Diagnosing Price Dispersion

Matthew Grennan[*]
University of Pennsylvania, The Wharton School & NBER
grennan@wharton.upenn.edu

Ashley Swanson
University of Pennsylvania, The Wharton School & NBER
aswans@wharton.upenn.edu

December 4, 2018

Abstract

Using detailed purchase order data for a large sample of US hospitals 2009-15, we document large price dispersion across hospital buyers for identical products in a variety of medical supply markets. We develop a model and identification strategy to determine the extent to which this dispersion – and the market power that supports it – is determined by brand preferences, search/contracting costs, and bargaining abilities. Estimates suggest that large markups are primarily driven by lack of price sensitivity among health care providers in their product usage decisions. Hospital administrator bargaining ability varies widely across hospitals, driving most of the observed price dispersion across hospitals. These results suggest that moves to reduce buyer search and contracting costs are unlikely to have much impact in decreasing margins or price dispersion, unless they can help to solve provider and administrator agency problems.

[*]We would like to thank Sebastian Fleitas, Kate Ho, Robin Lee, Aviv Nevo, Tobias Salz, Ali Yurukoglu, and audiences at Carnegie Mellon, McGill, Penn, Stanford, and the Federal Trade Commission for helpful feedback. The data used in this paper were generously provided, in part, by the ECRI Institute (www.ecri.org). We gratefully acknowledge financial support from the Wharton Public Policy Initiative and NSF Award 1559485. Stuart Craig, Gi Heung Kim, Changhwa Lee, and Donato Onorato provided excellent research assistance. Any errors are our own.

1 Introduction

The Law of One Price fails to hold in product areas as diverse as coal (Stigler 1961), automobiles (Goldberg and Verboven 2001), pharmaceuticals (Sorensen 2000; Starc and Swanson 2018), mutual funds (Hortacsu and Syverson 2004), video games (Dinerstein et al. 2018), and Chinese footwear (Roberts et al. 2016). This failure is important per se in business-to-business markets due to its downstream antitrust implications, and it is more broadly an indicator of large degrees of market power and thus inefficiency. The potential explanations for this phenomenon are numerous. On the demand side, preferences over brands, bundles of characteristics, safety, and price may vary across buyers. On the supply side, costs of production and distribution may vary across supplier-buyer pairs; in cases where prices are negotiated for each pair, heterogeneity in contract structure and relative bargaining weights may also impact prices. Finally, contracting frictions – encompassing the process of searching for suppliers and acquiring the information and cooperation needed to develop a new buyer-supplier relationship – may limit the sets over which prices are negotiated and usage decisions are made. The relative importance of demand-side factors, supply-side factors, and search/contracting frictions likely varies across markets, implying that the welfare effects of price dispersion – and any policy solutions focused on addressing price dispersion – will also depend on the relative contribution of these factors.

In this paper, we examine price dispersion for a wide variety of hospital supply markets. Hospital supplies, including medical devices, are estimated to account for 24 percent of the dramatic growth in inpatient hospital costs between 2001 and 2006 (Maeda et al. 2012), and there is substantial variation in prices of inputs across hospitals. For the top hospital supplies in our data, the average standard deviation of prices across hospitals for the same exact brand-month ranges from 6-31 percent of the mean price. This is nearly as high as the coefficient of variation estimated for common procedure prices paid to hospitals in different hospital referral regions in Cooper et al. (2018), which could be driven in part by input price variation; it is also at the top of the range of coefficients of variation observed in consumer goods markets.[1]

The product markets we consider vary in several important dimensions: in the complexity of their underlying technologies (e.g., gloves vs. implantable cardiac rhythm management devices), the strength of brand preferences, the heterogeneity (perceived or real) in safety and quality, the supplier concentration, the distribution models typically employed, and the relative importance of intermediaries in contracting. The top supply categories in our

[1]E.g., Scholten and Smith (2002) report dispersion measures of 1.6 percent to 20.7 percent for a variety of retail consumer goods such as cameras, batteries, contact lens solution, and lettuce.

data include product categories that are typically known as "physician preference items" or PPIs (e.g., cardiac and orthopedic implants) as well as other medical/surgical product categories that are more commodity-like (e.g., gloves). Conventional wisdom in the industry is that for PPIs, usage is driven by brand-loyal and price-insensitive physicians choosing which product to use to treat a given patient; for more commodity-like products, demand depends to a greater extent on hospitals' procurement and stocking decisions as a function of product characteristics and price. Hospitals typically rely on the services of group purchasing organizations (GPOs) to negotiate contracts for many products, but GPO prices are used only as a starting point for direct hospital-manufacturer negotiations for certain purchases, including physician preference items (Schneller 2009). In the data, though, we document large amounts of price dispersion across nearly every market, PPI or not.

Given the above variation, we present a model that allows for heterogeneity in preferences, relative bargaining weights, and search/contracting costs across buyer-supplier pairs.[2] Rather than imposing a particular parametric model of search behavior, we recover search cost estimates using a moment inequalities approach based on weak assumptions regarding the optimality of the observed set of products from which each hospital purchases. We estimate the model using rich new data on the top non-capital, non-pharmaceutical hospital supply purchases for a large sample of hospitals between 2009 and 2015. The context that makes the data available is that sample hospitals joined a benchmarking database for the explicit purpose of reducing their supply costs, which provides useful variation in prices to help identify demand across a wide variety of product categories. We estimate our model separately within each product category. We then use the estimates to explore the relative importance of supply and demand factors and information frictions in determining the degree of observed price dispersion – and the market power that supports the dispersion – in dramatically varying markets.

The primary challenge to estimating an empirical model that simultaneously allows for search costs, preference heterogeneity, and bargaining heterogeneity is to separately identify these different mechanisms. This issue is easily seen by considering each separate pair of factors individually. First, the presence of both search costs and preference heterogeneity introduces an identification problem similar to the familiar selection problem in the labor economics literature (beginning with Heckman 1979) and to the problem of "selection on moral hazard" in the insurance economics literature (e.g., Einav et al. 2013): the unobservable shocks in the demand equation may be correlated with the process that generated the

[2] Throughout this paper, we use the term "search costs" to refer to any costs of adding a brand to a hospital's consideration set. These could include search costs in the literal sense, as well as contracting frictions that may or may not be mediated by the presence of contracting intermediaries such as GPOs.

set of products under consideration. This may lead researchers to obtain biased estimates of demand parameters based on quantities purchased within the observed (endogenous) consideration set.

Second, the presence of both search costs and bargaining in a market may lead to bias in models that only account for one or the other. Allen et al. (2013, 2018) show in the Canadian mortgage market that the presence of high search costs (and/or brand loyalty) for some consumers reduces their bargaining leverage and mutes the effect of supply-side concentration on prices. In our context the effect of any source of market power, concentration or search, is further muted by the fact that hospitals exercise their monopsony power and negotiate prices that are even lower than competition and their outside options alone might suggest. A number of researchers have shown the importance of modeling the bargaining stage in models with differentiated products demand and negotiated prices.[3] Thus, the central issue we must contend with in our empirical analysis is that the relevant consideration set for the buyer (and competitor set for suppliers) is a function of search and bargaining and preference parameters, and each of these may be buyer-supplier specific.

Our identification approach proceeds as follows. First, we contend with the endogeneity of the consideration set using a similar logic as that in Hausman (1996) and Nevo (2001) – in those studies, prices of the same good in other markets are used as cost shifters. In the current study, exposure variables capturing interactions with the same vendor in other, unrelated product categories are used as search cost shifters. Our exclusion restriction is that exposure to vendors in unrelated product categories impacts the formation of the consideration set, but does not reflect correlated preferences over vendors across unrelated product categories, conditional on the consideration set and controls. The institutional details underlying this strategy are discussed in Section 2.2.

We jointly estimate differentiated products demand using observed brand shares within each hospital-year consideration set, and marginal costs and bargaining parameters in a standard Nash-in-Nash bargaining framework. Estimation of preferences relies on rich and plausibly exogenous variation in consideration sets over time; estimation of price sensitivity relies on price shocks occurring when hospitals subscribed to a benchmarking database, as documented in Grennan and Swanson (2018). Relative bargaining abilities for each brand-hospital pair are identified by the extent to which a brand's price changes as the added value of the brand to the hospital changes.

[3]Both Grennan (2013) in the context of hospitals negotiating for coronary stents (a PPI), and also Gowrisankaran et al. (2015) in the context of negotiations between hospitals and insurers, show that inelastic demand from end-users (insurance enrollees facing limited out-of-pocket price variation) would imply negative marginal costs when prices are modeled as the outcome of a Bertrand pricing game, but the estimates from a bargaining model imply more reasonable marginal costs.

Finally, we use observed consideration sets and demand and supply parameter estimates to infer search costs. Search costs rationalize which brands are included in vs. excluded from hospitals' consideration sets, given supply and demand parameters. Note that, since Goeree (2008) first formalized the modeling and estimation of consideration sets, most work on search has been done in two settings: (1) online, where detailed search data and information on the consideration set are available; and (2) in wage bargaining, where only equilibrium matches are available. Our setting is somewhere in between. We do not observe buyer search in the way that online studies observe it (e.g., Dinerstein et al. (2018) observe actual browsing data from eBay); however, we do leverage rich panel variation in demand realizations and consideration sets over time, which provides similar identifying information. Our moment inequalities approach identifies search costs using simple conditions: e.g., that if a brand is in the consideration set, it must have been optimal to search *at some point* (when brands are substitutes, the highest possible added value is relative to the outside good); and that if a brand is not in the consideration set, it must not have been worth searching (again when brands are substitutes, the lowest possible added value is relative to the full consideration set).

The demand and bargaining model estimates illustrate several interesting and intuitive patterns. First, while demand for all products is fairly price-insensitive, demand for physician preference items is several orders of magnitude less price-sensitive than demand for more commoditized, non-PPI medical/surgical products. Second, supplier firms tend to capture a greater portion of the total surplus generated by hospital supply negotiations for PPIs than for other products. These patterns are consistent with the widely-held belief that PPIs are overpriced due to misalignment between the objectives of hospitals (i.e., the purchasers) and physicians (i.e., the users) (Robinson 2008).

The search cost estimates suggest that contracting frictions are on the order of 10 percent of price, on average. This conceals quite a bit of heterogeneity, though. In a few categories, the breadth of the buyer-supplier relationship, measured by the percent of hospital spend with that vendor in *other* product categories, can decrease search costs by 5-30 percent.

With the estimated demand and bargaining models, we are able to make progress in decomposing the price variation across hospitals, and also understanding the sources of market power that make this variation possible. We do this by computing a series of counterfactual price equilibria where we exogenously vary choice set size and alternately shut down heterogeneity across hospitals in demand and bargaining. Holding choice set size fixed, we estimate that price variation across hospitals is driven primarily by bargaining (especially in commodities); but demand heterogeneity also accounts for substantial price variation. This is important for several reasons. First, while research suggests that hospitals have made substantial improvements in purchasing of commodities since the introduction of the prospective payment system, these results indicate that there is still considerable heterogeneity driven by relative bargaining power. This bargaining power variation could in turn capture heterogeneity in management (Bloom et al. 2014), information (Grennan and Swanson 2018), or something else (Lewis and Pflum 2015). Second, the results confirm that preference heterogeneity generates significant market power for PPIs – the welfare impact of this preference heterogeneity will depend on whether it is driven by true product differentiation in quality vs. physician-specific brand preferences generated by marketing.

Considering these same counterfactual price equilibria as choice set sizes increase sheds light on the role of search/contracting costs in generating markups, and the interactions of search with demand and bargaining. As choice set sizes increase, average prices decrease, but relatively little. The largest gains from lowering contracting frictions appear to be the expected consumer surplus gains from increasing choice, not increasing price competition.

This paper contributes to the growing literature on empirical models of negotiated price markets based on "Nash-in-Nash" bargaining (Crawford and Yurukoglu 2012; Grennan 2013, 2014; Gowrisankaran et al. 2015; Ho and Lee 2017). Whereas prior studies have typically taken the set of buyer-supplier relationships as given, we consider the role of frictions in the search/contracting process and how it interacts with demand and bargaining to generate markups and price dispersion across buyers. In this aspect, our study is most closely related to the Allen et al. (2018) study of search and negotiation of mortgage quotes and the Salz (2017) study of waste management contracts, but our modeling differs substantially due to the different data and institutional contexts. In particular, prices in our model are determined by "Nash-in-Nash" bargaining, whereas theirs are the result of auctions.

Our approach to search, using moment inequalities based on stability conditions, is also related in spirit to Ghili (2018), but differs in that our stability conditions are built to be weaker in the sense that they are potentially consistent with a wide variety of processes that have been considered the the industrial organization search literature (Sorensen 2003; Hortacsu and Syverson 2004; Hong and Shum 2006; Honka 2014). While limiting the counterfactuals that can be considered, our approach provides one path forward for cases like ours where buyers contract with a set of differentiated substitute suppliers for needs that are realized over time, a common scenario in business-to-business markets.

A final noteworthy contribution of this paper is the large number and variety of product markets we are able to analyze. The study adds a substantial number of markets to the body of evidence on price dispersion (Scholten and Smith 2002; Kaplan and Menzio 2015) in addition to the growing numbers of empirical case studies in bargaining and search just mentioned.

2 Data and Background on Hospital Purchasing

Health care in the hospital setting has high fixed capital costs in the form of facilities and equipment, but it also has high variable costs in the form of skilled labor, pharmaceuticals, and consumable supplies such as implantable medical devices. The price dispersion we document in this setting is particularly notable because, in the short run, hospitals are typically reimbursed a fixed amount by private or public insurers for the services they provide. Thus, consumable prices come directly from the hospital's bottom line, and the supplier in our data make up 24 percent of hospital operating costs (Craig et al. 2018). In this Section, we provide some background on how consumable supplies are used and purchased, and we describe the unique data set and research setting that allow us to analyze the determinants of price dispersion.

2.1 Hospital Purchase Order Data

The primary data used in this study come from a unique database of all consumable supply purchases made by a large number of US hospitals during the period 2009-2015. The data are from the PriceGuide™ benchmarking service (hereafter, "PriceGuide data") offered by the ECRI Institute, a non-profit health care research organization. For each transaction, we observe price, quantity, transaction month, and supplier for a wide range of product categories.

The reported price and quantity data are of high quality because they are typically transmitted as a direct extract from a hospital's materials management database. The PriceGuide™ benchmarking service compares each hospital's submitted data to that of others in the database and generates several analyses of the hospital's savings opportunities; thus, the hospitals have strong incentives to report prices accurately. Related to its materials management origins, the data is at the stock-keeping-unit (SKU) level, requiring us to use machine learning algorithms to group SKUs that belong to the same manufacturer-brand.[4] For stents and surgical staplers, we also validate our algorithms against data collected from manufacturer catalogs and find that our machine learning algorithm performs well in identifying brands. See Appendix A for details.

[4]The goal of the machine learning procedure is to identify the level of product at which hospital-supplier contracts are negotiated. E.g., for stents, prices are negotiated separately for each brand, and each brand subsumes a large number of SKUs. We use the RE-EM tree package to flexibly group SKUs (defined by a set of dummies for all potential alphanumeric characters in each SKU position) into brands based on observed price variation within each manufacturer-vendor combination. In Appendix C, we present results for alternative aggregations of SKUs and find our results qualitatively unchanged.

2.1.1 Representativeness of the benchmarking database sample

The hospitals in the purchase order data voluntarily joined a subscription service that allows them to benchmark their own prices and quantities to those of other hospitals 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. 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. As discussed in detail in Grennan and Swanson (2018), we do not find evidence that hospitals differentially join the database during times where prices are trending up or down relative to market price trends. However, we do note that the PriceGuide members are overrepresented in the American west and underrepresented in the south, and that they are larger than the average US hospital.

2.2 Product Categories

Each transaction in our data includes a product category identifier from the ECRI Institute's Universal Medical Device Nomenclature System (UMDNS).[5] We chose the products for this analysis as those in the top 100 UMDNS codes by total spending; we then removed products that were overly broad (e.g., "office supplies"), those that had missing or inconsistent quantity data, those whose files were too large for us to estimate supply and demand with a reasonable bound on computing resources, and those for which our "exposure" identification strategy (described below) is not sufficiently powerful.[6] After applying these filters, thirty top product categories remain. In many of our analyses, we separately consider product categories by class according to the Food and Drug Administration's classification system. Specifically, we distinguish FDA risk class III from FDA risk classes I-II; for brevity of notation, we refer to class III supplies as physician preference items (PPIs) and to classes I-II as non-PPIs.[7]

[5]A note on terminology: throughout this draft, we use "product category" to refer to the UMDNS grouping included in the transaction files. The UMDNS system is employed to classify any device or supply based on its intended purpose, with some distinctions for mechanism of action. It covers all medical devices and supplies, clinical laboratory equipment and reagents, and selected hospital furniture, among other items. For example, drug-eluting coronary stents have UMDNS code 20383. For a finer level of distinction, 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. Finally, for a coarser level of distinction, we use "product class" to refer to broad groupings of product categories: physician preference items (PPIs) such as drug-eluting coronary stents, and non-PPIs such as surgical gloves.

[6]See Appendix A for the lists of product categories removed at each step.

[7]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

The top panel of Table 1 summarizes our data for non-PPIs; product categories are listed in decreasing order of average yearly spend.[8] Some of these products are roughly commodities, such as surgical gloves: products that can be used in a hospital setting by staff members with a variety of roles and scopes of practice. Conditional on a few characteristics, such as material, we do not expect particular manufacturers or products to be strongly preferred.[9] Non-PPIs also include other medical and surgical items that may be used in moderately invasive procedures, but may or may not be associated with strong brand preferences; examples include catheters and bone grafts. Non-PPIs are a quite heterogeneous category: as shown in Table 1, these product categories vary in popularity, and spending per hospital-year varies from \$16 thousand to \$470 thousand. Similarly, price per unit varies from \$0.38 (linen underpads) to \$1,934 (bone grafts). Non-PPIs are purchased by 441 sample hospitals on average.[10] The variation in negotiated price across hospitals depends on the UMDNS code under consideration: for surgical gloves, the coefficient of variation across hospitals, within brand-time, is only 6 percent, while for batteries, the CV is 23 percent. There are many brands to choose from, but hospitals may have limited awareness of the total set of brands (and corresponding prices) available: the average hospital only purchases $|\mathcal{J}_h| = 7$ of the 144 unique brands available for the average non-PPI UMDNS code. Even more notably, it is usually the case that the most popular brand j^* is not purchased by the majority of hospitals: on average, the probability that j^* is in hospital h's consideration set $\mathcal{J}_h$ is only 0.34. For non-PPI product categories, particularly the more commodity-like categories, this is unlikely to be driven by preference heterogeneity; it seems much more likely that supply factors, such as contracting frictions, drive this variation.

PPIs are summarized in the bottom panel of Table 1. Examples include coronary stents, pacemakers, and prosthetic hip joints. For physician preference items, usage is driven by strong brand preferences of physicians, often surgeons, choosing which product to use to treat a given patient. These strong PPI brand preferences are frequently noted in policy

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.

[8] Table 1 summarizes data for final analytic sample. See Appendix A.1 for details regarding data cleaning.

[9] In the current analyses, many commodities are dropped due to inconsistencies across hospitals in how quantities are reported. Exam gloves, one of the most popular product categories in our data, are dropped for this reason.

[10] The full dataset contains 1,228 hospitals, but we restrict analysis in each product category to hospitals purchasing the given product category in significant volumes, those for which we observe the date they joined the benchmarking service, and those we were able to match to external hospital characteristics. To perform the analysis in the current study, we obtained permission to contract a trusted third-party to match facilities in the PriceGuide data to outside data on hospital characteristics from the American Hospital Association (AHA) annual surveys. The third-party then provided us with access to the merged data for analysis, with hospital-identifiable information removed. See Appendix A.1 for details.

discussions. PPIs tend to be expensive, high-tech products used in specific, advanced procedures – often cardiac and orthopedic procedures – and are thus not necessarily purchased by all hospitals: only 324 sample hospitals purchased the average PPI. At the extreme, only 232 sample hospitals purchased hemostatic media. The average hospital purchasing PPIs spends \$383 thousand per year on each PPI UMDNS code. The coefficients of variation for PPIs are in a similar range as those observed for non-PPIs, but given how expensive PPIs are (\$1,951 per unit on average), the dollar values of the price variations observed across hospitals are more extreme. These product categories can often only be purchased directly from one of a few manufacturers, and manufacturers exert substantial effort in marketing their brands to hospitals with relevant patients. It is thus not surprising that we observe relatively more comprehensive consideration sets: the average hospital's consideration set contains $|\mathcal{J}_h| = 10$ brands, of the $|\mathcal{J}| = 95$ available, and the modal brand is in most hospitals' consideration sets ($\overline{\Pr[j^* \in \mathcal{J}_h]} = 0.61$).

Table 1: Summary of Purchasing Categories

	N_h	Annual Spend \$1000s	p		$\lvert\mathcal{J}\rvert$	$\lvert\mathcal{J}_h\rvert$		$Pr[j^* \in \mathcal{J}_h]$	$Pr[j^* = j^*_h]$
			μ	$\frac{\sigma}{\mu}$		μ	$\frac{\sigma}{\mu}$		
Non-PPIs									
Bone Grafts	352	\$470	\$1,934	0.13	121	9	0.47	0.69	0.21
Kyphoplasty Kit	217	\$183	\$1,761	0.11	13	3	0.42	0.89	0.57
Bone Nails	470	\$175	\$1,404	0.15	149	15	0.46	0.77	0.21
Bone Implant Putty	414	\$168	\$1,064	0.15	166	12	0.52	0.36	0.09
Polymeric Mesh	262	\$100	\$921	0.12	242	17	0.51	0.38	0.08
Tissue Fusion Device	160	\$57	\$408	0.06	20	2	0.48	0.35	0.20
Orthopedic Fixation Systems	379	\$117	\$394	0.18	132	17	0.58	0.75	0.36
Cath., Misc	148	\$88	\$361	0.31	36	5	0.46	0.07	0.06
Suture Anchors	422	\$101	\$319	0.11	154	18	0.53	0.62	0.09
Surgical Staplers	591	\$111	\$215	0.14	244	9	0.49	0.20	0.07
Linear Staplers	529	\$81	\$160	0.12	131	6	0.57	0.21	0.12
GI Staples	542	\$118	\$133	0.18	158	6	0.54	0.27	0.16
Batteries	470	\$130	\$93	0.23	71	5	0.61	0.19	0.16
Laproscopic Clip Applier	488	\$62	\$92	0.08	58	3	0.47	0.26	0.16
Pulse Oximeter Probes	304	\$115	\$85	0.15	77	3	0.68	0.14	0.11
Trocars	593	\$60	\$40	0.18	204	7	0.54	0.29	0.11
Pneumatic Compression Cuffs	316	\$95	\$28	0.10	49	3	0.59	0.30	0.26
Liquid Adhesives	599	\$53	\$27	0.09	56	2	0.52	0.35	0.25
Sutures	647	\$16	\$8	0.14	450	11	0.57	0.23	0.11
IV Infusion Pumps	230	\$86	\$5	0.07	25	2	0.52	0.25	0.22
IV Administration Kits	636	\$74	\$4	0.09	292	5	0.61	0.10	0.06
IV Tubing Extensions	625	\$40	\$2	0.09	330	5	0.66	0.09	0.05
Surgical Gloves	664	\$91	\$1	0.06	198	7	0.61	0.29	0.11
Linen Underpads	521	\$30	\$0	0.08	73	2	0.57	0.16	0.10
Average (24)	441	\$109	\$394	0.13	144	7	0.54	0.34	0.16
Physician Preference Items									
Pacemakers	357	\$534	\$4,376	0.11	53	11	0.42	0.90	0.32
Humeral Shoulder Prosth.	264	\$211	\$2,486	0.19	106	15	0.41	0.44	0.12
Drug Eluting Stents	314	\$1,028	\$1,571	0.05	9	4	0.36	0.81	0.40
Allografts	322	\$146	\$1,530	0.12	241	10	0.77	0.16	0.07
Acetabular Hip Prosth.	458	\$265	\$1,418	0.23	143	17	0.56	0.71	0.30
Hemostatic Media	232	\$115	\$324	0.05	15	3	0.46	0.61	0.35
Average (6)	324	\$383	\$1,951	0.13	95	10	0.50	0.61	0.26

2.3 Institutions: Hospital Contracting Environment

Hospitals purchase thousands of product categories, and prices for each of these product categories are determined in negotiation. In determining which brands to contract, purchasers within a hospital must factor in clinical value, safety, cost, and other conditions of sale and service (e.g., for capital equipment, which is not in our data, 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 often used as a starting point for direct hospital-manufacturer negotiations for physician preference items and capital equipment (Schneller 2009).[11]

Physicians play a large role in choosing which consumable supplies are used in medical procedures, particularly for medical devices like PPIs. For this reason, suppliers and their representatives work closely with physicians and hospital support staff in order to promote their brands, provide training for new brands, and even provide on-site technical assistance in the operating suite (Montgomery and Schneller 2007). This implies that suppliers' representatives have highly specialized knowledge about their brands and physician users. As noted above, hospitals typically purchase a small share of the brands available in a given product category. This may be due to heterogeneity in preferences: we would expect a hospital to be more likely to contract a given brand and thus include it in its consideration set for use in procedures if it is particularly preferred by the hospital's affiliated physicians. This phenomenon may also be driven by supply factors that vary across hospital-vendor pairs, such as distribution costs and search/contracting frictions.

In our demand analysis, we use this logic to motivate an identification strategy based on supply factors that impact multiple dissimilar product categories. Supply factors at the hospital-vendor pair level may generate correlations in the composition of hospitals' consideration sets across product categories with different staff users, therapeutic indications, and levels of technological sophistication. For example, geographic proximity and distribution costs may lead to some hospitals being more likely to have contracts for both Medtronic tracheal tubes and Medtronic stents. Importantly, we argue that, conditional on appropriate controls, supply side factors are likely to generate correlations across dissimilar product categories in terms of which brands enter the *consideration set* (i.e., which brands are in the storeroom), but that the siloed and specialized nature of medical device sales and training implies that *user* (i.e., physician) preferences over particular brands are likely uncorrelated

[11] A GAO report from 2003 noted that there are hundreds of GPOs, some of which operate regionally; however, at the time of the study, seven national GPOs with purchasing volumes over $1 billion accounted for more than 85 percent of all hospital purchases nationwide made through GPO contracts.

across dissimilar categories, conditional on consideration sets.[12]

The UMDNS system imposes a hierarchy over product categories based on their intended use. For example, Figure 1 below displays part of this hierarchy as a tree structure with up to seven splits leading to coronary drug-eluting stents. Coronary drug-eluting stents (in pink) are directly under the parent coronary balloon-expandable stents, less directly under the parent non-active implantable devices. Polymeric mesh and bone grafts (in green) are also non-active implantable devices, while cardiac pacemakers (in yellow) are instead active implantable devices. The most distant product category from coronary drug-eluting stents in this example is tracheal tubes (in yellow), which are not under the parent "Implants and Prosthesis." In our empirical analysis, we will assume that, conditional on having contracts for both Medtronic tracheal tubes and Medtronic stents, users in those product categories' verticals will not have correlated preferences regarding the Medtronic brands at the point of usage.[13]

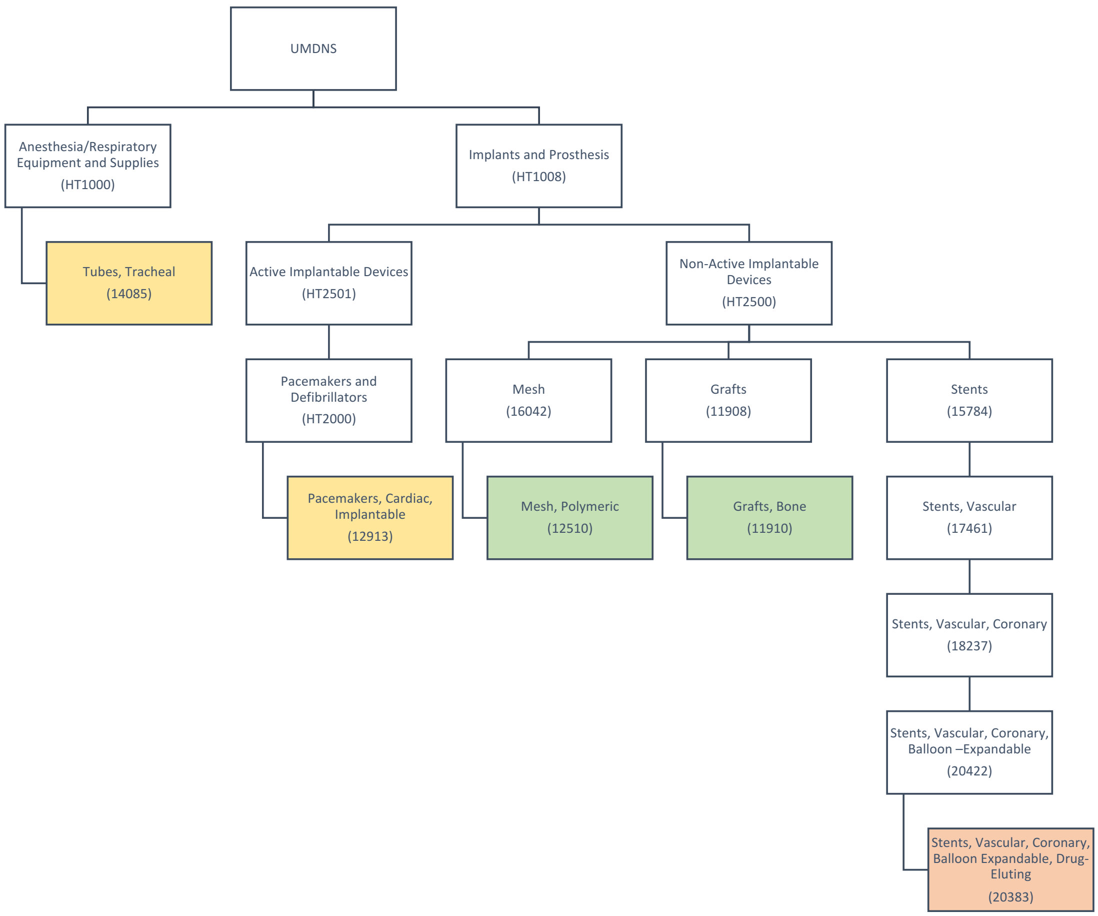
We use the above hierarchy to flag each pair of UMDNS codes (among the top 200 in overall spending) as likely similar ("near") or dissimilar ("far"). A likely near pair is one with more than one shared parent in the UMDNS tree structure. In the above example, tracheal tubes and cardiac pacemakers are classified as far from coronary drug-eluting stents, while polymeric mesh and bone grafts are classified as near. Starting with this set of candidate "far" pairs, we then identified each pair with significant vendor overlap – UMDNS code A and UMDNS code B have "significant vendor overlap" if at least 70 percent of spending in A is contributed by vendors that sell products in B. Finally, among the pairs with significant vendor overlap, we checked whether each product category would typically be used in the same or similar procedures; for example, balloon categories and coronary stents are used in the same procedures; cardiac valve prostheses and annuloplasty rings are used in different procedures for heart valve abnormalities. Our final list of dissimilar pairs of UMDNS codes are those that are "far" according to the UMDNS tree hierarchy, and that either *do not* have significant vendor overlap, or *do* have significant vendor overlap but are not use in the same or similar procedures.

In the identification strategy laid out in Section 3.1.1 below, we use exposure of a given hospital-vendor pair across dissimilar product categories by this definition to address a form of selection bias. Specifically, the brands in a given hospital's contracted consideration set are likely those that are particularly preferred by the hospital's physicians, leading to bias in the estimated preference parameters.

[12]Our exclusion restriction is discussed more precisely in Section 3.1.1 below.

[13]In the remainder of the paper, we use the term "vertical" to denote unique purchasing entities within hospitals.

Figure 1: UMDNS Code Hierarchy



2.3.1 Consideration set formation and pricing

In this business-to-business bargaining setting, we expect the size of consideration sets in a given product category to be a function of users' preferences, search and contracting costs, and the gradient between consideration set size and price. The classic "Nash-in-Nash" used in empirical bargaining studies generates a clear prediction that, if brands are substitutes, the addition of a brand to the set a hospital contracts with ($\mathcal{J}_h$) will weakly decrease the negotiated price of the other brands in the set. This is because an additional substitute reduces the gains from trade for inframarginal brands.

A more nuanced set of predictions can be obtained from models in the more recent empirical bargaining literature that allow firms to strategically employ small consideration sets in order to extract larger discounts from included vendors (e.g., Ho and Lee 2018). This practice is known as "standardization" in the hospital industry, and is thought to be a useful source of savings (Noether and May 2017); however, hard evidence on its ubiquity is scarce, and Grennan and Swanson (2018) found no evidence of it for an important PPI (coronary drug-eluting stents). Models with this type of exclusion may predict that, for a given set of potential suppliers $\mathcal{J}_h^{potential}$, prices will be *increasing* in the size of the set the hospital contracts with $\mathcal{J}_h$ (at the margin, if exclusion is optimally applied).

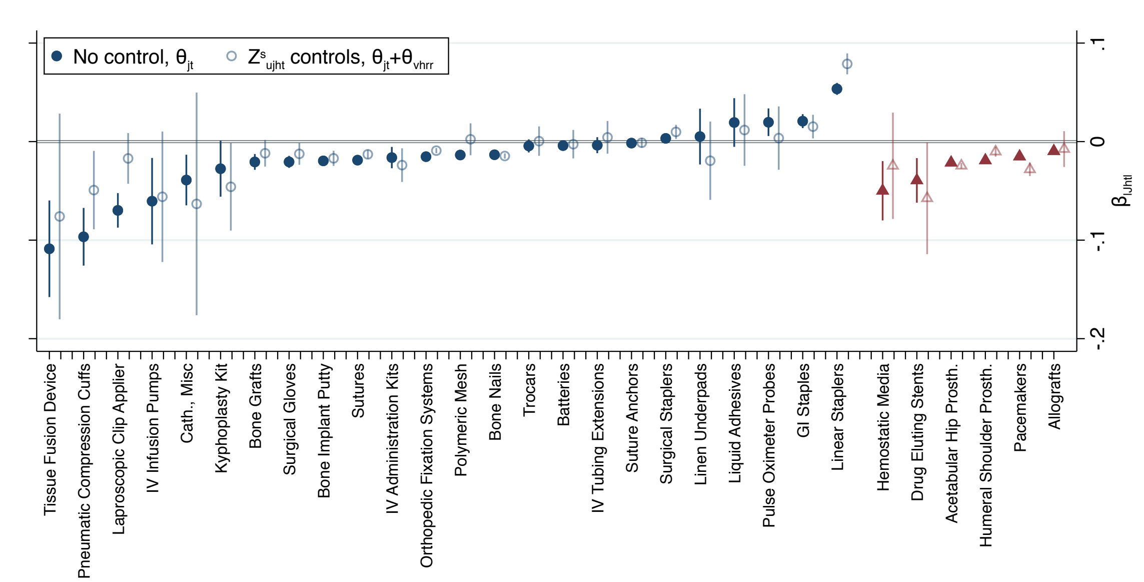
Figure 2 provides suggestive evidence on this point. It shows the results of a regression within each top UMDNS code u:

$$p_{jht} = \theta_{jt} + \beta|\mathcal{J}_{ht}|$$

where, in each regression, hospital-brand-year-level observations are quantity-weighted so that β can be interpreted as the average correlation between consideration set size $|\mathcal{J}_{ht}|$ and the hospital's average price for time t, net of composition effects absorbed in θ_{jt}.

The regression coefficients β are shown as solid markers for each top product category in Figure 2: we display coefficients and 95 percent confidence intervals for non-PPIs in blue, then for PPIs in red. Categories are ordered from left to right in increasing order of β. The coefficients in Figure 2 are in most cases small and negative, indicating a weak negative association between consideration set size and prices.

Figure 2: Correlation between p_{jht} and $|\mathcal{J}_{ht}|$



In an alternative version of this specification, we also control for "exposure" of the hospital to the relevant brand j's vendor in both similar ("near") and dissimilar ("far") product categories, as an informal proxy for hospital-vendor level supply factors (see further discussion of our identification strategy in Section 3.1.1 below). For example, if $\mathcal{D}_u$ is a set of product categories that are dissimilar to category u, then exposure of hospital h to brand

j's vendor in dissimilar categories is defined as:

$$Z^{\mathcal{D}}_{ujht} = \frac{\sum_{u' \in \mathcal{D}_u} E_{u'hv(j)t}}{\sum_{u' \in \mathcal{D}_u} \sum_{v'} E_{u'hv't}}$$

where E_{uvht} is total expenditure by hospital h on vendor v's brands in UMDNS code u at time t.[14] Exposure of hospital h to brand j's vendor in similar categories is defined in the same way, but for $\mathcal{S}_u = \mathcal{U} \setminus \mathcal{D}_u$.

The hollow markers in Figure 2 show the estimated β's controlling for exposure at the hospital-brand level in near and far categories; aggregate exposure of the hospital to all brands purchased $j \in \mathcal{J}_h$ in near and far categories; and hospital and vendor-region fixed effects.[15] These controls have little effect on the pattern of estimated βs: so this slight negative association between consideration set size and price holds within hospitals with similar predicted supply factors such as distribution costs and search/contracting frictions.

Each of these findings is consistent with the Nash-in-Nash model we estimate in Section 3.2, though they cannot exclude alternative models that allow for strategic restrictions of consideration set size (such as Nash-in-Nash with Threat of Replacement (Ho and Lee 2018)). These reduced form patterns may reflect the countervailing forces of standardization and gains from trade, and they do not account for a number of important omitted variables, such as the demand and bargaining parameters we propose to estimate. In Appendix D, we more directly consider the consideration set stability predictions of the Nash-in-Nash with Threat of Replacement model, using our estimated supply and demand parameters and alternative assumptions about hospitals' quality expectations and the level of product at which search takes place. This exercise suggests that observed consideration sets are inconsistent with firms employing strategic exclusion optimally. Thus, in this paper, we rely on the predictions of Nash-in-Nash bargaining to estimate supply side parameters, though we consider the role of standardization to be an important avenue for further research.

3 Structural Model of Heterogeneous Preferences, Negotiated Prices, and Search

We model hospital consideration set formation and demand decisions in a framework with negotiated prices. We assume hospital demand is derived from the preferences of its staff and the needs of its patient population. Those involved in hospital purchasing search for

[14] The denominator is a normalization to account for variation across hospitals in their propensity to purchase brands in dissimilar categories.

[15] "Region" is defined by hospital referral region, of which there are approximately 300 in the US.

brands to add to the consideration set given their beliefs about brand quality and bargaining parameters, and given the cost of adding an additional product to the consideration set. Upon termination of search, the hospital and vendor negotiate prices for each brand in the consideration set. Finally, conditional on the consideration set and negotiated prices, the hospital purchases a brand for each use case from that set according to its demand function. The timing is as follows:

1. Hospital h has ex ante beliefs regarding brand j and time t defined by:

   Preferences $\theta_h + \theta_{jt} + \theta_{v(j),hrr(h)} + \xi^o_{jht} + \xi^u_{jht}$ (unknown $\xi^u_{jht} \sim N(0, \sigma_\xi)$)

   Costs of production/distribution $C_j(q_{jh}; \gamma)$ (known)

   Bargaining $\frac{\beta_{jt}}{\beta_h}\nu_{jht}$ (unknown $\ln(\nu_{jht}) \sim N(0, \sigma_\nu)$)

   Search costs $sc_{jht} = X^{sc}_{jht}\psi$ (known)

2. Hospital consideration set $\mathcal{J}_{ht}$ determined and $\{\xi^u_{jht}\}, \{\nu_{jht}\}$ learned via search/contracting process.

3. Contract prices $p_{jht}(mc_j, \mathcal{J}_{ht}, \theta_{jht}, \beta_{jht})$ determined.

4. Quantities $q_{jht}(\mathcal{J}_{ht}, p_{jht}, \theta_{jht})$ realized.

Below, we describe each step of the search-bargaining-demand model, in reverse order of model timing.

3.1 Demand model

The utility of brand $j \in \mathcal{J} = \{1, ..., J\}$ for use case i (often a doctor/patient combination, for implantable medical devices) at hospital h is

$$u_{ijht} = \delta_{jht} + \varepsilon_{ijht}. \tag{1}$$

The use-specific i.i.d. unobservable $\varepsilon_{ijht} = \epsilon_{ight} + (1 - \lambda_g)\epsilon_{ijht}$ is the random coefficients representation (from Cardell 1997) of the nested logit model where ϵ_{ight} is a random component common to all goods in group g; and ϵ_{ijht} is the standard type I extreme value error term (with scale normalized to one). As a nesting parameter $\lambda_g \in [0, 1]$ approaches 1, there is more substitution among products within group g than to the outside good and other products outside g.

The mean utility across use cases is specified as

$$\delta_{jht} = \theta_h + \theta_{jt} + \theta_{v(j),hrr(h)} - \theta^p p_{jht} + \xi^o_{jht} + \xi^u_{jht}, \tag{2}$$

where $\theta_h + \theta_{jt}$ are hospital and brand-time specific dummy variables and their utility weights, $\theta_{v(j),hrr(h)}$ is a dummy for each vendor-region pair and its utility weight (to account for regional variation in vendors' sales and marketing activity), $\theta^p p_{jht}$ is the negotiated price for pair hj and its utility weight, and $\xi^o_{jht} + \xi^u_{jht}$ is a hospital-brand specific unobservable preference heterogeneity term. The non-standard element here is that this model admits the possibility that the hospital observes a nonzero ξ^o_{jht} *before* the consideration set is formed. We discuss the selection problem this induces and our method for addressing it in the identification and estimation section below.

Given contracts for a set of brands $\mathcal{J}_{ht}$ and flow of choice opportunities M_{ht}, we assume the hospital chooses the brand in the consideration set that maximizes utility for each use opportunity, so that quantities demanded are given by:

$$q_{jht} = M_{ht} Pr[u_{ijht} > u_{ihkt}, \forall k \in \mathcal{J}_{ht}] = M_{ht} \frac{e^{\frac{\delta_{jht}}{1-\lambda_g}}}{\sum_{k\in g} e^{\frac{\delta_{hkt}}{1-\lambda_g}}} \frac{\left(\sum_{k\in g} e^{\frac{\delta_{hkt}}{1-\lambda_g}}\right)^{1-\lambda_g}}{1 + \sum_g \left(\sum_{k\in g} e^{\frac{\delta_{hkt}}{1-\lambda_g}}\right)^{1-\lambda_g}}\ , \tag{3}$$

and hospital surplus across all contracted brands is given by:

$$\pi_h(\mathcal{J}_h) = M_{ht} \frac{1}{\theta^p} \ln \left(1 + \sum_g \left(\sum_{k\in g} e^{\frac{\delta_{hkt}}{1-\lambda_g}}\right)^{1-\lambda_g}\right)\ . \tag{4}$$

3.1.1 Demand identification and estimation

We follow the procedure in Berry (1994), setting choice probabilities implied by the demand model equal to market shares observed in the data, and inverting the system to yield a linear correspondence between a function of market shares and the mean utility for each brand:

$$\ln(s_{jht}/s_{0ht}) - \lambda_g \ln(s_{jht}/s_{ght}) = \delta_{jht} = \theta_h + \theta_{jt} + \theta_{v(j),hrr(h)} - \theta^p p_{jht} + \xi^o_{jht} + \xi^u_{jht}, \tag{5}$$

leading to the linear estimation problem $\ln(s_{jht}/s_{0ht}) = \lambda_g \ln(s_{jht}/s_{ght}) + \theta_h + \theta_{jt} + \theta_{v(j),hrr(h)} - \theta^p p_{jht} + \xi^o_{jht} + \xi^u_{jht}$.

Estimating this equation faces two well-known challenges in that theory suggests both $\ln(s_{jht}/s_{ght})$ and p_{jht} are correlated with the unobservable term $\xi^o_{jht} + \xi^u_{jht}$. We take an instrumental variables approach to solving this identification problem. For $\ln(s_{jht}/s_{ght})$, we follow much of the literature (e.g., Berry et al. 1995; Berry and Waldfogel 1999) in using $(|\mathcal{J}_{ght}|, |\mathcal{J}_{ght}|^2)$ as instruments, which leverages the fact that more variety will on average affect substitution independent of the individual brand's unobservable.[16]

For p_{jht}, we build on the results in Grennan and Swanson (2018) that show how access to benchmarking information generates a price shock that varies across brand-hospital combinations based on their place in the pre-information price and quantity distributions.[17] Specifically, we instrument for price with the full set of interactions between variables indicating: whether benchmarking information is available for that brand at that hospital at that time $1_{\{post_{jht}\}}$; the hospital-brand's quartile of quantity purchased, relative to all other hospital-brands in the year before information is available (e.g., $1_{\{q_{jht}>p75(q)\}}$); and the hospital-brand's quintile of price relative to all other hospitals purchasing the relevant brand in the year before information is available (e.g., $1_{\{p_{jht}>p80(p)\}}$).

We also wish to correct for a specific kind of sample selection problem that could be introduced if the consideration set formation process is somehow correlated with demand unobservables $\xi^o_{jht} + \xi^u_{jht}$. One version of such a problem would be if the hospital observes a nonzero ξ^o_{jht} *before* the consideration set is formed, and search is directed by this information; in this case, the brands in the consideration set may specifically be brands that are preferred (in expectation) by the hospital. This would tend to introduce a positive bias in the estimated fixed effects θ_h and θ_{jt}.

We address this problem by introducing a control function for $E[\xi^o_{jht}|j \in \mathcal{J}_{ht}]$ as in Petrin and Train (2009). Specifically, suppose that the search process can be approximated by the following reduced form:

$$1(j \in \mathcal{J}_{ht}) := \mathbb{1}(Z^s_{jht}\tilde{\phi} + \tilde{\varepsilon}_{jht} > 0) \tag{6}$$

where Z^s_{jht} is a set of instruments that impact search but not demand, and $\tilde{\varepsilon}$ is a shock to the search process which may in general be correlated with the demand observables and unobservables. Our approach then takes the following steps: (1) we estimate the reduced form search model as a probit of $\mathbb{1}(j \in \mathcal{J}_{ht})$ on the instruments Z^s_{jht} and controls; and (2) we include the term $f(Z^s_{jht}\tilde{\phi})$ as a regressor in the demand model, where $f(.)$ is the inverse Mills ratio. This specification of the control function has the intuitive property that it is small when the instruments push a brand into the choice set with probability one. Under the

[16] Instruments like $|\mathcal{J}_{ght}|$ may be endogenous in this setting (and others) due to the consideration set formation step. However, results are similar when we instrument instead using regional patterns of brand availability: $|\mathcal{J}_{g,hrr(h),t}|$.

[17] Grennan and Swanson (2018) found evidence that access to benchmarking leads to price decreases for a variety of product categories, particularly for hospitals and brands that involved high prices and large purchase quantities prior to benchmarking access.

story of search directed based on knowledge of the unobservable ξ^o_{jht}, this generates a clear prediction that the coefficient on the control function will be zero if no selection is present, and it will be positive and correct the fixed effect distribution downward if there is a positive correlation between ξ^o_{jht} and $\tilde{\varepsilon}$.

As discussed in Section 2.3, the unique data on *all* consumable supply purchases across numerous product categories in each hospital, paired with the phenomenon of many vendors supplying product categories in disparate "verticals" across the hospital, provides potential instruments Z^s_{jht} based on the exposure of the hospital to the vendor of brand j in verticals which are unrelated from a demand perspective. The set of instruments we currently employ for this purpose are *vendor-hospital* exposure – the total spend observed for the given hospital on the current brand's vendor *in dissimilar categories* in the same year, relative to *all spend* by that hospital in those dissimilar categories. We define "dissimilar" categories based on the UMDNS code hierarchy previously discussed. Letting $\mathcal{P}_u$ be the set of UMDNS code u's "parents," a category u' is in the "dissimilar" set $\mathcal{D}_u$ if u and u' share at most one parent: $|\mathcal{P}_u \cap \mathcal{P}_{u'}| \leq 1$. In the example hierarchy in Figure 1, the green highlighted categories are "similar" to coronary drug-eluting stents (e.g., bone grafts) and the yellow highlighted categories are "dissimilar" (e.g., tracheal tubes).

Given the dissimilar set $\mathcal{D}_u$, we construct the exposure instrument Z^s_{ujht} as:

$$Z^s_{ujht} = \frac{\sum_{u' \in \mathcal{D}_u} E_{u'hv(j)t}}{\sum_{u' \in \mathcal{D}_u} \sum_{v'} E_{u'hv't}}$$

where E_{uvht} is total expenditure by hospital h on vendor v's brands in UMDNS code u at time t. The denominator is a normalization to account for variation across hospitals in their propensity to purchase brands in dissimilar categories.

To illustrate the power of this identification strategy, Figure 3 shows the results of a regression within each top UMDNS code u:

$$1_{\{q_{ujht}>0\}} = \theta_{jt} + \beta_{near} 1_{\{Z^{s,near}_{jht} > p50(Z^{s,near})\}} + \beta_{far} 1_{\{Z^{s,far}_{jht} > p50(Z^{s,far})\}}.$$

In this specification, $Z^{s,far}_{jht}$ is our focal instrument: the exposure of hospital h to brand j's vendor at time t in *dissimilar* categories. We control for θ_{jt} in order to focus on variation in exposure *within* brand-time. We also control for $Z^{s,near}_{jht}$: exposure of h to j's vendor in *similar* categories; we argue that exposure to vendors in similar categories is more likely than exposure to vendors in dissimilar categories to be driven by correlated preferences across verticals. For each Z^s variable, we examine the effect of above-median exposure (as a rough proxy for "high" exposure) of the hospital-vendor pair in other categories. Thus, our coefficient of interest β_{far} captures the effect of above-median exposure to a given brand's vendor in dissimilar categories on a hospital's tendency to include that brand in its consideration set. The solid markers show the results of the specification when we control only for brand-time fixed effects. The hollow markers show the results of our preferred specification, in which we also include hospital fixed effects and vendor-HRR fixed effects. The former remove variation driven by a given hospital's overall preference for variety; the latter absorb variation driven by regional sales and marketing activity of specific vendors.

Figure 3 plots the estimated coefficients $\hat{\beta}_{far}$ and corresponding 95 percent confidence intervals, with and without hospital and vendor-HRR fixed effects, for each of our top UMDNS codes.[18] In each panel, the top estimates (blue circles) are for non-PPIs; the bottom (red triangles) for PPIs. In panel (a), we show results for the 70 (of the original 100) top spending product categories that were not deemed overly broad (e.g., "office supplies") and did not have missing or inconsistent quantity data (see Appendix A). In panel (b), we show results for the final 30 product categories in our analytic sample; these categories are required to have a powerful exposure first stage in our preferred specification and must also be sufficiently small as to not run up against a memory constraint in estimation.

Within both product classes, we see positive estimated effects of high exposure on brand adoption for the vast majority of product categories. In many cases, these effects are also statistically significant within category in the preferred specification (hollow markers). The effect of including hospital and HRR-vendor fixed effects is usually to attenuate the estimate of β_{far} slightly.

Taken together, these results suggest that exposure to vendors in dissimilar categories has a large positive effect on inclusion of those vendors' brands in the consideration set for the focal category. This identification strategy leads us to the demand model we take to the data:

$$\ln(s_{jht}/s_{0ht}) = \lambda \ln(s_{jht}/s_{ght}) + \theta_h + \theta_{jt} + \theta_{v(j)hrr(h)} - \theta^p p_{jht} + \theta^f f(Z^s_{jht}\hat{\hat{\phi}}) + \xi^u_{jht}, \quad (7)$$

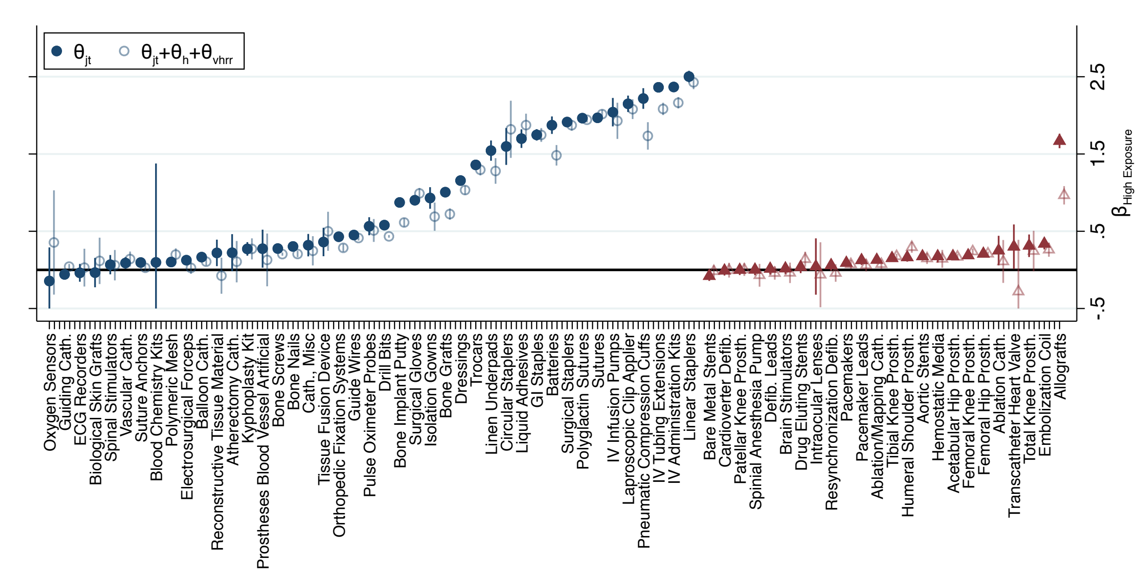
which is a linear instrumental variables specification, where we instrument for price and the nested logit term as described above, $\hat{\hat{\phi}}$ is obtained from a first stage probit of $1_{\{q_{ujht}>0\}}$ on excluded instruments and controls, and $f(\cdot)$ is the inverse Mills ratio.[19] Our preferred

[18] In order to facilitate comparison across widely varying UMDNS codes with different choice set sizes, we normalize each coefficient by dividing through by the mean probability of consideration set inclusion across hospital-brand-years in the UMDNS code.

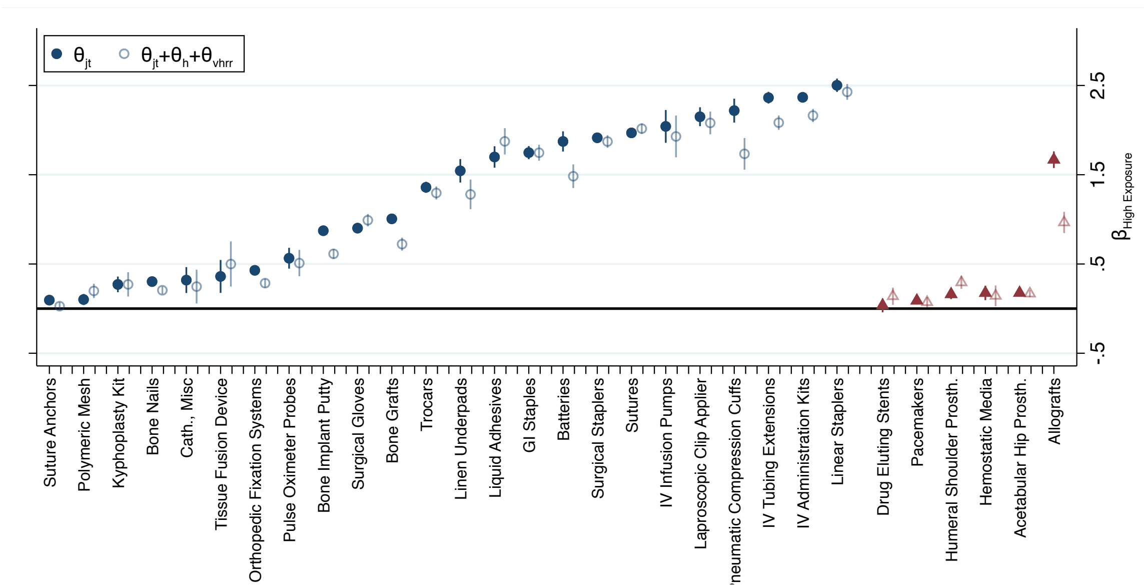
[19] In our current implementation, we include hospital, brand-time, and vendor-HRR fixed effects in the demand regression, but only include hospital and vendor fixed effects in the first stage, as the probit does not handle high-dimensional fixed effects well. In future iterations, we will test alternative specifications of the reduced form first stage.

Figure 3: Reduced Form Evidence of Exposure and Consideration Set Inclusion

(a) All UMDNS Codes



(b) Final UMDNS Codes



specification uses quintiles of $Z^{s,far}$ as our excluded instruments in the "exposure" first stage, and we control for quintiles of $Z^{s,near}$ in both the exposure first stage and in the demand model.

3.2 Supply model of negotiated prices

In the business-to-business market for a given hospital supply, the price for a given brand is buyer-specific. We assume that prices are determined between the hospital and the set of vendors with which it contracts as a Nash Equilibrium of simultaneous bilateral Nash Bargaining problems. Each price maximizes the bilateral Nash product, taking other prices as given:

$$\begin{aligned} p_{hj} &= \arg\max \left(q_{hj}p_{hj} - C_j(q_{hj})\right)^{b_j(h)} \left(\pi_h(\mathcal{J}_h) - \pi_h(\mathcal{J}_h \setminus j)\right)^{b_h(j)} \\ &= \frac{C_j(q_{hj})}{q_{hj}} + \frac{b_j(h)}{b_j(h)+b_h(j)} \left[\left(1 + \frac{\partial q_{hj}}{\partial p_{hj}} \frac{p_{hj} - \frac{C_j(q_{hj})}{q_{hj}}}{q_{hj}}\right) \frac{\pi_h(\mathcal{J}_h) - \pi_h(\mathcal{J}_h \setminus j)}{q_{hj}} + p_{hj} - \frac{C_j(q_{hj})}{q_{hj}} \right] \end{aligned} \quad (8)$$

where $C_j(q_{hj})$ is a function capturing the costs of manufacturing and distributing quantity q_{hj} of brand j to hospital h. The terms $b_j(h)$ and $b_h(j)$ are relative bargaining ability weights that capture the extent to which the optimal price depends on vendor profits vs. the expected additional hospital surplus in the case that a contract is agreed to for brand j: $\pi_h(\mathcal{J}_h) - \pi_h(\mathcal{J}_h \setminus j)$. All the terms in this pricing equation are assumed to be known to all market participants at the time of bargaining.

3.2.1 Supply identification and estimation

We jointly estimate the above (linearized) demand model with the control function correction term and the supply model using a generalized method of moments approach. This enables us to simultaneously recover the demand parameters θ, marginal costs, and mean relative bargaining weights.

We model marginal costs as a fraction of the price charged to the hospital paying the minimum price for each brand. This specification balances the challenges of estimating marginal costs in a market with large markups (Grennan 2013) and the sparsity of product characteristic data within a product category, with the desire to allow for flexible marginal costs across products within a category that may indeed be quite different from one another. We interpret the estimated parameter as the average minimum margin in each product category.[20,21]

$$mc_j := \gamma \min_{ht} p_{jht}. \tag{9}$$

This specification follows the approach in Grennan (2013, 2014) by assuming no unobservable term in marginal costs, and instead loads the supply side unobservable in the bargaining parameters. We prefer this specification as our prior is that, in our empirical setting, marginal costs are less likely to vary across hospitals and time in unobservable ways than bargaining outcomes. Intuitively, relative bargaining abilities for each brand-hospital pair are identified by the slope with which price changes as the added value of the brand changes, and marginal costs are identified as the intercepts in this relationship as added value approaches zero.

We think of the bargaining parameters themselves as representing characteristics of the hospital and brand that enter the negotiation, but are separate from the economic factors that define total surplus: cost, willingness-to-pay, and disagreement points. We model the relative ratio of the two bargaining parameters by:

$$\frac{b_{jt}(h)}{b_{ht}(j)} := e^{\beta_{jt} - \beta_h - 1_{\{Info_{jh}\}} X_{jh}^{pq} \beta_{jh}^{Info,pq} + \nu_{jht}}. \tag{10}$$

Here β_{jt} and β_h represent brand-year and hospital fixed effects, respectively. The term $1_{\{Info_{jh}\}} X_{jh}^{pq} \beta_{jh}^{Info,pq}$ represents whether the hospital has access to benchmarking information on the brand $1_{\{Info_{jh}\}}$ and where the hospital was in the price and quantity distributions relative to other hospitals prior to having that information X_{jh}^{pq}. Our inclusion of the benchmarking information regressors relates directly to our use of this variation in identifying demand, and is intended to capture the finding from Grennan and Swanson (2018) that benchmarking appears to solve an asymmetric information problem in which hospitals may be uncertain about a brand's supplier's negotiator type.[22] Finally, ν_{jht} is the econometric unobservable term for the supply side moments.

In order to recover the supply parameters, we rearrange the supply equation and take

[20] In unreported results, we analyze robustness of this assumption. We have tried estimating models with less flexibility across products, which tend to push marginal costs towards zero and which we believe overstate margins and understate true cost heterogeneity. Models using product characteristics $mc_j = X_j^{mc}\gamma$ seem to work better in the product categories for which they are available, but the potential richness of the relevant characteristics varies widely across categories, and combined with the size of the vendor and product spaces, collecting such data from manufacturer catalogs across all product categories has proven overwhelming. We provide estimates of such a model for selected categories where we have collected such data.

[21] In ongoing work, we are seeking to incorporate potential returns to scale in distribution at the product- and vendor-hospital level.

[22] This seems like the most natural way to map asymmetric information about bargaining parameters into the Nash-in-Nash framework. Providing noncooperative foundations of such a model, however, is not straightforward, and would presumably involve extending the ideas in Collard-Wexler et al. (2017) to asymmetric information bargaining as in Rubinstein (1985).

logs to obtain the following expression:

$$\begin{aligned}\ln(\nu_{jht}) &= -\beta_j + \beta_h + 1_{\{Info_{jh}\}} X^{pq}_{jh} \beta^{Info,pq}_{jh} + \ln(p_{jht} - mc_j) \\ &- \ln\left(\left(1 + \frac{\partial q_{jht}}{\partial p_{jht}} \frac{p_{jht} - mc_j}{q_{jht}}\right) \frac{\pi_h(\mathcal{J}_{ht}) - \pi_h(\mathcal{J}_{ht} \setminus j)}{q_{jht}}\right) .\end{aligned}$$

In this expression, prices p_{jht}, product characteristics $X^{mc}_j (= \min_{ht} p_{jht})$, and demand observables X^d_{jht} enter as data, and we condition out the bargaining regressors. That leaves only the marginal cost parameter γ to be recovered from this moment equation, which is identified under the assumption $E\left[\ln(\nu_{jht}) Z^S_{jht}\right] = 0$. For our supply side instrument we use the optimal instrument (Hansen 1982): $Z^S := \frac{\partial \ln(\nu)}{\partial \gamma}$.

We combine the supply and demand moments in a GMM estimator. We estimate demand and supply jointly, imposing supply optimality constraints:

$$mc_j \in [0, p_{jht}], \tag{11}$$

and

$$\frac{\partial s_{jht}}{\partial p_{jht}} \frac{p_{jht} - mc_j}{s_{jht}} \in [-1, 0]. \tag{12}$$

3.3 Search/contracting model

The demand and pricing models specified thus far are based upon a consideration set $\mathcal{J}_{ht}$ that has been determined by hospital h's search over the set of all possible brands available at a point in time $\mathcal{J}_t$. The search process we specify is intended to accommodate the following features: (1) Allowing for various sources of heterogeneity across brands in (beliefs about) preferences θ_{hj}; bargaining β_j, β_h; and search costs sc_{hj} seems important for fitting the data and intuitions about agent information and behavior. (2) Hospitals make repeated purchases from $\mathcal{J}_{ht}$, so the composition of $\mathcal{J}_{ht}$ matters (in a similar vein to optimal retailer assortment or portfolio choice problems). (3) Unless the full brand set $\mathcal{J}_t$ is small, the computability of an (optimal) search model relies on the *ordering* of brands not changing (too much) with the size/composition of the inframarginal consideration set $\mathcal{J}_{ht}$.

One cannot satisfy all of $[1-3]$ simultaneously and also meet the assumptions required for the algorithms commonly applied to simplify dynamic search problems (e.g. Chade and Smith 2006; Weitzman 1979). Our approach is to instead construct moment inequalities based on weak search assumptions. In this setting, we find that very weak assumptions are still informative. Moreover, adding inequalities based on stronger assumptions generates identical results to a full search model consistent with those assumptions (results varying the strength of the bounds assumptions available by request). Importantly, the approach is computationally tractable because our particular bounds can be computed prior to search cost estimation.

3.3.1 Search identification and estimation

We first consider how brands $j \in \mathcal{J}_{ht}$ provide upper bounds on search costs. Assuming that the set of firms purchased from in the data must be a subset of the firms searched, $\mathcal{J}_{ht} \subseteq \mathcal{J}_{ht}^{search}$, then it follows that any $j \in \mathcal{J}_{ht}$ must have been worth paying search costs for at some stage during the search process. When brands are all substitutes for one another, the weakest such assumption comes from the value of j versus the outside good:

$$E_{\xi,\nu}[AV_{jht}(\theta, \beta, \gamma; \emptyset \cup j)] > sc_{jht} \ \forall j \in \mathcal{J}_{ht}. \tag{13}$$

This assumption is weak in that it provides a greater value than adding j to any other set, and also in that it is potentially consistent with both simultaneous and sequential models of search (and in the case of sequential, any order of search).

Analogously, brands $k \in \mathcal{J}_t \setminus \mathcal{J}_{ht}$ provide lower bounds on search costs. Assuming there is at least one brand k that a given hospital has not searched, $k \in \mathcal{J}_t \setminus \mathcal{J}_{ht}^{search}$, then it must *not* have been worth paying search costs for this brand at any time during the search process. For substitutes, the weakest such assumption comes from value of k as part of the full possible choice set:

$$E_{\xi,\nu}[AV_{kht}(\theta, \beta, \gamma; \mathcal{J}_t)] < sc_{kht} \ k \in \mathcal{J}_t \setminus \mathcal{J}_{ht}^{search}. \tag{14}$$

Again, this assumption is weak in that it provides a smaller value than adding k to any other set, and also in that it is potentially consistent with both simultaneous and sequential models of search (and in the case of sequential, any order of search). To account for the possibility that $\mathcal{J}_{ht}^{search}$ may be much larger than $\mathcal{J}_{ht}$ (and maybe approaching $\mathcal{J}_t$), we further take the minimum of the bounds above over all potential $k \in \mathcal{J}_t \setminus \mathcal{J}_{ht}$.

For both bounds, estimating $AV_{jht}(\theta, \beta, \gamma; \emptyset \cup j)$ involves computing counterfactual equilibrium prices and quantities for the respective counterfactual choice sets (for each hospital-year). We do this using our estimated demand (θ) and supply (β, γ) parameters. We assume that $E[\xi_{jht}^o | j \in \mathcal{J}_{ht}]$ (or $E[\xi_{kht}^o | k \notin \mathcal{J}_{ht}]$) takes the predicted value from the control function in the demand estimation, and we numerically integrate over 20 draws each from the estimated densities of the demand residual ξ_{jht}^u and supply residual $\ln(\nu_{jht})$.[23]

[23]While market share data provides information on actual mean utility for the products in the consider-

We use these bounds and parameterize search costs by:

$$sc_{jht} = \psi^0 + \psi^{t-1} 1_{\{j \in \mathcal{J}_{ht-1}\}} + \psi^{far} Z^{far} + \epsilon \tag{15}$$

where ψ^0 estimates the mean search/contracting cost (in units of dollars of expected surplus per patient) for a product/vendor that is new to the hospital, ψ^{t-1} and ψ^{far} allow for lower search costs for products sourced by the hospital in the previous year or in other product categories in the current year, respectively. We implement the constrained optimization approach to computing the identified sets for this moment inequalities estimator described in Dong et al. (2017).

4 Quantifying Sources of Price Dispersion

4.1 Estimated parameters: demand, pricing, and search

Table 2 shows the estimated parameters from our demand and pricing models. The first parameters of interest are the demand parameters themselves. There is a wide range over the estimated nesting parameter λ, which captures substitution to the outside good (recall λ is an approximate measure of within-category correlation in substitution, with 0 characterizing the plain logit model, and 1 implying no substitution to the outside good). The class average λ is similar for PPIs and non-PPIs, both at about 0.2. However, class averages conceal a great deal of heterogeneity within PPIs and other medical/surgical products. In non-PPIs, the estimated nesting parameter ranges from close to zero (e.g., suture anchors) to 0.73 (e.g., pneumatic compression cuffs).

The price coefficient θ^p (that scales dollars into logit utils, and recall the extreme value type 1 normalization fixes the standard deviation of the error to $1 - \lambda * (\pi/\sqrt{6} \sim 1.28)$) is rather small across most categories, indicating little price sensitivity in product usage patterns in general. There are, however, large differences in price sensitivity across product categories. PPI usage is almost two orders of magnitude less price-sensitive than other medical/surgical supplies. This is consistent with what we would expect, given the relative amounts of brand-manufacturer-specific branding, and the relative importance of PPIs and

ation set, predicting mean utility for products not used is more difficult. Similar to work in teacher "value added", we apply Bayesian shrinkage procedures to the parameters $\theta_{jt}, \theta_h, \beta_{jt}, \beta_h$ to account for the fact that some are estimated from very few observations. We are also concerned that buyers may be more pessimistic/optimistic than our control function predicts, and have thus explored robustness to different assumptions on $E[\xi^o_{jht} | j \in \mathcal{J}_{ht}]$ (similar to Eizenberg (2014) in the context of firm entry). For example, this bound can be weakened further by assuming the hospital has extreme beliefs in the portion of the demand unobservable it knows prior to search, e.g. $\xi^o_{jht} = \max_{ht} \xi_{jht}$ for $j \in \mathcal{J}_{ht}$ and $\xi^o_{jht} = \min_{ht} \xi_{jht}$ for $k \notin \mathcal{J}_{ht}$.

non-PPIs in determining procedural outcomes. We find it reassuring that our relatively parsimonious demand model is able to empirically identify this anticipated feature of physician preference items via substitution patterns revealed by the data.

Table 2: Demand and Pricing Parameter Estimates

	N_h	p		λ	$\theta^p * 1,000$	AV^{CS}		$\frac{p-mc}{p}$		B	
		μ	$\frac{\sigma}{\mu}$			μ	$\frac{\sigma}{\mu}$	μ	$\frac{\sigma}{\mu}$	μ	$\frac{\sigma}{\mu}$
Non-PPIs											
Bone Grafts	352	$1,934	0.13	0.26	-0.174	$3,437	0.08	0.47	0.13	0.21	0.26
Kyphoplasty Kit	217	$1,761	0.11	0.41	-0.204	$2,951	0.16	0.25	0.31	0.14	0.45
Bone Nails	470	$1,404	0.15	0.09	-0.330	$1,893	0.12	0.61	0.09	0.32	0.23
Bone Implant Putty	414	$1,064	0.15	0.15	-0.248	$3,120	0.06	0.34	0.32	0.11	0.46
Polymeric Mesh	262	$921	0.12	0.00	-0.401	$2,296	0.05	0.24	0.35	0.09	0.46
Tissue Fusion Device	160	$408	0.06	0.00	-1.626	$536	0.06	0.26	0.14	0.17	0.21
Orthopedic Fixation Systems	379	$394	0.18	0.00	-0.892	$937	0.08	0.46	0.22	0.17	0.40
Cath., Misc	148	$361	0.31	0.26	-0.553	$1,394	0.07	0.31	0.22	0.09	0.68
Suture Anchors	422	$319	0.11	0.00	-2.224	$364	0.10	0.26	0.31	0.19	0.42
Surgical Staplers	591	$215	0.14	0.26	-1.224	$409	0.09	1.00	0.00	0.34	0.15
Linear Staplers	529	$160	0.12	0.14	-3.089	$249	0.09	0.21	0.34	0.14	0.49
GI Staples	542	$133	0.18	0.09	-2.125	$350	0.08	0.64	0.07	0.20	0.26
Batteries	470	$93	0.23	0.00	-3.288	$266	0.07	0.38	0.31	0.13	0.54
Laproscopic Clip Applier	488	$92	0.08	0.42	-2.652	$226	0.12	0.36	0.10	0.14	0.23
Pulse Oximeter Probes	304	$85	0.15	0.10	-3.418	$247	0.07	0.25	0.37	0.10	0.54
Trocars	593	$40	0.18	0.18	-7.609	$95	0.09	0.31	0.36	0.13	0.56
Pneumatic Compression Cuffs	316	$28	0.10	0.73	-5.564	$66	0.24	0.28	0.27	0.15	0.43
Liquid Adhesives	599	$27	0.09	0.61	-11.738	$42	0.21	0.20	0.34	0.12	0.51
Sutures	647	$8	0.14	0.16	-24.815	$33	0.03	0.20	0.38	0.05	0.53
IV Infusion Pumps	230	$5	0.07	0.24	-125.080	$6	0.11	0.19	0.24	0.15	0.34
IV Administration Kits	636	$4	0.09	0.13	-64.557	$13	0.04	0.18	0.34	0.06	0.43
IV Tubing Extensions	625	$2	0.09	0.30	-109.270	$7	0.05	0.17	0.34	0.05	0.44
Surgical Gloves	664	$1	0.06	0.23	-71.219	$10	0.02	0.96	0.00	0.09	0.07
Linen Underpads	521	$0	0.08	0.48	-48.341	$15	0.16	0.18	0.32	0.01	0.47
Average(24)	441	$394	0.13	0.22	-20.443	$790	0.09	0.36	0.24	0.14	0.40
Physician Preference Items											
Pacemakers	357	$4,376	0.11	0.33	-0.086	$4,697	0.12	0.77	0.04	0.42	0.15
Humeral Shoulder Prosth.	264	$2,486	0.19	0.00	-0.099	$8,663	0.06	0.59	0.14	0.15	0.32
Drug Eluting Stents	314	$1,571	0.05	0.38	-0.789	$613	0.22	0.19	0.22	0.33	0.30
Allografts	322	$1,530	0.12	0.00	-0.223	$3,739	0.06	0.50	0.11	0.17	0.23
Acetabular Hip Prosth.	458	$1,418	0.23	0.00	-0.404	$1,816	0.20	0.42	0.31	0.27	0.51
Hemostatic Media	232	$324	0.05	0.56	-0.736	$687	0.23	0.19	0.22	0.10	0.34
Average(6)	324	$1,951	0.13	0.21	-0.390	$3,369	0.15	0.44	0.17	0.24	0.31

The other output of the demand estimation reported here is the consumer surplus component of added value – the additional surplus accruing to the physician-patient-hospital for having access to brand j – implied by the estimated utility model. In part due to the low price sensitivity of demand, these are relatively large in dollar value across all categories, but increasingly so in the preference item categories. Coefficients of variation of the added value range from 0.15 in PPIs to 0.09 in non-PPIs – slightly larger than price coefficients of variation for PPIs and smaller for non-PPIs. These AV^{CS} estimates are the key input from the demand estimation that, combined with the granular price data, allows estimation of the bargaining model.

The bargaining model then provides estimates of marginal costs, which in turn define markups; and bargaining parameters, which rationalize the split of the total added value AV^{TS} between device vendors and hospitals, conditional on the consideration set. The bargaining parameters indicate that manufacturers of non-PPIs on average capture 14 percent of the total value added up for negotiation, while PPIs capture more, on average 24 percent. This is especially interesting in light of the fact mentioned earlier, that we assume the preferences estimated in the demand model are the relevant preferences for measuring added value in the bargaining model. The bargaining residual captures the relative weight put on vendor and hospital surplus to explain the price variation as a function of added value variation. Thus, a smaller relative share to vendors can be driven by the purchasing agents involved in negotiation perceiving brands as closer substitutes than provider substitution patterns would indicate. In that light, one explanation consistent with prior work on PPIs would be that the larger bargaining split is driven by the greater ability/desire of physicians to transmit their preferences to purchasing (or conversely, the inability of purchasing to move physician market share) for PPIs than for non-PPIs.

4.2 Search costs and breadth of buyer-supplier relationships

In Section 4.3, we will further examine the contributions of search, bargaining, and demand to the level and variation of markups. First, though, we consider the parameter estimates of the search cost functions. Of particular interest are the coefficients on our measures of the percent of hospital spending with a focal supplier in other product categories. These speak to one channel via which supplier breadth is argued to potentially enhance welfare, but they also have the potential to create switching costs that can raise markups of broad incumbents. Table 3 provides evidence on this matter.

For each product category, the first three columns of Table 3 report the search cost parameter estimates, and the last four columns provide summary statistics (mean and standard deviation) for the extent of purchasing persistence ($1(q_{jh,t-1} > 0)$) and breadth ($Z^{s,far}$) to facilitate understanding of the quantities involved. Parameters are point identified in all categories.[24] Looking at the category averages, search/contracting costs are substantial, but not overwhelming, averaging roughly 10 percent of price.

The averages mask a great deal of heterogeneity across product categories, though. In addition, the product categories with high baseline search costs are almost uniformly also the categories where vendor breadth matters most. But even in these categories breadth can

[24] The estimates here follow the approach outlined in theory and estimation, using mean estimates for expected unobservables ξ^{o}_{jht} (parameter estimates for looser bounds using $\xi^{o}_{jht} = \xi^{\max}_{\min}$ and also different bounds more akin to the "stability" inequalities in Ghili (2018) available by request). For the loosest bounds, most parameters are no longer point identified.

Table 3: Determinants of Search/Contracting Costs

	ψ^0	ψ^{far}	ψ^{t-1}	$Z^{s,far}$		$1\{t-1\}$	
				μ	σ	μ	σ
Non-PPIs							
Bone Grafts	18.5	0.1	0.2	0.12	0.08	0.44	0.31
Kyphoplasty Kit	421.8	-35.9	15.1	0.19	0.15	0.42	0.35
Bone Nails	31.7	0.0	-0.0	0.15	0.11	0.39	0.34
Bone Implant Putty	36.4	0.1	0.0	0.12	0.09	0.42	0.33
Polymeric Mesh	27.6	0.1	0.2	0.03	0.02	0.34	0.31
Tissue Fusion Device	141.7	-3.7	-44.0	0.06	0.05	0.36	0.27
Orthopedic Fixation Systems	24.5	-0.4	-0.0	0.13	0.08	0.42	0.33
Cath., Misc	141.5	-66.3	-7.5	0.04	0.02	0.48	0.25
Suture Anchors	4.4	-0.0	-0.0	0.01	0.01	0.39	0.33
Surgical Staplers	6.6	0.0	0.0	0.09	0.08	0.47	0.32
Linear Staplers	11.4	0.0	-0.0	0.09	0.09	0.48	0.31
GI Staples	13.5	-0.1	-0.3	0.09	0.08	0.43	0.33
Batteries	0.3	-0.0	-0.1	0.05	0.05	0.36	0.31
Laproscopic Clip Applier	4.6	-0.1	-0.0	0.09	0.09	0.53	0.32
Pulse Oximeter Probes	37.0	-0.1	-1.4	0.06	0.05	0.43	0.25
Trocars	1.9	-0.0	-0.0	0.09	0.08	0.50	0.31
Pneumatic Compression Cuffs	12.7	-1.9	-0.2	0.10	0.09	0.49	0.26
Liquid Adhesives	1.5	-0.0	-0.0	0.09	0.08	0.50	0.27
Sutures	0.2	0.0	-0.0	0.09	0.08	0.49	0.30
IV Infusion Pumps	1.4	-0.3	-0.9	0.12	0.10	0.54	0.26
IV Administration Kits	0.2	0.0	-0.0	0.15	0.11	0.50	0.27
IV Tubing Extensions	0.0	0.0	-0.0	0.14	0.10	0.51	0.26
Surgical Gloves	0.2	0.0	-0.0	0.12	0.11	0.55	0.27
Linen Underpads	0.6	-0.0	-0.0	0.12	0.10	0.51	0.25
Average(24)	39.2	-4.5	-1.6	0.10	0.08	0.46	0.30
Physician Preference Item							
Pacemakers	134.9	-0.1	0.6	0.08	0.07	0.42	0.33
Humeral Shoulder Prosth.	173.4	0.7	0.6	0.08	0.05	0.39	0.34
Drug Eluting Stents	617.4	-2.8	-29.3	0.04	0.04	0.44	0.31
Allografts	69.2	0.1	-0.4	0.04	0.02	0.36	0.28
Acetabular Hip Prosth.	45.0	-0.0	-0.0	0.08	0.07	0.43	0.33
Hemostatic Media	77.8	2.7	-2.4	0.05	0.07	0.49	0.31
Average(6)	186.3	0.1	-5.2	0.06	0.06	0.42	0.32

decrease search costs by up to two percent (at the mean percent of relative spend on the same vendor in *other* product categories).

It is difficult to tell directly from these parameter estimates the extent to which search/contracting frictions vs. product differentiation are the driving force behind the large estimated markups in these medical devices. Both seem to be nontrivial, but without an equilibrium model, it is not clear how to assess their relative role. Relatedly and similarly, it is difficult to characterize the relative contribution of demand versus bargaining heterogeneity in the observed price dispersion across hospitals. The next Section computes several counterfactuals in order to shed further light on these issues.

4.3 Decomposing price variation

Here we use parameter estimates to examine various decompositions of the prices observed in the data. If the price variation and market power that generates markups are driven by true brand differentiation in quality (where quality could either be vertical quality for the average use, or horizontal use-specific match quality), then that has quite different welfare implications than physician-specific brand preferences or search frictions that limit the choice set. To disentangle these factors, we explore equilibrium prices and consumer surplus in several counterfactual scenarios.

First, to better understand the drivers of price variation across hospitals, we condition on the observed choice sets in the data, but we counterfactually shut down heterogeneity across hospitals in bargaining ($B_{jht} = \frac{\beta_{jt}}{\beta_{jt}+\frac{1}{|\mathcal{H}|}\sum_{h\in\mathcal{H}}\beta_h}$) and demand ($\delta_{jht} = \theta_{jt} + \frac{1}{|\mathcal{H}|}\sum_{h\in\mathcal{H}}(\theta_h + \theta^f f(Z^s_{jht}) + \xi^u_{jht}) - \theta^p p_{jht}$), each in turn. We then compute equilibrium prices at these parameters, and recompute the coefficient of variation in price across hospitals (within product-time). This provides a measure of the extent to which the price dispersion in the data is being driven by bargaining or demand, conditional on the observed choice sets.

The results are summarized in the left panel (columns 1-4) of Table 4. In the third and fourth columns, we show the proportion of price dispersion ($\%\frac{\sigma}{\mu}$) remaining after taking out price variation driven by bargaining ($p^{-\sigma_B}$) and preferences ($p^{-\sigma_D}$), respectively. We find that in every product category, variation in prices across hospitals in the observed data is driven much more by variation in bargaining than by demand. This dominance of bargaining variation is slightly stronger in non-PPIs, but varies more within our product classes than across them; e.g., bargaining is far more important to price dispersion for bone nails than for linen underpads among the non-PPIs, and more important for humeral shoulder prostheses than for drug-eluting stents among the PPIs. All of the categories show some covariance between bargaining and demand estimates, with the price variation explained by each totaling more than the observed price variation in the data.

Table 4: Decomposing Variation in Prices and Markups

	$\mathcal{J}_{ht}$				$\mathcal{J}_t, q^w(\mathcal{J}_{ht})$		$\mathcal{J}_t$		$E[CS(\mathcal{J}_t) - CS(\mathcal{J}_{ht})]$
	p		$p^{-\sigma_B}$	$p^{-\sigma_D}$	p		p		
	μ	$\frac{\sigma}{\mu}$	$\%\frac{\sigma}{\mu}$	$\%\frac{\sigma}{\mu}$	μ	$\frac{\sigma}{\mu}$	μ	$\frac{\sigma}{\mu}$	
Non-PPI									
Bone Grafts	$1,927	0.12	9.1	102.2	$1,909	0.12	$1,772	0.04	$4,103
Kyphoplasty Kit	$1,715	0.14	17.1	99.7	$1,685	0.13	$1,681	0.06	$2,081
Bone Nails	$1,441	0.15	2.7	101.0	$1,436	0.15	$1,437	0.07	$2,803
Bone Implant Putty	$1,064	0.15	3.1	101.7	$1,060	0.15	$958	0.07	$3,941
Polymeric Mesh	$916	0.12	1.6	101.8	$915	0.12	$891	0.05	$4,023
Tissue Fusion Device	$437	0.07	16.7	112.0	$435	0.07	$443	0.03	$755
Orthopedic Fixation Systems	$393	0.19	1.4	101.8	$392	0.19	$377	0.07	$1,032
Cath., Misc	$394	0.20	8.6	121.3	$388	0.20	$337	0.23	$898
Suture Anchors	$330	0.12	1.6	101.3	$330	0.12	$323	0.05	$476
Surgical Staplers	$216	0.14	13.8	104.6	$211	0.15	$203	0.08	$690
Linear Staplers	$159	0.11	4.7	103.3	$158	0.11	$153	0.06	$286
GI Staples	$134	0.17	5.5	104.1	$132	0.17	$128	0.09	$380
Batteries	$100	0.22	0.0	121.4	$100	0.22	$85	0.08	$287
Laproscopic Clip Applier	$90	0.08	38.5	108.5	$87	0.09	$91	0.05	$232
Pulse Oximeter Probes	$88	0.11	6.0	125.1	$88	0.11	$74	0.12	$178
Trocars	$36	0.19	3.0	101.9	$36	0.19	$35	0.08	$125
Pneumatic Compression Cuffs	$27	0.11	49.4	121.0	$26	0.10	$26	0.08	$114
Liquid Adhesives	$28	0.10	42.6	110.6	$27	0.09	$26	0.05	$44
Sutures	$8	0.15	1.0	101.8	$8	0.15	$7	0.06	$36
IV Infusion Pumps	$5	0.06	22.9	133.9	$5	0.06	$5	0.04	$6
IV Administration Kits	$4	0.09	3.4	109.0	$4	0.09	$4	0.03	$15
IV Tubing Extensions	$2	0.09	9.2	106.0	$2	0.09	$2	0.03	$8
Surgical Gloves	$1	0.06	33.9	126.8	$1	0.06	$1	0.13	$10
Linen Underpads	$0	0.07	41.4	114.1	$0	0.07	$0	0.05	$14
Average(24)	$396	0.13	14.1	109.8	$393	0.12	$378	0.07	$939
Physician Preference Item									
Pacemakers	$4,203	0.12	20.3	101.4	$4,131	0.12	$4,216	0.07	$6,896
Humeral Shoulder Prosth.	$2,387	0.20	2.4	101.5	$2,380	0.20	$2,302	0.10	$11,456
Drug Eluting Stents	$1,504	0.06	30.2	103.0	$1,493	0.06	$1,522	0.04	$632
Allografts	$1,594	0.12	2.1	104.2	$1,592	0.12	$1,387	0.07	$5,375
Acetabular Hip Prosth.	$1,426	0.22	1.4	101.5	$1,424	0.22	$1,395	0.13	$2,700
Hemostatic Media	$314	0.09	54.7	110.4	$306	0.09	$307	0.04	$676
Average(6)	$1,905	0.14	18.5	103.7	$1,888	0.14	$1,855	0.07	$4,622

We also explore the comparative static of how markups and cross-hospital variation change as the choice set size increases. Specifically, we compute counterfactual prices and quantities in equilibrium when all hospitals have the same choice set, $\mathcal{J}_t$, consisting of all purchased brands across all hospitals in that year.[25] This corresponds to the potential set of suppliers hospitals would have access to in a world with no search/contracting frictions.[26]

The results are summarized in the center-right panel (columns 5-8) of Table 4. Columns 5-6 allow prices to adjust, but hold quantities fixed at their values under the true observed

[25]For cases where the hospital did not use this brand in the data, we set demand and bargaining unobservables to zero.

[26]For some product categories, this may be an extreme case to consider. Without specifying a specific model of search – and, as discussed, optimal search in this context is computationally infeasible – we are unable to consider counterfactual choice sets under intermediate search costs.

choice set $\mathcal{J}_{ht}$. Columns 7-8 allow both prices and quantities to adjust. Increasing the choice set size (holding quantities fixed) has a relatively small effect on prices, decreasing average by 0-4 percent, depending on the product category. This is presumably driven by the low price sensitivity and low percentages of the bargaining split to manufacturers.

In addition to price effects, the final column of Table 4 measures the effect of contracting frictions on overall consumer surplus, in dollars per device used. Related to the low price sensitivity and large added values estimated, the additional surplus from expanding the choice set is often large, easily outweighing any price considerations of search frictions.

5 Conclusion

Price dispersion across buyers for the same exact product must come from dispersion in marginal costs of distribution or dispersion in markups. Thus, absent the former, price dispersion is an indicator of market power, and understanding the economic forces underlying the price dispersion is critical for understanding impediments to market efficiency. In business-to-business markets, price dispersion across buyers due to variation in markups is of antitrust interest per se, due to its potential impact on downstream competition.

In this paper, we explore price dispersion in a large and policy-relevant market: hospital supply contracting. In detailed data from hospital purchase orders across a variety of product markets, we document substantial price dispersion across hospitals for the same products purchased from the same vendors. In spite of the fact that the demand side in this application solely includes hospitals, there is large variation across product categories in the preferences of end users, the concentration and bargaining power of suppliers, and the potential importance of information and search/contracting frictions. We document reduced form evidence suggesting that all of these features may play some role in the observed data.

We then develop a structural model allowing for heterogeneity across hospital buyers in demand for differentiated products, price negotiations, and frictions in the search/contracting process that determines who contracts with whom. We address the problem of potential selection of suppliers based on unobserved preferences by using a control function based on hospital exposure to a vendor in product categories that are likely unrelated in terms of user preferences, but potentially related at an administrative level through their impact on contracting costs. We also leverage exogenous variation in prices due to the introduction of benchmarking information to the hospitals in the sample. The "generic" nature of these identification strategies allows us to obtain credible estimates of demand and supply across a large variety of product categories. Finally, we estimate search/contracting frictions using a moment inequalities approach that is computationally feasible and accommodates various forms of the search/contracting process, which we do not observe directly.

Our estimates suggest that large markups are primarily driven by large perceived product differentiation and lack of price sensitivity among health care providers in their product usage decisions. This problem is especially severe in "physician preference items", where price sensitivity is nearly two orders of magnitude lower than in more commoditized non-PPI medical/surgical supplies. Hospital purchasing administrators are able to counteract this somewhat by exercising a large degree of monopsony power in their price negotiations, but this ability varies widely across hospitals, driving most of the observed price dispersion across hospitals.

These results suggest that moves to reduce buyer search costs, such as more information/transparency on suppliers and supplier consolidation to more "one stop shopping" product portfolios, are unlikely to have much impact in decreasing markups or price dispersion. Instead, the potential returns are likely greater to solving provider and administrator agency problems that generate low price sensitivity and large dispersion in bargaining outcomes.

References

Allen, J., Clark, R., and Houde, J.-F. (2013). The effect of mergers in search markets: Evidence from the canadian mortgage industry. *American Economic Review*, 104(10):3365–3396.

Allen, J., Clark, R., and Houde, J.-F. (2018). Search frictions and market power in negotiated price markets. forthcoming, Journal of Political Economy.

Berry, S. and Waldfogel, J. (1999). Free entry and social inefficiency in radio broadcasting. *RAND Journal of Economics*, 30(3):397–420.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics*, 25(2).

Berry, S. T., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Bloom, N., Sadun, R., and VanReenen, J. (2014). Does management matter in healthcare? Stanford Mimeo.

Chade, H. and Smith, L. (2006). You have full text access to this content simultaneous search. *Econometrica*, 74(5):1293–1307.

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.

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

Dinerstein, M., Einav, L., Levin, J., and Sundaresan, N. (2018). Consumer price search and platform design in internet commerce. *American Economic Review*, 108(7):1820–1859.

Dong, B., Hsieh, Y.-W., and Shum, M. (2017). Computing moment inequality models using constrained optimization. USC-INET Research Paper No. 17-21. Available at SSRN: https://ssrn.com/abstract=2990826.

Einav, L., Finkelstein, A., Ryan, S., Schrimpf, P., and Cullen, M. R. (2013). Selection on moral hazard in health insurance. *American Economic Review*, 103(1):178–219.

Eizenberg, A. (2014). Upstream innovation and product variety in the us home pc market. *The Review of Economic Studies*, 81(3):1003–1045.

Ghili, S. (2018). Network formation and bargaining in vertical markets: The case of narrow networks in health insurance. Working paper.

Goeree, M. (2008). Limited information and advertising in the u.s. personal computer industry. *Econometrica*, 76(5):1017–1074.

Goldberg, P. and Verboven, F. (2001). The evolution of price dispersion in the european car market. *Review of Economic Studies*, 68(4):811–848.

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). Transparency and negotiated prices: The value of information in hospital-supplier bargaining. Working paper.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054.

Hausman, J. (1996). Valuation of new goods under perfect and imperfect competition. In Bresnahan, T. and Gordon, R., editors, *The Economics of New Goods*, volume 58. National Bureau of Economic Research.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–162.

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

Ho, K. and Lee, R. S. (2018). Equilibrium provider networks: Bargaining and exclusion in health care markets. Manuscript.

Hong, H. and Shum, M. (2006). Using price distributions to estimate search costs. *RAND Journal of Economics*, 37(2).

Honka, E. (2014). Quantifying search and switching costs in the us auto insurance industry. *RAND Journal of Economics*, 45(4):847–884.

Hortacsu, A. and Syverson, C. (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds. *Quarterly Journal of Economics*, 119(2):403–456.

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

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

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.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(307-42).

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

Petrin, A. and Train, K. (2009). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research*, xlvi.

Roberts, M., Xu, D. Y., Fan, X., and Zhang, S. (2016). The role of firm factors in demand, cost, and export market selection for chinese footwear producers. Revise and resubmit, Review of Economic Studies.

Robinson, J. C. (2008). Value-based purchasing for medical devices. *Health Affairs*, 27(6).

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

Salz, T. (2017). Intermediation and competition in search markets: An empirical case study. Available at SSRN: https://ssrn.com/abstract=2961795.

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.

Sela, R. J. and Simonoff, J. S. (2011). Reemtree: Regression trees with random effects. R package version 0.90.3.

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

Sorensen, A. (2003). Insurer-hospital bargaining: Negotiated discounts in post-deregulation connecticut. *Journal of Industrial Economics*.

Starc, A. and Swanson, A. (2018). Preferred pharmacy networks and drug costs. Working paper.

Stigler, G. (1961). The economics of information. *Journal of Political Economy*, 69(3):213–225.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica*, 47(3).

ELECTRONIC APPENDICES – NOT FOR PRINT PUBLICATION

A Data Appendix

As noted in the main text, the hospital supply transactions data used in this project are typically transmitted as a direct extract from a hospital's materials management database to the PriceGuide™ benchmarking portal. 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. Moreover, as discussed in Grennan and Swanson (2018), at least for coronary stents, the distribution of prices observed in hospitals' pre-join benchmarking data is similar to that observed in an external, representative market research dataset.[27]

The raw transactions data contain 116 million observations for 2,876 members across 3,394 product categories and 2.7 million stock keeping units (SKUs). Our analyses include 30 important product categories, defined by their UMDNS codes. To arrive at this set of product categories, we first restricted to the top 100 categories by overall spending. From these, we excluded categories that were too broad (involving products that were not clearly substitutes) or where data quality seemed to be an issue. Regarding the first cut, we selected eight categories by hand that seemed excessively broad based on their UMDNS names: Blood Collection Tubes, Clinical Reagents, Computer Supplies, In-Vitro Diagnostic (IVD) Kits, Industrial Supplies, Office Supplies, Patient Education Material, and Pharmaceuticals.[28] Regarding the second cut, demand estimation requires that we analyze quantities across hospital-years, within each product category, for a well-defined unit. Although many medical and surgical product categories are sold by the unit (e.g., a single coronary stent), 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. Using these fields, we transformed the price and quantity variables into a price per single unit and quantity of single units purchased, respectively, and we dropped transactions that were missing either unit of measure or conversion factor data. Then, for each product category, we restricted the analysis to transactions that were reported in the modal unit of measure (i.e., if a product category is usually sold in boxes, we include only transactions reported in boxes in our analytic sample). Finally, we dropped product categories from the analysis if more than thirty percent of transactions were lost when we

[27] In our empirical analysis, hospital fixed effects will absorb any persistent misreporting. Nonetheless, there is some evidence that the data are incomplete. For example, we find it unrealistic that some broadly used categories (e.g. examination gloves) do not include data from all hospitals. This can occur when transactions are not submitted with informative free-text item descriptions and are accordingly not assigned UMDNS codes by ECRI.

[28] For example, "IVD Kits" include microbial detection kits costing $2.14 on average, as well as tests for antibiotic-resistant bacteria colonization costing $4,400 on average.

limited the sample to the modal unit of measure. This filter ensures that the included data are meaningfully representative of the category. At this stage, we dropped five additional categories: Examination Gloves, IV Solutions, MRI Contrast, Procedure Kits, and Surgical Packs.

We also dropped several product categories from our analysis due to incompatibilities with the identification strategy detailed in Section 3.1.1. In order to address the identification problem introduced by selection of brands into each hospital's consideration set, we use instruments based on exposure of hospitals to vendors across dissimilar product categories. This approach requires a measure of dissimilarity – a starting point for this classification approach is based on UMDNS codes' relative positions in a UMDNS hierarchy, which is missing for eleven of the remaining product categories: Antibiotic Orthopedic Cement, Cranial Bone Screws, IV Saline, Long Term IV Catheters, Spinal Bone Plates, Spinal Bone Screws, Spinal Rod Implants, Spinal Spacers, Trauma Bone Plates, Trauma Bone Screws, and Vascular Closure Devices.[29]

Our "one-size-fits-all" identification strategy is quite powerful for many heterogeneous product categories in our data, but (not unexpectedly) fails to be powerful for some products. For example, this may happen in cases with a very limited set of vendors and generally complete consideration sets; for such products, the search and contracting costs that are a primary focus of this paper are less meaningful. We only include in our analysis those products for which a joint F-test of significance of the excluded exposure instruments in our first stage has a p-value of less than 0.05. This results in our exclusion of 31 products that would not otherwise be excluded: fourteen non-PPIs such as Blood Chemistry Kits, ECG Recorders, and Oxygen Sensors; and seventeen PPIs such as Bare Metal Stents, Brain Stimulators, and Patellar Knee Prostheses.

Finally, for reasons of practicality, we exclude product categories with prohibitively large datasets – specifically, we exclude product categories requiring greater than 50GB of RAM for supply and demand estimation. This results in our exclusion of nine products that were not otherwise excluded. Seven of these are non-PPIs with large numbers of brands: Balloon Catheters, Bone Plates, Bone Screws, Dressings, Drill Bits, Guide Wires, and Polyglactin Sutures. Two are PPIs: Femoral Hip Prostheses and Tibial Knee Prostheses.

[29]In Appendix A.1 below, we also describe several sampling refinements at the observation, hospital, or brand level. Five small, but expensive, product categories do not survive these refinements: Cochlear Stimulators, Incontinence Neurostimulator, Mammary Prostheses, Skin Expanders, Vagus Nerve Stimulators, and Ventricular Assist Device.

A.1 Data Cleaning and Final Sample

Above, we describe the refinement of the product categories we include in our analysis. Within the included set of product categories, we performed several additional refinements to the sample to address variable availability, suspected errors, and management of outliers.

- First, we limit to usable transaction data (with non-missing memberid, SKU, and manufacturer; and with positive quantity purchased). We also remove transactions with suspected price errors (i.e., brands with mean price an order of magnitude below the median price across brands, and transactions that are integer multiples of other common prices for the same SKU) and transactions with prices in the tails of the observed distribution. The main goal of the latter filter is to remove products that are erroneously included in the product category and/or which are not substitutable with the majority of products in the category. For example, the "surgical staplers" category includes many stapler refill cartridges. In practice, we limit to transactions with prices between the 5^{th} and 95^{th} percentiles of the price distribution.

- In addition to dropping transactions not in the modal unit of measure in each UMDNS code or with a missing conversion factor (e.g., 10 units per box), we also drop transactions with a suspicious unit of measure: with a unit of measure of either "BX" or "CS" and with a conversion factor of 1.

- Keep manufacturers with sufficient observations ($N_{hjt} > 200$) that we can assign brand IDs in the machine learning procedure detailed below.

- Limit to hospitals and health systems (as opposed to laboratories); such that we observe the member hospital's timing of database join in the login data and can use the price shock observed after database join to identify price elasticities; and that merge onto the AHA survey data with non-missing location (HRR), total admissions, outpatient admissions, overall bed size, bed size of each department (obstetric beds, cardiac ICU beds, etc.), and full-time-equivalent staff. The AHA variables are primarily used to construct the total market size in each category.

- Remove small hospital-years, for which demand estimation is less well-behaved and the assumption that the consideration set is equivalent to the set of products purchased in that hospital-year is less palatable. In practice, we remove hospital-years with $q_{ht} < 30$ and hospital-years below the 5^{th} percentile in terms of total quantity.

A.2 Identifying Brands in the Transaction Data

We utilized machine learning methods to categorize SKUs into brand IDs, in order to appropriately control for brand-specific price trends. The absence of a brand identifier 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 most thorough method we employed to identify brands, for a subset of products, 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 for all product categories due to the large number of manufacturers and SKUs. Additionally, many manufacturers' websites were found to be difficult to navigate, particularly once we extended the analysis beyond high-dollar physician preference items. Finally, the item description field was often uninformative as to brand. Hence, we used an algorithmic approach to assign brand identifiers for the other product categories.

Our preferred algorithm implements the Random Effect-Expectation Maximization (RE-EM) estimation method from Sela and Simonoff (2011), which is an adaptation of a recursive partitioning tree algorithm to allow for group effects. With no particular assumption made about the significance of each letter within a SKU, recursive partitioning tree allows us to obtain overfitting-proof groupings that minimizes the 10-fold cross validation error. Furthermore, the group effects in the RE-EM estimation method allow us to control for systematic heterogeneity in price across hospital-time.

Given a transaction $i = 1, \ldots, N$ where N is the size of the dataset within a UMDNS code, price p_i of the transaction, dummy matrix Z_i indicating each transaction's hospital-time group, group effect b_i, and attribute vector $D_i = \{d_{i1}, \ldots, d_{iL}\}$ where d_{il} is the lth digit of the SKU associated with transaction i, the RE-EM proceeds as follows:

1. Initialize estimated group effect $\hat{b_i}$ to zero.
2. Iterate through the following steps until the estimated hospital-time effect $\hat{b_i}$ converges.
   (a) Estimate a regression tree with recursive partitioning on price adjusted by hospital-time group effect, $p_i - Z_i\hat{b_i}$ with attributes D_i. Take the terminal nodes, $j \in J$, of the tree and create an indicator variable, $I(D_i \in j)$.
   (b) Fit a linear model, $p_i = Z_i b_i + I(D_i \in j)\mu_p + \epsilon_i$ and extract $\hat{b_i}$ from the model.
3. Once $\hat{b_i}$ converges, take the final grouping $j \in J$ and use it as the new product identifier for each i.

At each iteration of step (2a), the tree is pruned using 10-fold cross validation at each split; the model retains the simplest tree with cross validation error no more than one standard error away from the tree with the minimum cross validation error.

With this method, we categorized 19,806 SKUs across 30 UMDNS codes into 3,100 RE-EM brands. For surgical staplers and drug-eluting coronary stents, which we validated by hand, we identified 3.8 RE-EM brands per "true" stapler brand, and 0.8 RE-EM brands per "true" drug-eluting stent brand.

B Monte Carlo Evidence on Performance of Control Function Selection Correction

As discussed in Section 3.1.1, we use a control function approach to correct for potential bias introduced by our estimation of preference parameters within an endogenously-formed consideration set. In this Appendix, we use a simple Monte Carlo simulation to illustrate the identification problem and to demonstrate how the control function performs in addressing it.

First, suppose that the consideration set formation process can be well-approximated by the following reduced-form index model:

$$\mathbb{1}(j \in \mathcal{J}_h) := f(\phi_h + \phi_j + Z_{jh}\phi^z + \epsilon_{jh}) \tag{16}$$

where we will specify f using a Probit link or, alternatively, a linear probability model (LPM). Second, in a slightly simplified representation of our demand model, suppose that end user h chooses brand j from the consideration set $\mathcal{J}_h$ upon each use opportunity i to maximize utility represented by:

$$u_{ijh} = \theta_h + \theta_j + \xi_{jh} + \varepsilon_{ijh} \tag{17}$$

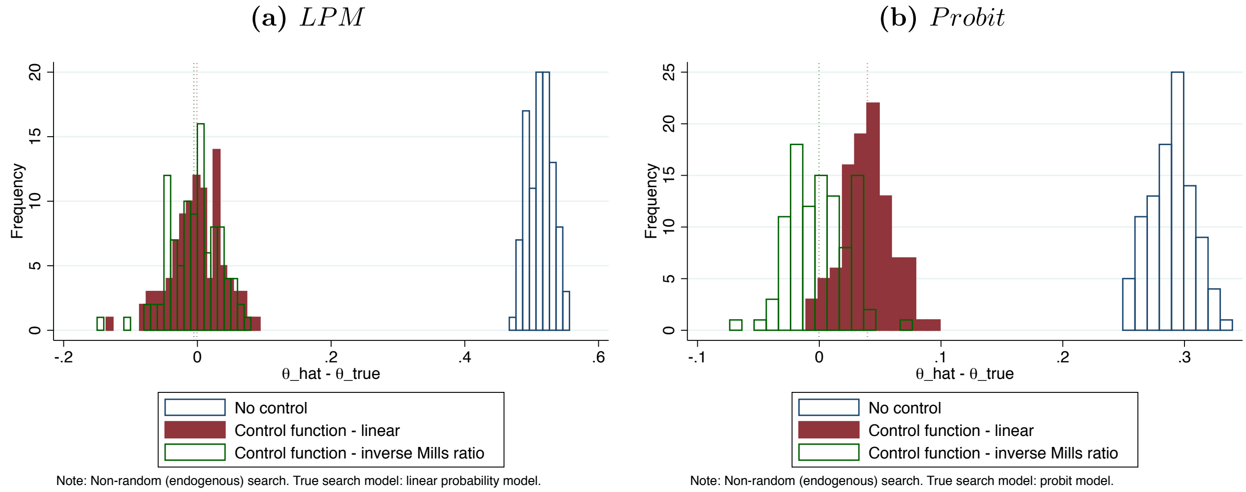
The use-specific i.i.d. unobservable ε_{ijh} is the standard type I extreme value error term (with scale normalized to one), and ξ_{jh} is the unobserved average "match" value between hospital h and brand j.

We allow for endogeneity between demand and "search" via correlation between the search cost shock ϵ_{jh} and ξ_{jh}:

$$\begin{pmatrix} \epsilon \\ \xi \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{pmatrix}\right]$$

In our Monte Carlo simulation, each iteration has $N_i = 500$ use cases for each of $N_h = 50$ hospitals forming consideration sets and then choosing among $N_j = 2$ brands. We draw each independent variable from the following distributions: $\theta_h \sim U[.1, .7]$; $\theta_j \sim U[.3, .7]$; $\phi_h \sim U[0, .2]$; $\phi_j \sim U[.2, .4]$; $Z \sim \mathcal{N}(.1, .1)$. Finally, we let ξ_{jh} and ϵ_{jh} be bivariate normal with means zero, variances of 0.5, and covariance $\rho = 0.5$. In Figure A1, we display the distribution of the bias in estimated brand fixed effects: $\hat{\theta}_j - \theta_j$. The left panel imposes that the true reduced-form search model is a linear probability model; the right panel imposes that the true search model is a Probit.

Figure A1: Monte Carlo Results: Distribution of $\hat{\theta}_j - \theta_j$



The blue bars show the distribution of the bias in the estimated brand fixed effects when demand is estimated assuming no endogeneity. In each panel, the bias is positive, suggesting that a brand's average quality is overestimated when it is inferred only from those hospitals that liked it well enough to put it in their consideration sets. The red bars show the distribution of the bias when we implement a linear control function correction – we regress a dummy for consideration set inclusion linearly on hospital and brand dummies and our excluded instrument Z, then include the residual linearly as a control in the demand model. The green bars show the distribution of the bias when we implement a Probit-based control function correction – we perform a Probit regression of consideration set inclusion on hospital and brand dummies and our excluded instrument Z, then include the inverse Mills ratio of the predicted index as a control in the demand model.

As shown in the Figure, the procedure does a good job of correcting the bias due to consideration set endogeneity when the control function is implemented correctly based on the "true" search model. That is, the estimates based on the linear control function are unbiased when the search model was truly linear, and the estimates based on the inverse Mills ratio control function are unbiased when the search model was truly a Probit. The inverse Mills ratio control function also works well when the search model was truly linear, but the estimates based on the linear control function are still slightly biased upward when the search model was truly a Probit. Thus, the inverse Mills ratio procedure is somewhat more robust to model misspecification.

C Main Results With Different Levels of Product Categorization

In this Section, we present summary statistics, and supply and demand estimates under alternative models of brand identification. In the "Step-Through Algorithm" employed in Grennan and Swanson (2018), we define brands by stepping through each vendors' SKUs from left to right, stopping when the price variation explained within hospital-period levels off; this results in a finer categorization of products than our baseline RE-EM algorithm. In the "Vendor" results described in the following tables, we instead simply assume that supply and demand are each determined at the vendor level.

Several patterns stand out. In Table A1 (A2), we observe larger (smaller) choice $\mathcal{J}$ and consideration $\mathcal{J}_h$ sets, as expected. We observe similar price dispersion for the step-through and RE-EM brands, but we observe much larger dispersion for the vendor level of aggregation; this is not unexpected, as the vendor level of aggregation involves coarser controls for time trends.

In Table A3, we observe similar supply and demand parameters as in our baseline results in Table 2. However, we observe higher nesting parameters, lower price sensitivity, and higher added values at the vendor level of aggregation. This suggests that there is more substitution among vendors than to the outside good than we observe at the brand level, and we likely underestimate price sensitivity when we aggregate to the vendor level by insufficiently controlling for within-vendor variation in quality.

Table A1: Step-Through Algorithm – Summary of Purchasing Categories

	N_h	Annual Spend \$1000s	p		$\|\mathcal{J}\|$	$\|\mathcal{J}_h\|$		$Pr[j^* \in \mathcal{J}_h]$	$Pr[j^* = j^*_h]$
			μ	$\frac{\sigma}{\mu}$		μ	$\frac{\sigma}{\mu}$		
Non-PPIs									
Kyphoplasty Kit	217	\$182	\$2,465	0.08	28	5	0.49	0.55	0.24
Bone Grafts	352	\$470	\$1,587	0.12	256	11	0.53	0.74	0.20
Bone Nails	470	\$175	\$1,441	0.14	397	21	0.56	0.65	0.18
Bone Implant Putty	414	\$168	\$966	0.13	269	13	0.53	0.35	0.09
Polymeric Mesh	262	\$100	\$791	0.11	329	18	0.51	0.42	0.08
Orthopedic Fixation Systems	379	\$117	\$349	0.17	251	22	0.65	0.75	0.39
Cath., Misc	148	\$85	\$327	0.15	42	5	0.46	0.07	0.06
Suture Anchors	422	\$101	\$311	0.11	183	18	0.53	0.61	0.10
Surgical Staplers	591	\$111	\$217	0.14	263	9	0.51	0.21	0.07
Linear Staplers	529	\$81	\$157	0.12	178	7	0.59	0.19	0.10
GI Staples	542	\$118	\$126	0.19	231	6	0.66	0.32	0.18
Laproscopic Clip Applier	488	\$62	\$91	0.08	59	3	0.49	0.27	0.16
Batteries	470	\$270	\$74	0.18	128	5	0.63	0.12	0.07
Pulse Oximeter Probes	304	\$116	\$71	0.17	122	4	0.76	0.17	0.12
Trocars	593	\$60	\$38	0.15	275	8	0.61	0.26	0.07
Pneumatic Compression Cuffs	316	\$95	\$29	0.13	44	3	0.56	0.30	0.26
Liquid Adhesives	599	\$53	\$24	0.09	49	2	0.53	0.40	0.32
Sutures	647	\$16	\$6	0.14	603	11	0.77	0.30	0.15
IV Infusion Pumps	230	\$86	\$5	0.08	34	2	0.55	0.22	0.21
IV Administration Kits	636	\$74	\$4	0.12	281	5	0.62	0.17	0.09
IV Tubing Extensions	625	\$40	\$2	0.11	373	5	0.68	0.08	0.05
Surgical Gloves	664	\$91	\$1	0.06	533	15	0.78	0.17	0.04
Linen Underpads	521	\$30	\$0	0.18	53	2	0.54	0.23	0.16
Average (23)	453	\$117	\$395	0.13	217	9	0.59	0.33	0.15
Physician Preference Items									
Pacemakers	357	\$534	\$4,282	0.11	101	12	0.42	0.84	0.24
Drug Eluting Stents	314	\$1,028	\$1,588	0.05	14	5	0.42	0.81	0.41
Allografts	322	\$146	\$1,434	0.11	443	12	0.85	0.17	0.10
Hemostatic Media	232	\$115	\$321	0.05	19	3	0.47	0.61	0.35
Average (4)	306	\$456	\$1,906	0.08	144	8	0.54	0.61	0.27

Table A2: Vendor – Summary of Purchasing Categories

	N_h	Annual Spend \$1000s	p		$\lvert\mathcal{J}\rvert$	$\lvert\mathcal{J}_h\rvert$		$Pr[j^* \in \mathcal{J}_h]$	$Pr[j^* = j^*_h]$
			μ	$\frac{\sigma}{\mu}$		μ	$\frac{\sigma}{\mu}$		
Non-PPIs									
Bone Nails	470	\$176	\$1,505	0.18	10	3	0.44	0.91	0.58
Kyphoplasty Kit	217	\$183	\$1,497	0.37	4	2	0.43	0.97	0.85
Bone Grafts	352	\$465	\$1,409	0.32	15	4	0.45	0.98	0.77
Bone Implant Putty	414	\$168	\$982	0.40	20	4	0.46	0.75	0.42
Polymeric Mesh	262	\$101	\$628	0.36	22	6	0.31	0.71	0.21
Cath., Misc	148	\$87	\$338	0.46	28	4	0.43	0.07	0.06
Suture Anchors	420	\$100	\$314	0.16	7	4	0.29	0.94	0.52
Surgical Staplers	591	\$112	\$212	0.28	22	3	0.39	0.30	0.14
Linear Staplers	529	\$83	\$162	0.24	22	2	0.48	0.31	0.26
GI Staples	541	\$119	\$138	0.23	21	2	0.42	0.55	0.19
Laproscopic Clip Applier	488	\$63	\$84	0.18	25	2	0.46	0.33	0.28
Pulse Oximeter Probes	304	\$125	\$53	0.30	34	2	0.56	0.17	0.13
Batteries	470	\$131	\$53	0.38	40	4	0.56	0.19	0.16
Trocars	593	\$61	\$38	0.24	42	3	0.46	0.38	0.17
Liquid Adhesives	599	\$54	\$18	0.18	34	2	0.49	0.40	0.31
Pneumatic Compression Cuffs	313	\$96	\$16	0.18	15	1	0.41	0.35	0.32
Sutures	645	\$16	\$7	0.40	32	2	0.54	0.38	0.32
IV Infusion Pumps	227	\$87	\$4	0.09	11	1	0.37	0.35	0.33
IV Administration Kits	636	\$75	\$4	0.36	46	2	0.54	0.34	0.19
IV Tubing Extensions	625	\$39	\$1	0.29	40	2	0.55	0.35	0.21
Surgical Gloves	663	\$90	\$1	0.27	25	2	0.46	0.42	0.26
Average (21)	453	\$116	\$355	0.28	25	3	0.45	0.49	0.32
Physician Preference Items									
Pacemakers	354	\$532	\$4,350	0.12	5	3	0.30	0.99	0.68
Humeral Shoulder Prosth.	264	\$212	\$2,195	0.22	9	3	0.41	0.62	0.26
Drug Eluting Stents	314	\$1,029	\$1,610	0.05	6	3	0.28	0.83	0.39
Allografts	322	\$149	\$1,379	0.34	20	3	0.47	0.77	0.45
Acetabular Hip Prosth.	458	\$264	\$1,148	0.24	12	4	0.40	0.79	0.33
Hemostatic Media	230	\$116	\$274	0.09	8	2	0.46	0.87	0.71
Average (6)	324	\$384	\$1,826	0.18	10	3	0.39	0.81	0.47

Table A3: Step-Through Algorithm – Demand and Pricing Parameter Estimates

	N_h	p		λ	$\theta^p * 1,000$	AV^{CS}		$\frac{p-mc}{p}$		B	
		μ	$\frac{\sigma}{\mu}$			μ	$\frac{\sigma}{\mu}$	μ	$\frac{\sigma}{\mu}$	μ	$\frac{\sigma}{\mu}$
Non-PPIs											
Kyphoplasty Kit	217	$2,465	0.08	0.52	-0.102	$4,880	0.11	0.23	0.25	0.11	0.34
Bone Grafts	352	$1,587	0.12	0.32	-0.138	$4,648	0.05	0.21	0.37	0.07	0.48
Bone Nails	470	$1,441	0.14	0.19	-0.289	$2,051	0.11	0.52	0.12	0.27	0.24
Bone Implant Putty	414	$966	0.13	0.21	-0.228	$3,298	0.05	0.25	0.36	0.07	0.48
Polymeric Mesh	262	$791	0.11	0.00	-0.347	$2,710	0.04	0.22	0.31	0.07	0.41
Orthopedic Fixation Systems	379	$349	0.17	0.00	-0.756	$1,143	0.05	0.50	0.15	0.14	0.32
Cath., Misc	148	$327	0.15	0.29	-0.324	$2,331	0.06	0.50	0.07	0.08	0.27
Suture Anchors	422	$311	0.11	0.00	-1.534	$455	0.08	0.64	0.06	0.31	0.17
Surgical Staplers	591	$217	0.14	0.26	-0.769	$802	0.05	0.83	0.02	0.18	0.17
Linear Staplers	529	$157	0.12	0.14	-3.099	$250	0.09	0.20	0.34	0.13	0.49
GI Staples	542	$126	0.19	0.05	-2.298	$346	0.08	0.55	0.10	0.17	0.30
Laproscopic Clip Applier	488	$91	0.08	0.42	-4.275	$139	0.13	0.22	0.21	0.14	0.33
Batteries	470	$74	0.18	0.00	-2.766	$338	0.04	0.24	0.34	0.07	0.61
Pulse Oximeter Probes	304	$71	0.17	0.18	-0.552	$1,545	0.03	0.22	0.36	0.01	0.57
Trocars	593	$38	0.15	0.22	-7.298	$97	0.08	0.28	0.35	0.12	0.53
Pneumatic Compression Cuffs	316	$29	0.13	0.81	-3.887	$84	0.30	0.29	0.28	0.14	0.55
Liquid Adhesives	599	$24	0.09	0.64	-6.364	$80	0.19	0.20	0.29	0.06	0.50
Sutures	647	$6	0.14	0.01	-25.041	$38	0.02	0.20	0.40	0.04	0.55
IV Infusion Pumps	230	$5	0.08	0.20	-94.077	$9	0.08	0.16	0.34	0.09	0.45
IV Administration Kits	636	$4	0.12	0.19	-76.901	$10	0.06	0.20	0.36	0.09	0.48
IV Tubing Extensions	625	$2	0.11	0.33	-140.310	$5	0.07	0.26	0.21	0.09	0.34
Surgical Gloves	664	$1	0.06	0.94	-37.904	$2	0.09	0.24	0.15	0.14	0.23
Linen Underpads	521	$0	0.18	0.56	-54.302	$13	0.20	0.24	0.34	0.01	0.62
Average (23)	453	$395	0.13	0.28	-20.155	$1,099	0.09	0.32	0.25	0.11	0.41
Physician Preference Items											
Pacemakers	357	$4,282	0.11	0.26	-0.069	$6,815	0.08	0.96	0.00	0.38	0.12
Drug Eluting Stents	314	$1,588	0.05	0.21	-0.996	$568	0.18	0.18	0.23	0.34	0.29
Allografts	322	$1,434	0.11	0.00	-0.246	$3,797	0.05	0.19	0.40	0.07	0.50
Hemostatic Media	232	$321	0.05	0.56	-0.756	$677	0.22	0.17	0.22	0.09	0.33
Average (4)	306	$1,906	0.08	0.26	-0.517	$2,964	0.13	0.38	0.21	0.22	0.31

Table A4: Vendor – Demand and Pricing Parameter Estimates

	N_h	p		λ	$\theta^p * 1,000$	AV^{CS}		$\frac{p-mc}{p}$		B	
		μ	$\frac{\sigma}{\mu}$			μ	$\frac{\sigma}{\mu}$	μ	$\frac{\sigma}{\mu}$	μ	$\frac{\sigma}{\mu}$
Non-PPIs											
Bone Nails	470	$1,505	0.18	0.73	-0.102	$2,502	0.38	0.92	0.02	0.40	0.29
Kyphoplasty Kit	217	$1,497	0.37	0.28	-0.190	$3,894	0.28	0.48	0.31	0.20	0.58
Bone Grafts	352	$1,409	0.32	0.47	-0.020	$29,007	0.08	0.54	0.23	0.03	0.58
Bone Implant Putty	414	$982	0.40	0.26	-0.177	$3,575	0.15	1.00	0.00	0.22	0.41
Polymeric Mesh	262	$628	0.36	0.53	-0.185	$2,522	0.17	0.54	0.29	0.15	0.65
Cath., Misc	148	$338	0.46	0.22	-0.594	$1,374	0.08	0.32	0.35	0.09	1.01
Suture Anchors	420	$314	0.16	0.14	-0.687	$1,032	0.09	1.00	0.00	0.23	0.17
Surgical Staplers	591	$212	0.28	0.72	-0.569	$646	0.35	0.77	0.06	0.23	0.48
Linear Staplers	529	$162	0.24	0.69	-1.015	$590	0.40	0.39	0.29	0.16	0.78
GI Staples	541	$138	0.23	0.69	-0.903	$491	0.37	0.87	0.03	0.23	0.43
Laproscopic Clip Applier	488	$84	0.18	0.73	-1.358	$382	0.39	1.00	0.00	0.23	0.40
Pulse Oximeter Probes	304	$53	0.30	0.77	-0.998	$424	0.27	0.28	0.36	0.08	1.22
Batteries	470	$53	0.38	0.32	-2.121	$331	0.13	0.46	0.31	0.11	0.74
Trocars	593	$38	0.24	0.38	-5.434	$113	0.15	0.48	0.25	0.16	0.51
Liquid Adhesives	599	$18	0.18	0.71	-2.217	$238	0.26	0.27	0.34	0.03	0.72
Pneumatic Compression Cuffs	313	$16	0.18	0.95	-0.925	$572	0.56	0.85	0.02	0.08	0.99
Sutures	645	$7	0.40	0.87	-3.837	$91	0.36	0.41	0.35	0.09	1.24
IV Infusion Pumps	227	$4	0.09	0.92	-15.796	$44	0.34	0.18	0.34	0.05	1.14
IV Administration Kits	636	$4	0.36	0.40	-47.467	$13	0.18	0.45	0.34	0.15	0.73
IV Tubing Extensions	625	$1	0.29	0.44	-93.726	$6	0.17	1.00	0.00	0.20	0.34
Surgical Gloves	663	$1	0.27	0.82	-73.730	$4	0.40	1.00	0.00	0.23	0.48
Average (21)	453	$355	0.28	0.57	-12.002	$2,279	0.26	0.63	0.19	0.16	0.66
Physician Preference Items											
Pacemakers	354	$4,350	0.12	0.46	-0.011	$54,726	0.13	0.52	0.12	0.04	0.26
Humeral Shoulder Prosth.	264	$2,195	0.22	0.50	-0.095	$4,963	0.28	0.97	0.01	0.32	0.31
Drug Eluting Stents	314	$1,610	0.05	0.83	-0.173	$1,143	0.40	0.23	0.16	0.27	0.32
Allografts	322	$1,379	0.34	0.00	-0.210	$4,056	0.14	0.55	0.26	0.17	0.59
Acetabular Hip Prosth.	458	$1,148	0.24	0.49	-0.188	$2,426	0.23	0.76	0.06	0.28	0.36
Hemostatic Media	230	$274	0.09	0.47	-0.706	$891	0.20	0.32	0.15	0.11	0.30
Average (6)	324	$1,826	0.18	0.46	-0.230	$11,367	0.23	0.56	0.13	0.20	0.36

D Role of Strategic Exclusion

One of our key observations in the raw data is that each hospital sources its medical and surgical supplies from only a small subset of the vendors available in the market. One explanation for this observation is that search and contracting frictions prevent hospitals from making purchases from many different vendors. Another potentially important explanation is that hospitals strategically exclude some vendors in order to strengthen their bargaining leverage (Ho and Lee 2018). This Appendix explores whether the patterns in our data are consistent with strategic exclusion. Specifically, we conduct two different tests regarding whether strategic exclusion alone, without any other frictions such as search, is sufficient to explain the observed consideration sets and prices in these data.

We begin by testing whether the observed consideration sets and prices are consistent with a stability condition in the Nash-in-Nash with Threat of Replacement model of strategic exclusion in Ho and Lee (2018) (hereafter, NiNTR). A hospital h's consideration set $\mathcal{J}_h$ and prices $\boldsymbol{p}_h$ violate the stability condition if higher bilateral gains-from-trade are available if the hospital replaces one of the included brands $j \in \mathcal{J}_h$ with one of the excluded brands $k \in \mathcal{J} \setminus \mathcal{J}_h$ and pays the vendor of brand k its reservation price, defined to be the minimum price k would be willing to accept to be included in h's consideration set. That is, stability requires:

$$\pi_h(\mathcal{J}_h) - \pi_h(\mathcal{J}_h \setminus j) + q_{hj}(p_{hj} - mc_j) \geq \pi_h((\mathcal{J}_h \setminus j) \cup k) - \pi_h(\mathcal{J} \setminus j) + q_{hk}(p^{res}_{hk} - mc_k).$$

In this condition, the price of the excluded brand p_{hk} is one that makes the brand's vendor indifferent between selling and not selling, so that price equals k's marginal cost mc_k. If equilibrium prices are NiNTR prices, then the price with each included brand j is the minimum between the outside option price and Nash-in-Nash price, where the outside option price is the price at which the hospital is indifferent between purchasing from j and purchasing from the excluded brand k' that gives the highest gains from trade among all of the excluded brands.

To demonstrate how well this stability condition holds in our data, we calculate gains-from-trade for each hospital and brand pair using the demand and cost estimates presented in Section 4. Note that we employ a similar method for calculating gains-from-trade for counterfactual hospital-brand pairs as in Section 3.3: we assume that $E[\xi^o_{jht} | j \in \mathcal{J}_{ht}]$ (or $E[\xi^o_{kht} | k \notin \mathcal{J}_{ht}]$) takes the predicted value from the control function employed in the demand estimation, and we employ alternative assumptions regarding hospitals' assumptions on ξ^u. We then calculate the fraction of markets (hospital-years) that violate the stability condition for each product category. Figure A2 shows the fraction of markets violating stability.

Each bar around the markers represents one standard error estimated from a nonparametric bootstrap, resampling at the market level. The solid markers in Figure A2 summarize the stability violations assuming that the hospital-brand specific unobservable preference heterogeneity term for each brand outside the observed consideration set (ξ^u_{hkt} for each $k \in \mathcal{J} \setminus \mathcal{J}_{ht}$) is $\mathbb{E}(\xi^u)$, the mean of the estimated distribution of the demand residual. Under this implied model of expectations – that hospitals do not know ξ^u_{hkt} for brands outside their consideration sets – it is difficult to rationalize the observed consideration sets and prices based on strategic exclusion alone. For the average product category, 86% of hospital-years' consideration sets violate stability for at least one brand.

One possible explanation for this pattern is that the "violations" are driven by hospitals' unobserved preferences for brands ξ^u, in which case we would be overestimating the unobserved preferences for excluded brands. To address this concern, we repeat the same exercise assuming that hospitals have very low preferences for the excluded brands. Specifically, we assume that each hospital's unobserved preference for an excluded brand ξ^u_{hkt} is at the 25^{th} percentile of the estimated distribution of the demand residual. The results are presented in the hollow markers in Figure A2. Under these assumptions, we estimate that, for the average product category, 76% of hospital-years' consideration sets violate stability for at least one brand. We have also run the same exercise using the 1^{st} percentile of the estimated ξ^u_{hkt} distribution, and get qualitatively the same results.

Taken together, these results suggest that observed consideration sets and prices are not consistent with optimal strategic exclusion as in the NiNTR model. It also bears noting that this result is not driven by any particular model of the level of product at which consideration sets are formed. In Figure A3 below, we use alternative levels of aggregation to check stability. In the previous results, we defined brands using the RE-EM approach described above in Appendix A.2. In the left panel of Figure A3 below, we define brands more finely using the step-through algorithm (see Appendix C and Grennan and Swanson (2018)); in the right panel of Figure A3 below, we define brands more coarsely using vendor identifiers.

As expected, stability violations appear more pervasive in panel (a) and less pervasive in panel (b). However, our main takeaway is that the majority of markets exhibit stability violations under all specifications.

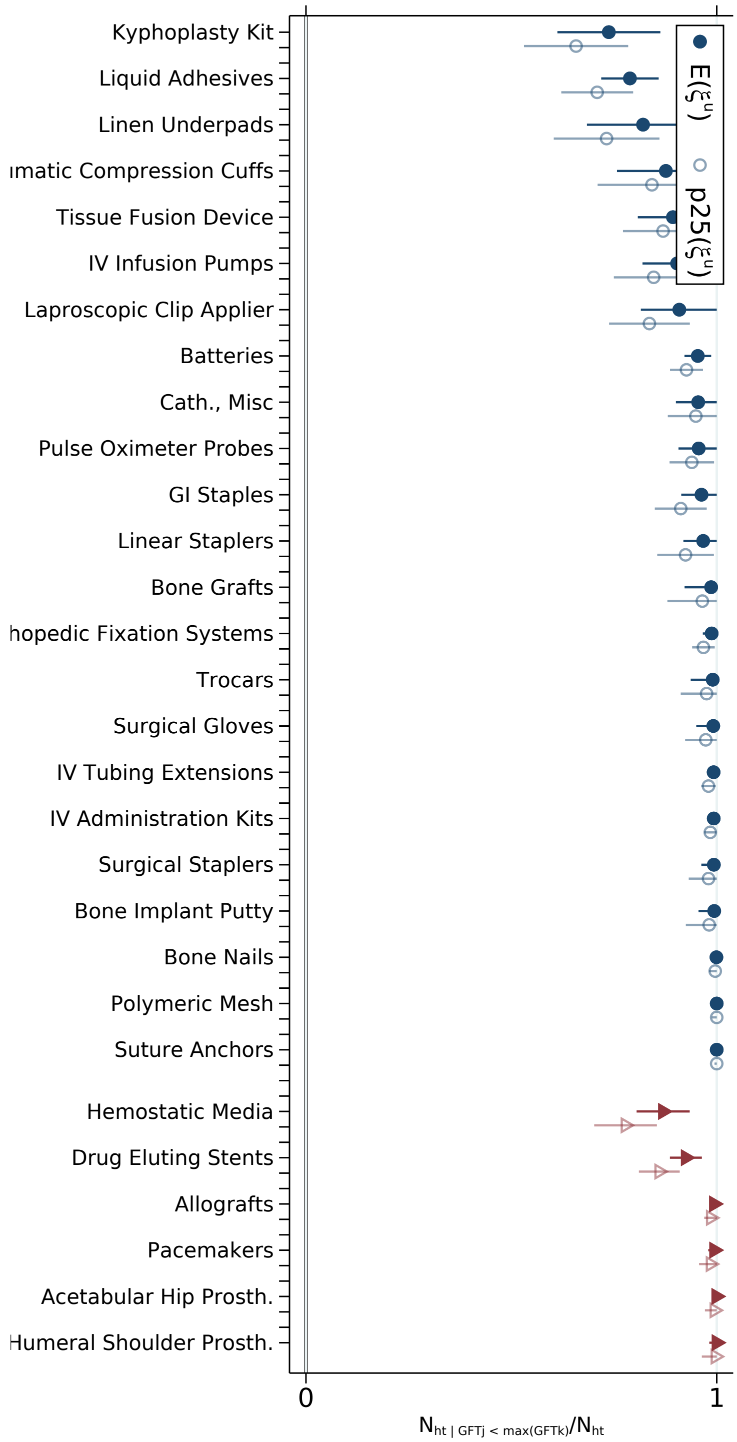
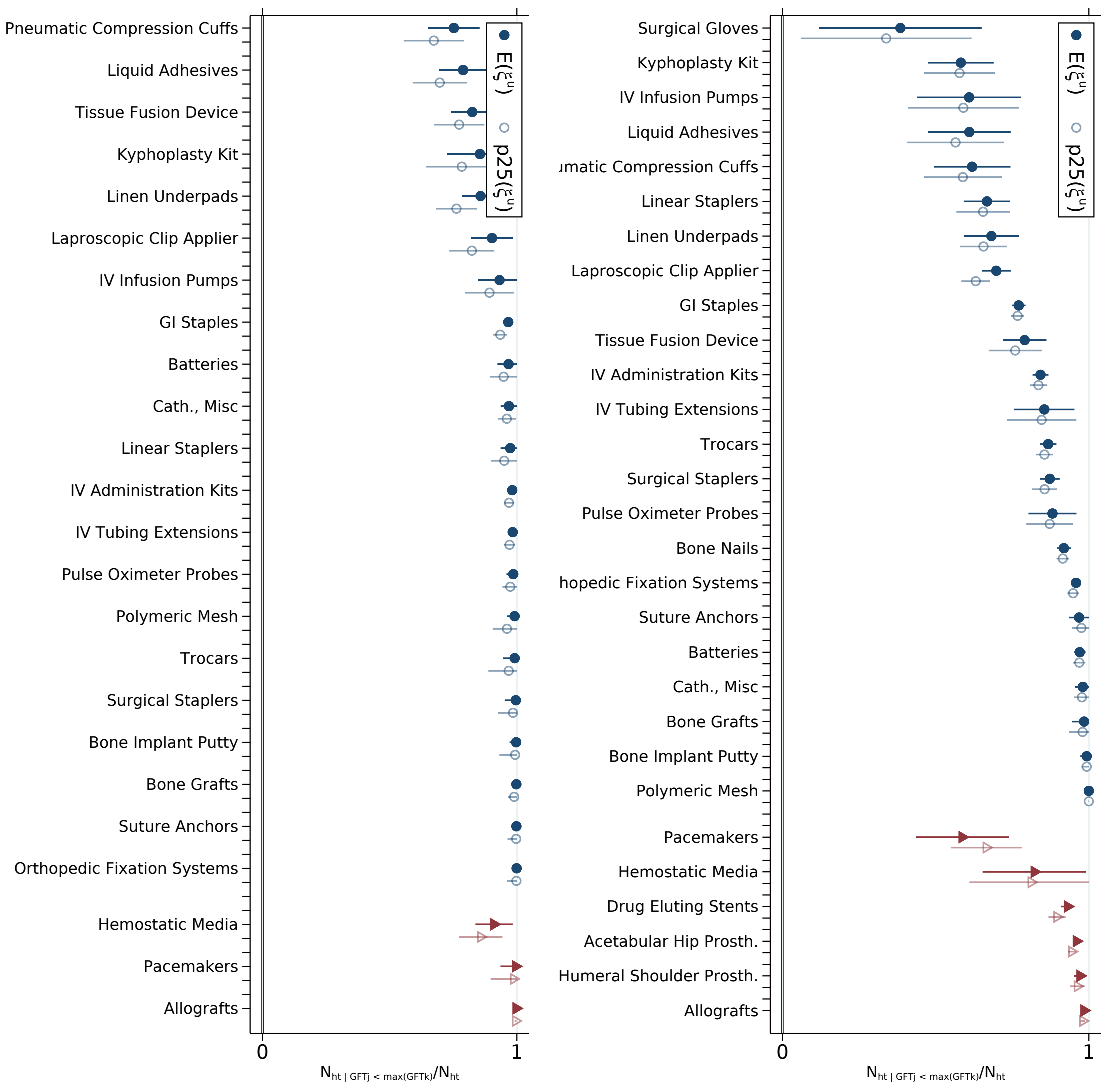


Figure A2: Average Fraction of Hospital-Years that Violate Stability – j =RE-EM Brand



(a) j=Step-Through Brand

(b) j=Vendor

Figure A3: Average Fraction of Hospital-Years that Violate Stability – Alternative j Definitions