixed effects vs. brand-month fixed effects. In each case, we show spend-weighted average treatment effects within each product class, for the unrestricted and restricted samples.

Table A11: Treatment Effects with Alternative Fixed Effects

	Unrestricted Sample						$|\mathcal{H}_j| \geq 100$ Sample					
	ATE		TE_{high^p}		$TE_{high^p,high^q}$		ATE		TE_{high^p}		$TE_{high^p,high^q}$	
	\$	%	\$	%	\$	%	\$	%	\$	%	\$	%
$\theta_h + \theta_j + \theta_t + \gamma_j(t - tmin_j)$												
Commodities	-0.10	-0.007	-3.48	-0.084	-3.75	-0.051	0.37*	0.004	-2.18†	-0.056†	-3.13†	-0.069†
	(0.23)	(0.014)	(993.66)	(5.768)	(333.35)	(2.256)	(0.22)	(0.007)	(0.46)	(0.016)	(0.81)	(0.019)
Other Med/Surg	0.91	-0.002	-32.69	-0.059	-39.58	-0.086	-2.73	-0.002	-25.82†	-0.049†	-43.12†	-0.075†
	(2.63)	(0.002)	(5023.61)	(2.844)	(1523.34)	(1.886)	(3.08)	(0.003)	(6.93)	(0.010)	(11.50)	(0.015)
PPIs	36.78	-0.001	-349.07	-0.072	-506.62	-0.087	-29.31	-0.005	-237.32†	-0.069†	-486.63†	-0.108†
	(27.44)	(0.002)	(2514.22)	(3.420)	(3389.52)	(4.044)	(34.88)	(0.004)	(52.02)	(0.008)	(105.63)	(0.018)
$\theta_{hj} + \theta_t + \gamma_j(t - tmin_j)$												
Commodities	0.24	0.001	-0.36	-0.009	-0.33	-0.009	0.50**	0.010	-0.09	-0.012	-0.50	-0.038†
	(0.19)	(0.004)	(0.41)	(0.006)	(0.60)	(0.011)	(0.20)	(0.006)	(0.45)	(0.012)	(0.83)	(0.011)
Other Med/Surg	0.16	-0.000	-6.36**	-0.014†	-0.16	-0.002	-3.30	-0.002	-12.10**	-0.021†	-25.15†	-0.033†
	(2.24)	(0.002)	(3.06)	(0.004)	(5.47)	(0.006)	(2.93)	(0.003)	(5.86)	(0.008)	(9.23)	(0.012)
PPIs	30.96	0.002	-123.25†	-0.025†	-92.65	-0.026†	-27.72	-0.003	-198.47†	-0.045†	-276.28†	-0.073†
	(21.80)	(0.002)	(42.74)	(0.004)	(57.87)	(0.005)	(31.97)	(0.003)	(51.68)	(0.008)	(89.28)	(0.017)
$\theta_h + \theta_{jy}$												
Commodities	-0.05	-0.001	-3.35†	-0.072†	-3.88†	-0.073*	0.28	0.002	-2.32†	-0.064†	-3.19†	-0.082†
	(0.21)	(0.013)	(0.43)	(0.027)	(0.66)	(0.040)	(0.21)	(0.006)	(0.44)	(0.016)	(0.75)	(0.023)
Other Med/Surg	-1.81	-0.005**	-35.43†	-0.059†	-43.53†	-0.078†	-5.06	-0.006*	-29.46†	-0.054†	-47.27†	-0.081†
	(2.58)	(0.002)	(4.03)	(0.006)	(5.43)	(0.016)	(3.10)	(0.003)	(6.56)	(0.009)	(15.05)	(0.015)
PPIs	-36.00	-0.008†	-417.52†	-0.081†	-601.89†	-0.100†	-79.45**	-0.011†	-295.05†	-0.077†	-568.09†	-0.117†
	(25.31)	(0.002)	(48.48)	(0.004)	(84.13)	(0.007)	(32.25)	(0.003)	(51.81)	(0.008)	(105.52)	(0.017)
$\theta_{hj} + \theta_{jy}$												
Commodities	0.17	0.002	-0.52	-0.013†	-0.58	-0.016**	0.39*	0.007	-0.28	-0.019	-0.72	-0.053†
	(0.19)	(0.004)	(0.39)	(0.005)	(0.56)	(0.008)	(0.20)	(0.006)	(0.42)	(0.012)	(0.76)	(0.017)
Other Med/Surg	-1.86	-0.003	-8.72†	-0.017†	-8.89**	-0.006	-5.44*	-0.005	-14.87**	-0.025†	-28.09†	-0.038†
	(2.22)	(0.002)	(3.10)	(0.004)	(3.90)	(0.005)	(3.03)	(0.003)	(5.80)	(0.008)	(10.01)	(0.012)
PPIs	-38.54*	-0.005†	-189.55†	-0.033†	-206.32†	-0.039†	-78.09†	-0.009†	-246.15†	-0.052†	-342.61†	-0.081†
	(23.09)	(0.002)	(46.82)	(0.004)	(74.41)	(0.005)	(30.10)	(0.003)	(51.66)	(0.008)	(90.21)	(0.016)
$\theta_h + \theta_{jm}$												
Commodities	0.04	-0.003	-3.11†	-0.058**	-3.91†	-0.061	0.39*	0.003	-2.22†	-0.059†	-3.17†	-0.080†
	(0.23)	(0.011)	(0.46)	(0.023)	(0.77)	(0.041)	(0.22)	(0.007)	(0.46)	(0.016)	(0.77)	(0.022)
Other Med/Surg	-1.42	-0.003	-33.37†	-0.057†	-46.54†	-0.081†	-2.53	-0.003	-25.73†	-0.049†	-47.17	-0.083†
	(3.04)	(0.003)	(6.37)	(0.009)	(8.30)	(0.016)	(3.42)	(0.004)	(7.21)	(0.010)	(62.15)	(0.021)
PPIs	46.24	-0.000	-351.19†	-0.077†	-561.28†	-0.095†	-28.86	-0.006	-220.50†	-0.069†	-466.85†	-0.105†
	(31.32)	(0.002)	(57.60)	(0.005)	(97.90)	(0.008)	(38.35)	(0.004)	(56.23)	(0.009)	(103.52)	(0.018)
$\theta_{hj} + \theta_{jm}$												
Commodities	0.28	0.002	-0.36	-0.013†	-0.68	-0.018**	0.49**	0.009	-0.15	-0.015	-0.68	-0.052†
	(0.20)	(0.003)	(0.42)	(0.005)	(0.66)	(0.007)	(0.21)	(0.006)	(0.44)	(0.013)	(0.80)	(0.017)
Other Med/Surg	-1.34	-0.002	-8.87*	-0.019†	-11.80	-0.013	-3.69	-0.003	-12.47**	-0.022†	-29.35**	-0.042**
	(2.57)	(0.003)	(4.99)	(0.007)	(8.13)	(0.011)	(3.05)	(0.003)	(5.90)	(0.008)	(12.77)	(0.018)
PPIs	29.25	0.002	-150.63†	-0.030†	-170.29*	-0.035†	-29.29	-0.003	-198.75†	-0.045†	-284.08†	-0.072†
	(28.57)	(0.002)	(55.41)	(0.005)	(101.30)	(0.007)	(35.43)	(0.004)	(54.51)	(0.008)	(94.47)	(0.017)

Authors' calculations from PriceGuide data, 2009-2014. $N_{hjt}^{(1)} = 516,582$; $N_{hjt}^{(2)} = 1,344,515$; $N_{hjt}^{(3)} = 703,544$. $N_h^{(1)} = 748$; $N_h^{(2)} = 701$; $N_h^{(3)} = 601$.
Fixed effects for each panel indicated in first column. Standard errors in category-specific regressions clustered at hospital-brand level. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

The results indicate that inclusion of hospital-brand fixed effects has an important effect

on the estimates – in the first panel, which includes hospital and brand and month fixed effects, as well as brand-specific linear trends, the standard errors are quite large, capturing the fact that many product categories in the expanded analysis are more heterogeneous than drug-eluting stents and estimates can be sensitive to compositional changes. Within the set of results with hospital-brand fixed effects, the results with different controls for time trends are qualitatively similar: for example, among PPIs, the high-price, high-quantity hospital-brand results range from −0.025 to −0.039 in the unrestricted sample, and from −0.073 to −0.081 in the restricted sample. Similarly, the unrestricted sample treatment effects for high-price, high-quantity hospital-brands range from −0.009 to −0.018 for commodities and from −0.002 to −0.013 for other medical/surgical products.

Finally, Table A12 examines sensitivity of our preferred fixed effects specification to whether it is estimated as a pooled regression of all price quintiles together (left panel) vs. separate regressions of each price quintile (right panel). In each panel, we show results for the bottom and top quintiles of the price distribution, for all hospital-brands, low-quantity hospital-brands, and high-quantity hospital-brands.[59]

Several patterns emerge in this comparison. First, the results for the top quintile of the price distribution are both qualitatively and quantitatively similar regardless of whether the specification is estimated separately by quintile or pooled. The top-quintile average effects by class are −0.013 for commodities, −0.017 for other medical/surgical, and −0.033 for PPIs in the pooled specification; these estimates change to −0.012 for commodities, −0.012 for other medical/surgical, and −0.026 for PPIs in the separate-quintiles specifications. While the increased richness of the control set in the separate-quintiles specifications has the expected effect of decreasing the precision of the estimates and reduces the magnitudes slightly, the general findings are robust to specification: we document small, significant effects of transparency; larger effects for PPIs; and particularly large effects for relatively high-quantity purchasers of PPIs. We also note that the qualitative patterns across product categories are similar across specification: e.g., some of the largest effects are for high-price, high-quantity joint prostheses.

Second, the results for the bottom quintile of the price distribution are sensitive to specification. The bottom-quintile average effects by class are 0.013 for commodities, 0.003 for other medical/surgical, and 0.014 for PPIs in the pooled specification; however, these estimates change to −0.005 for commodities, −0.009 for other medical/surgical, and −0.007 for PPIs in the separate-quintiles specifications. The qualitative pattern across product

---

[59]The primary difference between these specifications is that the pooled regression imposes the same brand-year trends across hospitals in all quintiles of the price distribution, whereas the non-pooled version allows brand-year trends to vary based on hospitals' position in the price distribution.

categories also changes: many of the positive and significant bottom-quintile results for joint prostheses observed in the left panel shrink dramatically or disappear in the right panel. These differences in the bottom-quintile results may be due to differential brand-specific price trends for low-price hospitals. For example, if low-price hospitals are those that have relatively high bargaining parameters, it may be the case that those hospitals achieve high discounts regardless of changes in the market environment and accordingly, exhibit less steep downward trends than other hospitals.

Table A12: Treatment Effects for Important Product Categories, Alternative Quintiles Treatments

ID	Name	Class	Pooled Regressions							Separated Regressions by Price Quintile						
			ATE	All Q		Low Q		High Q		ATE	All Q		Low Q		High Q	
				1	5	1	5	1	5		1	5	1	5	1	5
10279	Bandages Elastic	1	-0.011	0.007	-0.01	-0.017	-0.005	0.038	-0.018	-0.011	-0.008	-0.012	-0.021*	-0.01	0.019	-0.015
11315	Dressings	1	0.000	0.021**	-0.021	0.027*	-0.02	0.014**	-0.022*	0.000	0.007	-0.017	0.017	-0.014	-0.002	-0.023
11883	Gloves Surgical	1	-0.005	-0.018	-0.016	-0.017	-0.019	-0.019	-0.009	-0.005	-0.008**	-0.015	-0.008**	0.026	-0.010*	-0.004
12368	Drapes Surgical	1	0.007	0.004	-0.024	-0.005	-0.033*	0.016	0.001	0.007	-0.013	-0.02	-0.018	-0.006	-0.001	0.006
12745	Needles Injection Hypodermic	1	0.000	0.014**	-0.020**	0.008	-0.018*	0.026*	-0.028	0.000	-0.003	-0.016*	-0.003	0.002	0.000	-0.028
13655	Containers Specimen	1	0.018	0.082	0.011	0.174	0.031*	-0.029	-0.045**	0.018	-0.014	0.01	-0.003	0.005	-0.017	-0.050**
923840	Drill Bits	1	0.004	0.019†	-0.011	0.009	-0.01	0.032†	-0.012	0.004	0.002	-0.01	-0.001	0.002	0.007	-0.016
	Commodities	1	0.002	0.013*	-0.013†	0.021	-0.012**	0.005	-0.016**	0.002	-0.005*	-0.012**	-0.004	0.006	-0.003	-0.015**
10688	Catheters Vascular Angiography	2	0.004	0.025**	-0.013	0.019	0.021	0.035*	-0.078**	0.004	0.017	-0.01	0.021*	0.015	0.018	-0.063**
10894	Clip Appliers	2	0.008**	0.035**	-0.008	0.026	0.004	0.052*	-0.043**	0.008**	0.011	-0.006	0.003	0.021**	0.033	-0.050**
11502	Forceps Electrosurgical	2	0.000	-0.002	-0.001	0.007	0.000	-0.006	-0.001	0.000	-0.004	0.000	-0.017	0.001	-0.001	0.001*
11910	Grafts Bone	2	-0.001	-0.008	0.000	-0.004	0.000	-0.011	0.000	-0.001	-0.020†	0.001	-0.038**	0.001	-0.012	0.002
11925	Guide Wires	2	0.004	0.007	-0.001	0.008	-0.004	0.006	0.005	0.004	-0.009**	0.000	-0.007	0.000	-0.012*	0.004
12830	Orthopedic Cement	2	-0.004	-0.007	-0.017	0.017	-0.022	-0.035	-0.005	-0.004	-0.006	-0.019	-0.009	-0.001	-0.003	-0.021
13050	Plates Bone	2	-0.003	0.004	-0.017**	0.001	-0.011	0.008	-0.026**	-0.003	-0.011**	-0.018**	-0.012*	0.007	-0.009	-0.031**
13909	Sutures Synthetic Nonabsorbable Polypropylene	2	0.000	-0.019	-0.017	-0.004	-0.027**	-0.049	0.008	0.000	-0.103**	-0.013	-0.101	-0.015	-0.076	-0.006
13910	Sutures Natural Nonabsorbable Silk	2	0.009	0.007	-0.022	0.002	-0.016	0.015	-0.047	0.009	-0.014**	-0.016	-0.025†	0.006	0.001	-0.072
14085	Tubes Tracheal	2	0.005	0.015**	-0.046†	0.01	-0.060†	0.020**	-0.011	0.005	-0.003	-0.026*	-0.005	-0.024	-0.004	0.021
16071	Lenses Intraocular Posterior Chamber	2	0.008*	0.009	0.005	0.016	0.006	0.003	0.003	0.008*	0.01	0.008	0.011	0.005	0.011	0.011
16078	Nails Bone	2	0.009	0.002	-0.024	-0.013	-0.031*	0.016*	-0.015	0.009	-0.01	-0.019	-0.024**	-0.015	0.003	-0.007
16104	Wires Bone	2	0.013	0.053**	-0.031	0.055**	-0.026	0.050**	-0.04	0.013	-0.006	-0.014	-0.015	-0.03	0.018	-0.015
16655	Catheters Cardiac Ablation	2	0.000	0.011	-0.017**	-0.015	-0.01	0.036	-0.032**	0.000	0.011	-0.012	0.009	0.002	0.024	-0.028*
17184	Catheters Vascular Angioplasty Balloon	2	-0.005**	0.003	-0.030†	0.001	-0.028†	0.006	-0.035†	-0.005**	-0.015†	-0.024†	-0.018†	-0.015**	-0.011†	-0.029†
17471	Sutures Synthetic Absorbable Polyglactin	2	-0.025	-0.056	-0.11	0.037	-0.184**	-0.122	0.214	-0.025	-0.075	-0.091	-0.007	-0.107	-0.081	0.289
17600	Bone Matrix Implants	2	-0.011**	-0.007	-0.007	-0.017	0.002	0.003	-0.027**	-0.011**	-0.006	-0.007	-0.022	0.004	0.007	-0.025**
17846	Catheters Vascular Guiding	2	0.004	0.01	-0.004	0.019**	-0.006	-0.001	0.003	0.004	0.005	-0.005	0.011	-0.003	0.005	0.006
18609	Catheters Cardiac Mapping/Ablation	2	0.005	0.042**	-0.003	0.057**	-0.005	0.013	0.004	0.005	-0.004	0.003	-0.01	-0.01	-0.024	0.016
20317	Staplers Surgical	2	0.001	-0.03	-0.034	-0.032	-0.049	-0.029	0.007	0.001	-0.024	0.032	-0.036*	0.036	-0.014	0.002
20318	Staplers Surgical Linear	2	0.001	-0.005	-0.002	-0.019	-0.034**	0.012	0.056	0.001	-0.014	-0.003	-0.025	0.005	0.000	0.037
20453	Stent/Grafts Vascular Aortic	2	0.000	-0.005	0.001	-0.015	0.003	0.004	-0.004	0.000	-0.002	0.002	0.004	0.001	-0.003	0.004
22538	Suture Anchors	2	0.000	0.005	-0.008	0.006	-0.011	0.004	-0.001	0.000	0.000	-0.005	0.007	-0.009	-0.008	0.003
22908	Sutures Synthetic Nonabsorbable Nylon Monofilament	2	-0.004	0.046**	-0.068**	0.023*	-0.078**	0.078**	-0.047	-0.004	-0.039†	-0.034	-0.005	-0.021	-0.09	-0.028
34196	Screws Bone Spinal	2	-0.034*	0.004	-0.061**	0.019	-0.079**	-0.004	-0.016	-0.034*	-0.003	-0.062	0.01	-0.081	-0.006	-0.022
	Other Med/Surg	2	-0.003	0.003	-0.017†	0.004	-0.021†	0.002	-0.006	-0.003	-0.009†	-0.012†	-0.011†	-0.009	-0.006*	-0.003
12324	Lenses Intraocular	3	0.001	0.051**	-0.009*	0.081**	-0.009	0.013	-0.008	0.001	0.04	-0.007	0.061	0.002	-0.002	-0.008
12913	Pacemakers Cardiac Implantable	3	0.009*	0.035†	-0.028**	0.038†	-0.028†	0.032†	-0.028	0.009*	-0.001	-0.009	-0.002	-0.005	-0.003	-0.005
15766	Orthopedic Internal Fixation Systems Spinal	3	-0.017*	0.003	-0.034**	0.015	-0.029	-0.005	-0.041*	-0.017*	-0.026*	-0.02	-0.031	-0.040**	-0.02	-0.027
15870	Prostheses Cardiac Valve Biological	3	-0.002	0.008	-0.015	0.011*	-0.021	0.005	0.003	-0.002	-0.003	-0.012	-0.005	0.004	-0.002	-0.014*
16040	Stents Ureteral	3	0.000	0.005	-0.020*	0.006	-0.023*	0.003	-0.015	0.000	-0.002	-0.016	0.005	-0.011	-0.039	-0.019
16084	Prostheses Joint Hip Acetabular Component	3	-0.006	0.044†	-0.086†	0.047†	-0.072†	0.042†	-0.102†	-0.006	0.001	-0.079†	0.006	-0.037**	-0.004	-0.108†
16095	Prostheses Joint Hip Femoral Component	3	0.004	0.047†	-0.065†	0.037†	-0.080†	0.053†	-0.049**	0.004	0.003	-0.057†	0.001	-0.038**	0.003	-0.048**
16097	Prostheses Joint Knee Femoral Component	3	0.003	0.045†	-0.056†	0.059†	-0.055†	0.037†	-0.058†	0.003	0.012*	-0.052†	0.015	-0.025*	0.011*	-0.048†
16098	Prostheses Joint Knee Tibial Component	3	-0.002	0.041†	-0.079†	0.052†	-0.099†	0.034†	-0.048**	-0.002	0.013**	-0.061†	0.007	-0.074†	0.01	-0.029*
16111	Prostheses Mammary Internal	3	0.003	0.025†	-0.014	0.033†	-0.013	0.015*	-0.016	0.003	0.009	-0.008	0.021†	0.001	-0.005	-0.011
16921	Prostheses Joint Shoulder Humeral Component	3	-0.001	0.001	0.000	0.015	0.019	-0.007	-0.028	-0.001	-0.001	-0.016	-0.003	-0.015	-0.003	-0.038
17165	Allografts	3	0.008	0.043*	-0.033	0.069*	-0.04	0.02	-0.024	0.008	0.03	-0.026	0.069	-0.028	-0.013	-0.003
17241	Stimulators Electrical Spinal Cord Analgesic	3	-0.006	-0.008	-0.007	-0.019	0.003	0.000	-0.027*	-0.006	-0.017†	0.006	-0.002	0.021	-0.018†	-0.039**
18253	Grafts Skin Biological	3	-0.008†	-0.016**	0.003	-0.017†	-0.002	-0.008	0.011	-0.008**	-0.023**	0.007	0.027**	0.005	0.001	0.019†
18504	Defibrillator/Cardioverter/Pacemakers Implantable	3	-0.013*	0.011	-0.042†	0.007	-0.045**	0.013*	-0.028	-0.013**	-0.004	-0.034**	-0.006	-0.046†	0.002	-0.018
20376	Defibrillator/Cardioverter/Pacemakers Implantable Resynchronization	3	-0.002	0.013	-0.017	-0.006	-0.014	0.031	-0.026	-0.002	-0.006	-0.021	-0.008	-0.033**	-0.001	-0.02
20383	Stents Vascular Coronary Balloon-Expandable Drug-Eluting	3	-0.013†	-0.013**	-0.019	-0.015**	-0.005	-0.009	-0.075†	-0.013†	-0.020†	-0.014	-0.022†	-0.009	-0.013*	-0.051†
20422	Stents Vascular Coronary Balloon-Expandable	3	-0.010†	-0.025**	-0.040†	-0.017*	-0.046†	-0.039**	-0.031**	-0.010*	-0.032†	-0.019	-0.032†	-0.015	-0.039†	-0.015
22543	Occluders Vascular Intravascular Embolization Coil	3	0.004	0.008	0.002	0.027†	-0.009	-0.018	0.018	0.004	-0.003	0.002	0.009	-0.007	-0.018	0.022
923150	Orthopedic Fixation Systems Implantable Spine Interbody Fusion	3	-0.013**	0.007	-0.056†	0.015	-0.062†	0.000	-0.047†	-0.013*	-0.025*	-0.043†	-0.031	-0.037*	-0.032*	-0.038†
	PPIs	3	-0.005†	0.014†	-0.033†	0.014†	-0.033†	0.014†	-0.039†	-0.005**	-0.007**	-0.026†	-0.006*	-0.024†	-0.006*	-0.030†

Authors' calculations from PriceGuide data, 2009-2014. Specifications include hospital-brand and brand-year fixed effects; alternative fixed effects in Table A11. Left panel: estimates from pooled regression of all price quintiles together. Right panel: estimates from separate regressions of each price quintile. Standard errors clustered at hospital-brand level. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

In sum, we document a robust pattern of significant savings due to transparency among previously high-priced hospital-brands, which are particularly strong for relatively high-quantity hospital-brands. Given the stability of these focal estimates, in the main text, we report the results of our preferred pooled-quintiles specification with hospital-brand and brand-year fixed effects, which tend to have smaller standard errors due to the sparsity of monthly data for some brands in the expanded analysis. We also document some evidence consistent with an externality effect in which prices for low-price hospital-brands increase under transparency; these results are quite sensitive to specification, however.

F   Usage Pattern Changes and Demand

Stents and certain other expensive medical technologies are "physician preference items" where physician demand is based on strong preferences and is relatively insensitive to price – Grennan (2013) estimates own-elasticities centered around -0.32 for bare metal stents and -0.52 for drug-eluting stents. In general, however, the price benchmarking information treatment could influence quantities as well. Here, we perform a set of analyses to provide a check of this hypothesis and also provide proof of concept for how this analysis might be incorporated in the case of products where demand is more sensitive to price.

There are two primary ways in which quantities might adjust to benchmarking price information and subsequent renegotiations: (1) In a context where contracts specify quantities or market shares in addition to price, renegotiations to obtain better prices might also involve large quantity or share commitments—this effect was tested and ruled out in our analysis in Appendix C. (2) In a context where quantity is responsive to price, negotiation of better prices would lead to increased usage in the brands with the largest relative price decreases. We analyze this second case here.

We run the regression specifications allowing for heterogeneous treatment effects of information depending on pre-join prices and quantities, but here with quantity q_{hjt} as the dependent variable (results are qualitatively similar and so unreported for market shares and log transformations):

$$
\begin{aligned}
Q_{hjt} \;=\; & \beta_{quintile}^{Info} * \mathbb{1}_{\{post_{hjt}\}} * \mathbb{1}_{\{quintile_{hj,pre}\}} \\
& + \theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j}) + \varepsilon_{hjt}
\end{aligned}
$$

where $\beta_{quintile}^{Info}$ estimates the treatment effect, by price quintile. The results are shown in Table A13.

If quantity were responsive to price (with downward sloping demand), then we would expect quantity/share increases in exactly the areas we see relative price decreases. Because information leads to decreases in prices for brands in the high-price, high-quantity part of the pre-information distribution, we should expect potential quantity increases for those brands and decreases for other brands (whose prices haven't changed, but have become higher relative to the brands with price decreases). This is not the case in Table A13, where some specifications indicate that quantities may be changing slightly post join, but no specification shows that these changes vary significantly (economically or statistically) *as one moves across the pre-join price distribution.*

Version	Pre-info price quintiles ($\beta^{Info}_{quintile}$)				
	1	2	3	4	5
$\theta_h + \theta_j + \theta_t + \gamma_j * (t - t_{min_j})$	0.2	1.4	1.3	0.7	2.4**
	(1.5)	(1.9)	(1.7)	(1.7)	(1.1)
$\theta_h + \theta_{jt}$	−0.1	2.0	1.6	1.0	2.9**
	(1.6)	(2.0)	(1.9)	(1.8)	(1.3)
$\theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j})$	3.2†	2.7	3.8†	2.9**	3.0†
	(1.3)	(1.7)	(1.3)	(1.3)	(1.1)
$\theta_{hj} + \theta_{jt}$	4.0†	3.9*	5.2†	4.2†	5.0†
	(1.5)	(2.0)	(1.5)	(1.5)	(1.3)

Authors' calculations from PriceGuide data, 2009-2014. $N = 32,453$ member-brand-months. Includes 508 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript (†) indicates significant difference from zero at the 1% level; (**) at the 5% level; (*) at the 10% level.

Table A13: Treatment Effect *on Quantity* Estimates Across the Price Distribution

F.1 Censoring

The data used in this study is based on purchase orders issued by hospitals, so the data are "censored" in the sense that we only observe a price when a transaction takes place for a given brand-hospital-month. In our final analyses, we investigate whether this censoring could be a potential source of bias – e.g., if is correlated with the pricing process in a way that affects the interpretation of our results.

There are three distinct cases of these censored ($q_{hjt} = 0$, $p_{hjt} = .$) observations: (1) brands the hospital never contracts with; (2) brands a hospital contracts with at some point, but not the full time the brand is available; (3) brands that are currently contracted with, but not purchased in a given month due to sampling variation.

Cases (1) and (2) relate to the issue of whether hospitals contract with the "full support" of manufacturers, which is analyzed in Appendix C as a part of our analysis of potential standardization and share-based contracts. There, we document that almost half of hospitals contract for the "full support" set of manufacturers/brands (among the set of larger hospitals where sampling variation is less likely a problem). Of course, one cannot observe prices for the censored cases, so there is no way to compare prices for censored vs. uncensored observations. However, we can test for patterns in prices across hospitals that contract more vs. less with a given manufacturer. There we find no economically or statistically significant evidence that prices vary when hospitals contract exclusively with or devote more of their market share to a manufacturer (or brand).This holds in specifications across hospitals and within hospitals (and hospital-brand pairs). That we find no relationship on this margin we

can observe provides us some comfort regarding the margin we cannot.

Our quantity analyses thus far provide suggestive evidence on cases (2) and (3): if quantity were responsive to price (with downward sloping demand), then we would expect quantity/share increases in exactly the areas we see relative price decreases. This was not the case in Table A13. However, we find it instructive to dig deeper on points (2) and (3) by examining censoring explicitly: we examine the reduced form association between each hospital's position in the pre-join price and quantity distributions and the degree of censoring before and after join. That is, for each hospital-brand for which we were able to determine relative pre-join price and quantity position, we generated a panel of all months in which that hospital and brand existed in the database, with a dummy indicating whether we observed nonzero purchase for that hospital-brand.

The results are shown in Table A14. First, we note that, although the quintiles of the price distribution are mutually exclusive, this analysis only includes hospital-brands that were observed pre-join – thus, we generally observe fewer nonzero purchases (negative coefficients) for all hospital-brands in the post-join period as hospitals switch to newly entering brands. Second, the largest reduced form effects are in the middle of the pre-join price distribution: data are not differentially censored post-join for brands that were found to have relatively high prices. To put these numbers in perspective, consider the largest coefficient: -0.15 for the middle quintile hospital-brands in the preferred specification. Post-join, hospital-brands in the middle of the pre-join price distribution were 15 percentage points more likely to be censored; overall, observations for that quintile have nonzero purchase 49 percent of the time, and full discontinuation of all usage of those brands in the post-join period would decrease the nonzero purchase rate to 18 percent. That is, full discontinuation would more than double the censoring rate we observe. More to the point, we do not read too much into the average rate of censoring/discontinuation these results reveal, but rather focus on the fact that high-priced hospital-brands exhibit slower censoring post-join than mid-range hospital-brands.

Version	Pre-info price quintiles ($\beta_{quintile}^{Info}$)				
	1	2	3	4	5
$\theta_h + \theta_j + \theta_t + \gamma_j * (t - t_{min_j})$	-0.06	-0.01	$-0.11^{\dagger}$	$-0.11^{\dagger}$	-0.07^*
	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
$\theta_h + \theta_{jt}$	-0.07^*	-0.00	-0.07^{**}	$-0.12^{\dagger}$	-0.04
	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)
$\theta_{hj} + \theta_t + \gamma_j * (t - t_{min_j})$	-0.03	0.01	$-0.15^{\dagger}$	$-0.10^{\dagger}$	-0.03
	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
$\theta_{hj} + \theta_{jt}$	-0.04	0.00	$-0.10^{\dagger}$	$-0.10^{\dagger}$	0.00
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)

Authors' calculations from PriceGuide data, 2009-2014. $N = 30,183$ member-brand-months. Identification focuses on timing of join; sample contains 331 members. Version 1 includes hospital, brand, and month fixed effects, plus linear brand-specific time trends. Version 2 includes hospital and brand-month fixed effects. Version 3 includes hospital-brand and month fixed effects, plus linear brand-specific time trends. Version 4 includes hospital-brand and brand-month fixed effects. Figure shows Version 3 results. Standard errors clustered at hospital (Versions 1 and 2) or hospital-brand (Versions 3 and 4) level shown in parentheses. Superscript ($\dagger$) indicates significant difference from zero at the 1% level; ($**$) at the 5% level; ($*$) at the 10% level.

Table A14: Reduced Form Effect of Join on Censoring