Search and Wholesale Price Discrimination

Guillermo Marshall*

May 7, 2019

Abstract

Firms often choose not to post prices in wholesale markets, and buyers must incur costs to discover prices. Inspired by evidence of customized pricing (e.g., some customers pay up to 70% more than others) and search costs, I estimate a search model to study how personalized pricing impacts efficiency in a wholesale market. I find that price discrimination decreases total surplus by 11.6% and increases the sellers' profits by up to 52.1%. These effects are partially explained by price discrimination softening competition through a decrease in search incentives, illustrating how price discrimination may magnify the efficiency costs of search frictions.

Keywords: Price discrimination, search, wholesale market, vertical transactions
JEL classifications: L11, L13, D22, D43

*Department of Economics, University of Illinois at Urbana-Champaign. e-mail: gmarshll@illinois.edu. Acknowledgements: I thank Igal Hendel, Aviv Nevo, Rob Porter, Katja Seim (Editor), and two anonymous referees for detailed feedback as well as Rob Clark, Fernando Luco, Álvaro Parra, Tiago Pires, and numerous seminar participants at various institutions for helpful suggestions. All mistakes are my own.

1 Introduction

In contrast to the Bertrand paradox, many markets for homogeneous goods feature market power and heterogeneity in the prices paid by buyers. Examples include markets for intermediate goods, where sellers choose not to post prices because of customer heterogeneity, and as a result, buyers must incur costs to discover prices. In this article, I study a wholesale market where some buyers pay up to 70 percent more than others for the same good on the same day, and where pricing patterns suggest price discrimination based on search costs. Search costs are a concern because they reduce efficiency by limiting customer mobility to low-price sellers. Aside from the allocative inefficiencies generated by price discrimination, price discrimination affects competition intensity by changing the incentives of customers to search and, as a consequence, may either lessen or magnify the effects of search costs on efficiency.

I study the market for wholesale food in an urban area in the US, which serves as an example of a competitive market for homogeneous goods. The sellers are food distributors selling restaurant supplies, and the buyers are restaurants that make repeated purchases. By exploiting a unique dataset with data at the transaction level, I find three facts that are opposite to what one would expect from a competitive market for homogeneous goods. First, sellers enjoy market power. I discuss below the evidence supporting search costs as the source of this market power. Second, and as mentioned above, the extent of price and markup heterogeneity for each product is large. Third, these differences in prices and markups are systematic at the buyer level, suggesting that sellers actively practice price discrimination.

Inspired by these facts, I propose and estimate a structural search model to measure the welfare implications of customized pricing when there is heterogeneity in both search costs and demand. Importantly, the model is able to replicate the key stylized facts that I document throughout the article. In the model, search costs make it costly for buyers to move to low-price sellers, thereby giving sellers market power. By not committing to posted prices, sellers set buyer-specific prices by trading off profit with the probability of serving a particular customer.

Customer heterogeneity gives sellers incentives to price discriminate. When buyers have perfect information about prices, price discrimination has an ambiguous effect on total welfare (Schmalensee 1981, Varian 1985, Holmes 1989, Corts 1998). Price discrimination may, however, increase or decrease the buyers' propensity to search when they do not know all prices in the market and, hence, change the intensity of price competition. Although it is always a dominant strategy for a seller to customize prices, this intensified competition can result in lower equilibrium profits with price discrimination. How price discrimination affects market efficiency in environments with information frictions is, therefore, an empirical question.

The data I use in this study contain detailed information on all transactions that were completed in an eight-month period by a single food distributor. Each record is a transaction, defined as a customer–product–day combination, and includes information such as the unit price, unit wholesale cost, quantity, and customer observables. The richness of these data is crucial for my study. The panel structure of the data allows me to observe interruptions in business relationships, which point to reasons behind customers' decisions to move their business to an alternative seller. Moreover, the wholesale cost information associated with each transaction allows me to compute markups and understand how prices depend on customer heterogeneity while controlling for costs.

I find the following evidence consistent with search costs. First, the probability of a customer–seller breakup is positively correlated with price changes. This relationship between breakups and prices is consistent with the comparative static result of a sequential search model in which the likelihood that the price will surpass the customer's reservation price and trigger search increases in price. A fixed reservation price implies that only a price increase can trigger search, as buyers will never reject prices that are lower than the prices which they have previously accepted. Second, I find that a significant share of the variance in price–cost margins can be explained by time-invariant differences across customers, which suggests customized pricing. In particular, by using a proxy for the net search benefits of a buyer (i.e., the importance of each product in the customer's total expenditure), I find evidence that sellers set higher prices for customers who have a lower propensity to search, which is consistent with pricing in a search model where sellers personalize prices by trading off profit with the probability of serving a customer.

I then develop a model of customer search and pricing that captures these features. In the model, customers are heterogeneous in their demand function, price sensitivity, and in their search costs. When facing a price quote, a customer must decide whether to accept the price quote or to search sequentially for a better price among other sellers. The equilibrium search probability is increasing in price with a price elasticity that varies at the customer level. Sellers offer a single product and face wholesale costs that are independently drawn from a cost distribution that is common to all sellers. Price heterogeneity comes from both cost heterogeneity across sellers and sellers practicing price discrimination. Inertia in buyer–seller matches is a consequence of both search costs and the assumption that customers get a free quote from the seller with whom they have most recently transacted. I assume sellers observe each customer's demand and average search cost and, as a consequence, set buyer-specific prices by trading off profit with the probability of serving a customer. Motivated by descriptive evidence, I abstract away from the possibility of long-term contracts in the model.

In my framework, the strategies of the rival sellers only enter the objective function of each seller through the customers' probability of search, which only depends on the equilibrium price distribution through an index or aggregator (i.e., the equilibrium net value of matching with a new seller). This feature of the model significantly reduces the dimensionality of the state space of each seller's problem and makes analyzing a market with a large number of players tractable. By assuming that the data were generated by a symmetric stationary equilibrium, it suffices to observe a single firm following the equilibrium strategy to identify the primitives of the model, as each buyer's equilibrium net value of matching with a new seller can be treated as a parameter to be estimated. The identification of the equilibrium net value of matching with a new seller follows from the model prediction that a seller will best-respond by setting a lower price when the buyer has a greater propensity to search, which in turn is a function of the equilibrium net value of matching with a new seller.

Using the estimates of the structural model, I measure how price discrimination affects competition and welfare, relative to the equilibrium where sellers set uniform prices. I find that customized pricing reduces the total surplus of completed transactions by 11.6 percent relative to the uniform price case, and that the sellers' expected profits on average increase by 0.3 to 52.1 percent. These results can be explained as follows. Customers with above average search costs face higher price quotes in the price discrimination equilibrium relative to the uniform price equilibrium. They find it too costly to increase their search activity to compensate for the higher initial price quotes and pay higher prices in the price discrimination equilibrium. That is, price discrimination softens the intensity of price competition for these buyers. Customers with below average search costs face higher price quotes in the uniform price equilibrium. They, however, find it relatively cheap to increase their search intensity to compensate for these higher price quotes and thus ultimately pay similar prices in both equilibria. This implies that customers on average pay higher prices in the price discrimination equilibrium—a result driven by high-search-cost buyers. Here, price discrimination thus magnifies the efficiency costs of search frictions.

This article is related to several strands in the literature. A large body of theoretical work has studied the welfare effects of price discrimination. Schmalensee (1981) and

Varian (1985) study the case of a monopolist; Katz (1984), Katz (1987), Holmes (1989), Corts (1998), Yoshida (2000), and Armstrong and Vickers (2001) extend the analysis to competitive environments. The former articles focus on distributional inefficiencies and output changes caused by price discrimination, whereas the latter incorporate the competition intensity effects of price discrimination as well.

A few empirical articles have studied price discrimination in the context of vertical or wholesale markets. Dafny (2010) studies price tailoring in the market that gathers health-insurance carriers selling employer-sponsored health insurance, whereas Villas-Boas (2009) and Grennan (2013) study the welfare effects of price discrimination in the wholesale market for coffee and for medical devices, respectively. Other empirical articles studying how price discrimination affects welfare in diverse economic environments include work by Goldberg (1996), Leslie (2004), Busse and Rysman (2005), Mortimer (2007), Hendel and Nevo (2013), Marshall (2015), and Beckert et al. (2015).

This article also relates to the literature on search costs and price competition (see, for instance, theoretical contributions by Diamond 1971, Reinganum 1979, Varian 1980, Burdett and Judd 1983, Stahl 1989). On the empirical side of this literature, Sorensen (2000) relates heterogeneity in the net benefits of searching to price heterogeneity in the prescription drug market. Hong and Shum (2006) exploit mixed strategy equilibrium conditions to estimate search costs using only data on price distributions. The approach in Hong and Shum (2006) has been extended to allow for differentiated products (Hortaçsu and Syverson, 2004) and consumer learning (De los Santos et al., 2017). Aurichio and Braido (2018) study price dispersion in the Brazilian auto-insurance market and argue that the price dispersion can be explained by the dynamic incentives of brokers when consumers face search and switching costs. Salz (2017) studies the impact of intermediaries in a market with search frictions, and shows that intermediaries—who are disproportionately used by high search cost buyers—increase welfare by increasing the share of consumers who can afford to compare prices and, thus, intensify competition.

Closest to this article are a series of articles by Allen et al. (2014a,b,c), who present evidence of price dispersion in the Canadian mortgage industry. Their analysis suggests that the price variation can in part be explained by price discrimination based on bargaining leverage (or propensity to search). Further, they find that changes in competition differentially impact consumers depending on their search costs, and a significant share of the welfare loss caused by search costs is due to price discrimination. My article builds on their work by studying the mechanism through which price discrimination magnifies the efficiency effects of search costs in a wholesale market with customer heterogeneity.

The rest of the article is organized as follows. A description of the data as well as descriptive evidence is presented in Section 2. Section 3 presents the empirical model, and the estimation and model estimates are discussed in Section 4. Section 5 presents the analysis of the welfare effects of price discrimination, and Section 6 concludes.

2 Data and Descriptive Analysis

Background and Data Description

The data were collected from a food distributor located in an urban area in the US. The distributor serves about 1,000 customers, which are mainly restaurants, and sells several hundred different products. Products are defined at the bar-code level. The sample period covers eight consecutive months.

The distributor employs a sales force to engage with buyers. Each salesperson manages a number of accounts and sets the prices for each transaction. The sales force is given complete discretion in choosing prices, but is given profit-maximization incentives.[1] Aside from a base salary, the salespeople keep a fixed share of the profit of each transaction that they complete. The profit of each transaction is computed using a product-specific unit cost that is announced every day by the distributor through its salesforce software.[2] From the perspective of a salesperson, the unit cost announced by the distributor is the relevant marginal cost for the pricing decisions. The unit cost of a product only varies over time (i.e., there is no within-day variation across customers). The distributor uses linear prices only.

In the data, each record is a transaction (defined as a customer–product–day combination). For each transaction, the data include the date, identifiers for the customer and product, the unit price, the unit cost announced to the salesforce, the quantity, and an identifier for the salesperson that completed the transaction. In addition, the data also include the name and ZIP Code of each customer.

A key variable that I define is an indicator for whether a customer stops purchasing a *particular* product from the food distributor (henceforth, "a breakup"). Although I do not directly observe when breakups occur, I construct an indicator based on the time elapsed between a customer's purchases of a product. That is, I classify a

[1] Jindal and Newberry (2018) study another setting where a seller chooses not to commit to posted prices. In their setting, the seller sets a posted price but incentivizes consumers to negotiate prices, which generates heterogeneity in the prices paid by buyers.

[2] If an item is sold below cost, the salesperson has to reimburse the distributor for the losses.

seller–customer–product relationship as broken when a customer has not purchased the product from the food distributor for an unusually long period of time. Specifically, the time elapsed since a recorded transaction must exceed the greater of 50 business days or the observed maximum number of days observed between transactions (defined at the customer–product level) for the relationship to be classified as broken.[3] By definition, transactions that are close enough to the end of the sample period are never classified as breakups.

As discussed in the next subsection, there are 2,734 products in the dataset. Many of these products are rarely transacted. For this reason, most of the descriptive analysis will be conducted on the subsample of products with at least 250 transactions (192 products). I also repeat the descriptive analysis restricting the sample to the seven products with the highest number of transactions (henceforth, the "top 7" subsample). The top 7 products include flour, shortening, bleach, trash bags, plastic forks, napkins, and canned peppers, and have at least 1,000 recorded transactions each. The structural analysis will be conducted on the top 7 subsample.

Motivating Facts

Table 1 presents summary statistics. Panel A shows that the distributor sold 2,734 different products in this time period, with an average of 70.4 transactions per product. Panel B shows that the subset of products with at least 250 transactions was composed of 192 products, with an average of 521.1 transactions per product. Henceforth, the descriptive analysis restricts attention to products with at least 250 transactions, unless otherwise noted.

Panel B also shows product-level averages for two variables capturing the frequency of transactions. The average number of transactions per customer–product combination was 13.8, with some products having an average number of transactions per customer as low as 5.6 (10th percentile) and as high as 26.6 (90th percentile).[4] The table also shows that the average number of days before a customer repurchased a product was 10.8 days, although there is significant variation across products (the 10th and 90th percentiles were 6.3 and 16 days, respectively).[5]

Panel C shows that the number of customers who purchased at least one product

[3]The analyses based on this breakup indicator are robust to changes to this criterion.

[4]See Figure C1 in the Online Appendix for the distribution of transactions per customer–product combination. The mean and the 10th and 90th percentiles are 11.2, 1, and 27 transactions, respectively.

[5]The number of days between purchases is defined as the number of days between two completed transactions. See Figure C2 in the Online Appendix for the distribution of the product-level average of the number of days between two transactions by a particular customer.

during the sample period was 553. Customers on average purchased 9.5 different products in a given week, with some purchasing as many as 21 products per week (90th percentile). The average and median time between the first and last transactions of a customer across all products were 67 and 87 percent of the sample period, respectively, suggesting that most customers were making repeated transactions.

Panel D of Table 1 reports summary statistics at the transaction level, where a transaction is defined as a customer–product–day combination. The table shows that the average unit price in a transaction was 27.6 dollars, and the average quantity was 5.6 units. The average number of days between transactions, defined at the customer–product level, was 10.3 days.[6]

Panel E replicates Panel D, but restricts attention to the seven products with the greatest number of transactions. These products have an average number of transactions of 1,624.9. These are costly to store (e.g., a 50 pound bag of flour, a case with six one-gallon bottles of bleach) and nonseasonal. I isolate these products to show that the data patterns discussed below are not driven by easily storable goods or seasonality. The average unit price for these products was 20.7 dollars, and customers only purchased an average of 2.8 units of these products in a given transaction.

Figure 1 captures a key pattern in the data. The figure shows variation in prices for two products using transactions that were completed on one particular day. Because the prices correspond to one particular day, every salesperson faced the same marginal cost when choosing these prices (see the discussion in the previous subsection). Hence, the price variation in Figure 1 translates directly into variation in price–cost margins at the buyer level. The histograms show that the highest to lowest price ratios are 1.24 and 1.2 for products A and B, respectively. The figure suggests significant price heterogeneity across buyers and provides evidence in favor of customized pricing.

Table 2 extends the analysis of the variation in prices across buyers by presenting a price–cost margin decomposition for a wider set of products. The table shows that 64.4 percent of the markup variation is systematic at the product level (Column 1), whereas 96.2 percent of it is systematic at the customer–product level (Column 2).[7] The high share of margin variation that is systematic at the customer–product level provides further support for customized prices and suggests that margins largely vary

[6]The variable for the number of days between purchases has fewer observations than the other variables in this panel because it is not defined for the last transaction of each customer–product combination.

[7]The markup variation at the product level captures both product effects and average differences across products in the types of buyers who purchase the product. A regression of markups on customer fixed effects only has an R^2 of 11.9 percent, suggesting that the product effects dominate.

as a function of characteristics at the customer–product level that are fixed over time.

Column 5 of Table 2 shows that the margin variation that is systematic at the customer–product level does not get smaller when restricting attention to products that are relatively costly to store (i.e., top 7 products). This result alleviates the concern that the price variation may be in part driven by the ability of certain restaurants to manipulate their quantity in order to obtain better prices (e.g., "I'll double my order if you give me a lower price"), as manipulating quantities is more costly for the hard-to-store products in the top 7 subsample. A similar pattern is found in Table C1 in the Online Appendix, which restricts attention either to products that are unlikely to be purchased by the same customer (e.g., same-brand products offered in different sizes) or to products that are likely to be purchased together (e.g., oil and french fries). These results suggest that the systematic price variation at the customer–product level is not driven by goods that are likely or unlikely to be bundled.

To study the importance of observables in explaining margin differences, I add controls for average volume (i.e., greater volume is associated with lower prices) and ZIP Code in Column 3 of Table 2. These variables only increase the R^2 to 0.67, suggesting that either other factors are explaining the price differences that are systematic at the customer–product level or the specification is not flexible enough.[8] A similar conclusion is reached when repeating this analysis on the subsample of the top 7 products (Column 6 of Table 2). This motivates me to conduct a variation of this analysis where I cluster customers based on their similarity in observables, and use cluster fixed effects to control for observables in a more flexible way. Specifically, for each product in the top 7 subsample, I cluster customers based on their similarity in a set of variables: indicators for restaurant type (i.e., pizza restaurant, meat restaurant, fish restaurant, or other), average quantity per transaction, and the fraction of a customer's total expenditure throughout the sample period that was spent on the product. I make use of the k-medoid clustering algorithm using the Gower distance to measure dissimilarity across customers (Friedman et al., 2001), and compute $k = 10$ clusters of customers per product (see Appendix B for details).[9] Column 7 of Table 2 presents the results with cluster fixed effects, and shows that the cluster fixed effects can explain a large share of the variance in price–cost margins that is systematic at the customer–product level, suggesting the importance of observables in explaining price differences.[10]

[8] Dropping the quantity percentile variable from Column 3 decreases the R^2 to 0.668.

[9] I use the Gower distance to accommodate the use of continuous and discrete variables. I discuss the choice of $k = 10$ clusters in Section 4.

[10] And also unobservables that can be predicted based on observables.

Observation 1. *There is significant variation in prices and margins at the customer–product level.*

Table 3 studies how the average price paid by a customer for a given product throughout the sample period varies as a function of a proxy for the value of paying a low price (i.e., an observation is a customer–product combination). This proxy is defined as the fraction of a customer's total expenditure throughout the sample period that was spent on each product, and it is defined at the customer–product level.[11] Table 3 shows that customers on average pay lower prices for products that represent a greater share of their overall input expenditure. The same pattern holds when controlling for a measure of transaction size, or when restricting attention to highly transacted products.[12] This result suggests that sellers understand that customers may be more price sensitive for products that represent a larger share of their overall input expenditure, and sellers practice price discrimination accordingly.[13]

Observation 2. *Prices are lower for products that represent a significant portion of a customer's input expenditure.*

Table 1 provides information about breakups (i.e., the event of a customer breaking the relationship with the distributor for a particular product).[14] Panel D shows that breakups are infrequent in that they happen only after 4.4 percent of the transactions. Panel E shows that the frequency of breakups does not significantly change when restricting the sample to products that are highly transacted and are not seasonal (i.e., the top 7 products include items like flour, shortening, and trash bags). Lastly, Panel C shows that in most cases customers do not break multiple relationships with the distributor at the same time. The average number of breakups per customer–week combination was 0.493 (which, given the average number of products per customer–week, gives a probability of a breakup of 0.052), with 90 percent of all customer–week combinations presenting two breakups or fewer.

Observation 3. *Breakups occur after 4.4 percent of the transactions.*

[11] That is, the proxy is defined as (customer's total expenditure on product j)/(customer's total expenditure).

[12] Table C2 in the Online Appendix replicates this exercise at the transaction level and recomputes the proxy for each transaction, excluding the expenditure of that transaction. The same pattern holds.

[13] Customers being more likely to search for products that represent a larger share of their overall input expenditure produces a similar correlation between price and share of expenditure via selection. The empirical model will consider both forces: price discrimination and selection.

[14] See the previous subsection for the definition of the breakup variable.

Example 1 (Customer with n completed transactions).

time:	t_1	t_2	$\cdots$	t_k	t_{k+1}	$\cdots$	t_n	t_{n+1}
price:	p_1	p_2	$\cdots$	p_k	p_{k+1}	$\cdots$	p_n	*?*
cost:	c_1	c_2	$\cdots$	c_k	c_{k+1}	$\cdots$	c_n	c_{n+1}

I next study how breakups at the customer–product level are associated with price changes. For ease of exposition, consider Example 1, which shows the data for a customer–product combination with n completed transactions at times $t_1, \ldots, t_n$. Let us say that the customer needed the product again at t_{n+1}, but did not purchase the product from the food distributor at t_{n+1}. That is, the customer chose to break its relationship with the distributor after t_n. One possible explanation for this breakup is that the customer requested a price quote from the distributor at time t_{n+1}, but the price quote was unattractive, leading the customer to immediately break its relationship with the distributor at t_{n+1} *without* completing the transaction at t_{n+1}.

I explore the empirical relevance of this explanation in what follows. However, I face the challenge of not observing price quotes in failed transactions (e.g., the price quote p_{n+1} in the failed transaction at t_{n+1} in Example 1) nor the exact timing of these failed transactions. To overcome this challenge, I exploit two facts. First, I observe the food distributor's marginal cost of selling the product at t_{n+1} as long as at least one customer purchased that product at t_{n+1}, and second, most customers purchase products in regular time intervals. Thus, for every completed transaction of a customer at time t_k, I approximate the time of the next transaction using the formula $t_k + av.\,time\,between\,transactions$, where $av.\,time\,between\,transactions$ is a measure of the time between transactions of a given customer–product combination.[15] Using the marginal cost data, I then approximate the price change faced by the customer at time t_{k+1} using $p_{k+1} - p_k \approx c_{k+1} - c_k$, where c_k is the food distributor's marginal cost of selling the product at time t_k. I henceforth call this measure *counterfactual price change* (i.e., $p_{k+1} - p_k \approx c_{k+1} - c_k$), which is defined from the perspective of a completed transaction at time t_k, and can be computed regardless of whether the customer broke its relationship with the distributor after t_k.

Table 4 presents a linear probability model of breakup for product j after the transaction at time t_k on an indicator of a positive counterfactual price change at time t_{k+1} (i.e., an indicator for a price increase at the future time when the customer was

[15]More specifically, I approximate t_{n+1} using $\hat{t}_{n+1} \approx t_n + \sum_{i=2}^{n}(t_i - t_{i-1})/(n-1)$, where n is the total number of transactions of that customer–product combination (i.e., t_n plus the average time between purchases measured at the customer–product level).

expected to return to purchase the good), where an observation is a completed transaction. That is, I study whether a customer's breakup after a completed transaction at time t_k can be explained by a price increase when the customer was expected to return to purchase the good at t_{k+1}. Columns 1 and 2 show that the counterfactual price change indicator is associated with breakups when using the full sample and also when restricting the sample to highly transacted products (i.e., top 7 products). Table C3 in the Online Appendix replicates this exercise but also includes a measure of past price increases, and the table shows that only the counterfactual price change indicator is associated with breakups when restricting the sample to highly transacted products (i.e., top 7 products).

The last two columns of Table 4 repeat the breakup analysis using a different dependent variable: time to next purchase. For each completed transaction, this variable is defined as the number of days before the customer returned to purchase the product. Because I only observe transactions that took place during the sample period, the variable time to next purchase is right censored.[16] Because the censoring is of greater significance for transactions that took place near the end of the sample period, I restrict the analysis to transactions that took place in the first half of the sample period. The estimates mirror the first two columns of Table 4. Columns 3 and 4 show that the counterfactual price increase measure is associated with longer times between transactions, with the longer times between transactions likely caused by breakups. Table C3 in the Online Appendix replicates this exercise but also includes a measure of past price increases, and similar to the findings in the previous paragraph, the table shows that only the counterfactual price change indicator is associated with breakups when restricting the sample to highly transacted products (i.e., top 7 products).[17] Combined, the evidence in Table 4 suggests that if a customer faces a price increase at a given point in time, the customer breaks the relationship at that point without completing a transaction.

Observation 4. *The data are consistent with price increases immediately triggering breakups.*

Lastly, Table 1 provides information about the frequency of price changes. Panel

[16] In this part of the analysis, I impute information for the last transaction of each customer–product combination in the sample as follows. I define the time to next purchase variable as the number of days between the date of the last observed transaction and the final date in the sample period.

[17] Table C4 in the Online Appendix replicates this exercise but restricts the sample to customer–week combinations with fewer than two breakups (i.e., the 90th percentile of the number of breakups at the customer–week level) to alleviate the concern of breakups caused by bankruptcy. The results do not significantly change.

D shows that customers faced price changes (defined at the customer–product level) in fewer than 7 percent of the transactions. When conditioning on a marginal cost change, the empirical probability of a price change increases to 24.3 percent. When restricting the sample to highly transacted products, these numbers increase to 14.7 percent when considering all transactions and to 51.5 percent when conditioning on transactions with cost changes.

Observation 5. *Cost changes do not always trigger price changes.*

Discussion

The patterns in the data are consistent with price discrimination and search costs. Consistent with price discrimination are the extent of price variation that is systematic at the customer–product level and the fact that customers who have greater incentives to find a low price on average pay less. A search model can explain the fact that customers respond to price increases by breaking their relationships and are otherwise unlikely to do so—i.e., customers respond to price when the price exceeds their reservation price, but do not otherwise respond to price. A search model can also explain the fact that prices on average decrease when the customer has greater incentives to find a low price—i.e., incentives to find a low price can be thought of as a measure of propensity to search.

In what follows, I analyze the data through the lens of a search model where sellers practice price discrimination. The model is able to reproduce the stylized facts discussed in this section and makes further predictions, which I do not rely on in estimation and which I can compare to patterns in the data to evaluate model fit.

3 Model

In this section, I develop a search model that can replicate the stylized facts discussed in the previous section. Each product is modeled in isolation. I consider an environment with N buyers and M sellers, which I am going to think of as firms (hence, the use of the "it" pronoun). The sellers are ex-ante identical, and buyers are heterogeneous both in their demand functions and search costs. The sellers offer a homogeneous good j that buyers wish to purchase, and have the ability to quote personalized prices.

Each period, buyer i receives a price quote, p, from the seller with whom it last did business. The buyer decides whether to complete the transaction at the quoted price or instead break the relationship and search sequentially for a better price quote elsewhere.

The buyer makes the breakup decision using i) its search cost, which determines how costly it is to obtain another price quote from an alternative seller, and ii) its beliefs about the price distribution it will face when searching.

A seller decides the price it will quote buyer i based on its marginal cost as well as buyer i's primitives (i.e., demand function and search cost).[18] Sellers are ex-ante identical but are ex-post heterogeneous in that they draw different marginal costs from a stationary cost distribution.[19] In equilibrium, each buyer will face a buyer-specific distribution of price quotes, where the within-buyer price heterogeneity is caused by the cost differences across sellers. A similar model was studied by Reinganum (1979).

An equilibrium in this game is defined as a set of strategy functions, s^*, such that each player i is maximizing its value given s^*_{-i}, and the beliefs of all buyers and sellers are consistent with s^*. The next subsections analyze the equilibrium strategies of buyers and sellers.

With respect to modeling assumptions, I abstract away from long-term contracts. This assumption is motivated by the observations that prices respond to cost changes and buyers respond to price changes by breaking their relationship with the distributor. That is, there is no evidence of the distributor being committed to a price or the buyers being committed to purchasing from the distributor. I also do not explicitly model bargaining, but one can reinterpret the model as sellers making take-it-or-leave-it offers as a function of the bargaining strength of the buyers, which in the context of the model is their propensity to search (see Allen et al., 2014a,b,c for a similar discussion).

Buyer's Problem

Buyers have systematic differences in their search costs and demand functions. Buyer i's average search cost and demand function for product j are given by sc_{ij} and $q_{ij}(p; \xi_{ij}) = \exp\{\alpha p + \xi_{ij}\}$, respectively, where ξ_{ij} is a demand shock that varies across transactions and is an i.i.d. draw from a buyer–product specific distribution, H_{ij}. Each period, buyer i receives a price quote, p, from the seller with whom it last did business.

Whenever facing a seller, buyer i must decide whether to accept the seller's price quote or to decline it and search for a better price quote elsewhere. When facing price

[18] I assume sellers observe each buyers' primitives. This assumption is motivated by Table 2, which suggests that about 96 percent of the price differences across buyers is time invariant.

[19] Unfortunately, I do not have data to test for whether firms are ex-ante symmetric or a symmetric equilibrium is being played. This is a weakness of the analysis.

quote p, buyer i chooses to search iff

$$\delta E[V_{ij}] - sc_{ij} + \varepsilon \geq CS_{ij}(p; \xi_{ij}),$$

where $E[V_{ij}]$ is the expected value of matching with a new seller, δ is a discount factor, $sc_{ij} - \varepsilon$ is buyer i's realized search cost for product j, and

$$CS_{ij}(p; \xi_{ij}) \equiv \int_p^\infty q_{ij}(s; \xi_{ij}) ds$$

is the buyer's surplus at price p given demand shock ξ_{ij}. I assume that whenever a buyer matches with a seller, it draws a new demand shock ξ_{ij} from the distribution of demand shocks, H_{ij}. Under this assumption, the expected value of matching with a new seller, $E[V_{ij}]$, does not depend on realized values of ξ_{ij}, which will simplify estimation.

Customer i's search cost for product j has a time-invariant component, sc_{ij}, and a buyer–seller specific component, ε, which is an i.i.d. draw from a distribution with a log-concave cumulative distribution function G. That is, whenever a buyer matches with a seller, it faces a new draw ε from the distribution G. This assumption is only made to provide smoothness to the search decision process and implies that the buyer will choose to search with probability $1 - G(CS_{ij}(p; \xi_{ij}) - \delta E[V_{ij}] + sc_{ij})$ when faced with price quote p.[20]

Lastly, the expected value of matching with a new seller, $E[V_{ij}]$, is given by

$$E[V_{ij}] = \int \max \{CS_{ij}(p; \xi_{ij}),\ \delta E[V_{ij}] - sc_{ij} + \varepsilon\} \, dG(\varepsilon) dH_{ij}(\xi_{ij}) dF_{ij}(p), \qquad (1)$$

which considers the option value of continued search if the price quote offered by the new seller results in a payoff that is inferior to the value of searching (i.e., if $CS_{ij}(p; \xi_{ij}) < \delta E[V_{ij}] - sc_{ij} + \varepsilon$). In the equation, the expectation is with respect to the buyer–seller search cost component (ε), the demand shock ξ_{ij} (with cumulative distribution H_{ij}), and the new seller's price quote (with cumulative distribution F_{ij}). The expectation with respect to the price quote makes use of the buyer's equilibrium beliefs about the distribution of price quotes it faces. The expectation with respect to the demand shock ξ_{ij} captures the assumption that the buyer draws a new demand shock when rematching with a seller.[21]

[20]The estimation results show that the buyer–seller specific component, ε, represents less than 1 percent of the overall variance of search costs.

[21]Although the model assumes that a buyer's demand shock can vary from seller to seller, the

Seller's Problem

A seller of product j sets its price for buyer i by maximizing its expected profit,

$$\max_p \; G(CS_{ij}(p;\xi_{ij}) - \delta E[V_{ij}] + sc_{ij})q_{ij}(p;\xi_{ij})(p-c), \tag{2}$$

where $G(CS_{ij}(p;\xi_{ij}) - \delta E[V_{ij}] + sc_{ij})$ is the probability that buyer i will accept the price quote p, and c is the seller's marginal cost. When solving its pricing problem, the seller observes $q_{ij}(p,\xi_{ij})$ and sc_{ij}. The seller also has equilibrium beliefs about the distribution of price quotes faced by the buyer and, thus, equilibrium beliefs about the buyer's expected value of matching with a new buyer, $E[V_{ij}]$.

The first-order necessary condition generates the following price rule (where ξ_{ij} was dropped for notational ease):

$$p^* = c - \frac{G(CS_{ij}(p^*) - \delta E[V_{ij}] + sc_{ij})q_{ij}(p^*)}{G(CS_{ij}(p^*) - \delta E[V_{ij}] + sc_{ij})q'_{ij}(p^*) - g(CS_{ij}(p^*) - \delta E[V_{ij}] + sc_{ij})q_{ij}(p^*)^2}. \tag{3}$$

The optimal price balances the trade-off between the likelihood of serving the customer—which decreases in price—and the profit earned conditional on serving the customer—which increases in price for prices below the monopoly price. I discuss properties of the optimal price as well as other model predictions in the next subsection.

Predictions of the Model

I next discuss some properties of the best-response functions and the equilibrium objects. First, one of the properties of the solution to the sellers' problem is that when the seller faces a buyer with a greater net value of matching with a new seller (i.e., $\delta E[V_{ij}] - sc_{ij}$), the seller best responds by setting a lower price (all else equal). That is, when the threat of search is greater, the seller sets a lower price to reduce the likelihood of search. This property of the best response function of sellers plays a key role in the estimation of the model.

Second, in equilibrium, the net value of matching with a new seller (i.e., $\delta E[V_{ij}] - sc_{ij}$) is decreasing in the search cost, sc_{ij}. Combined with the observation in the previous paragraph, this implies that an increase in a buyer's search cost leads the buyer to face greater price quotes (all else equal).

Third, the rate at which a buyer searches in equilibrium is decreasing in the buyer's

estimation results suggest that the size of this variation only amounts to about 11 percent of the overall variation in ξ_{ij} (i.e., the fraction of the overall variance in ξ_{ij} that is within customer–product).

search cost (all else equal). The rate at which a buyer searches can be expressed as the expected search probability, which is the ex-ante probability that the buyer will search when meeting a seller. This probability is given by $E[1-G(CS_{ij}(p;\xi_{ij})-\delta E[V_{ij}]+sc_{ij})]$, where the expectation is with respect to price, demand shock, and search cost (i.e., all the variables affecting the search decision).

These properties are summarized in the following proposition and are further discussed in the Appendix. Proposition 1.(i) will play a key role in the estimation of the model, and Proposition 1.(iii) will be tested using the estimated prediction of the model for breakups, which are not used in estimation.

Proposition 1 (Model predictions)**.**

i) A seller's price is decreasing in the buyer's net value of matching with a new seller, $\delta E[V_{ij}] - sc_{ij}$.

ii) A seller's equilibrium price quote is increasing in the buyer's search cost.

iii) The rate at which a buyer searches in equilibrium is decreasing in its search cost.

4 Model Estimation

Estimation Sample

The model is able to reproduce all the stylized facts discussed in Section 2 except for the fact that cost changes do not always trigger price changes (Observation 5). I restrict the sample in two ways to abstract away from the issue of price stickiness. First, I focus on the sample of highly transacted products (i.e., top 7)—which have a higher rate of price revisions—and, second, I aggregate the data to the customer–product–month level. That is, if a customer has more than one transaction for a particular product in a given month, I aggregate these by computing the total quantity and a quantity-weighted average price and cost across all these transactions. The resulting estimation sample presents a rate of price changes (conditional on cost change) of 66 percent (up from the sample-wide figure of 24 percent).

After imposing these sample restrictions, the estimation sample has 4,479 transactions (or customer–product–month combinations) and 772 unique customer–product combinations. The average price is 20.43, which is almost identical to the average price in Panel E of Table 1, and the average quantity is 6.99, which is about three times higher than the average quantity in Panel E of Table 1, likely reflecting the temporal aggregation of the sample. The average (median) number of transactions per customer–

product combination is 5.8 (6), and the number of transactions per customer–product combination ranges between 2 and 8 transactions (percentiles 1 and 99, respectively).

Estimation Overview

I estimate two sets of primitives: the demand functions and the search costs. The demand functions and the search costs vary at the customer–product level. Each customer–product combination ij faces a search cost with two components: a time-invariant component, sc_{ij}, and a customer–product–seller specific search cost shock, $\varepsilon \sim G$. Throughout the estimation, I focus on estimating the time-invariant component of search costs and assume that the distribution of search cost shocks, G, is standard normal.[22] In what follows, I provide an overview of the estimation steps—see Appendix B for a detailed discussion.

First, I use data on quantity and price for each customer–product combination to estimate demand functions. The specification of the demand function of customer–product ij is given by $q_{ij}(p) = \exp\{\alpha p_{ijt} + \xi_{ijt}\}$, where $\xi_{ijt} = \xi_{ij} + \Delta\xi_{ijt}$ is a demand shock that varies across transactions t. To estimate the model, I assume that the innovation in demand shocks satisfy the moment conditions: $E[z_{ijtl}\Delta\xi_{ijt}] = 0$, $\forall l \in \{1 \dots L\}$. These moment conditions state that the innovation in demand shocks and a set of L instruments are orthogonal random variables. Based on these moment conditions, I estimate the parameters of the demand functions (price coefficient and customer–product fixed effects) using a two-stage least squares estimator.

With demand estimates in hand, I turn to estimating search costs. Search costs can be estimated using a full-solution method, where the equilibrium of the model is computed for every trial vector of search costs, and the parameter search is over the vector of search costs that can best explain the data. Although a full-solution method is feasible, it is computationally impractical. I employ an alternative approach—discussed in Flinn and Heckman (1982)—and estimate search costs using a two-step method that only requires the econometrician to compute the equilibrium of the game once. The method exploits two key properties of the model. First, the price-optimality condition in equation (3) depends on cost (c_j), demand (q_{ij} and CS_{ij}), and on the equilibrium object $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$, which I call the equilibrium net value of matching with a new seller (see equation (1) for a definition of $E[V_{ij}]$). Proposition 1 shows that the optimal price set by a seller is decreasing in $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$. That is, conditional

[22] When discussing the results, I show that the model estimates do not change significantly when making alternative parametric assumptions.

on observables, the model predicts lower prices for buyers with a greater value of θ_{ij} (i.e., a greater propensity to search). Second, $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$ is strictly monotonic in sc_{ij}. That is, there is a unique value of sc_{ij} that is consistent with θ_{ij}. The two-step method—which I discuss in detail in Appendix B—consists of first estimating $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$ for every customer–product ij using a GMM estimator, and then recovering sc_{ij} in a second step using the estimates $\hat{\theta}_{ij}$ and the structure of the model.

To estimate the equilibrium net value of matching with a new seller of each customer–product combination ij, I use the price-optimality condition in equation (3) to form the moment condition

$$E\left[p_{ijt} - c_{jt} - \nu_{ijt} + \frac{G(CS_{ij}(p_{ijt}) - \theta_{ij})q_{ij}(p_{ijt})}{G(CS_{ij}(p_{ijt}) - \theta_{ij})q'_{ij}(p_{ijt}) - g(CS_{ij}(p_{ijt}) - \theta_{ij})q_{ij}(p_{ijt})^2}|NS\right] = 0 \tag{4}$$

for every customer–product combination ij, where subscript t is used to enumerate the transactions of customer–product combination ij. These moment conditions require that the observed prices be optimal, and Proposition 1 shows that prices are a monotonic function of θ_{ij} (all else equal). That is, θ_{ij} is identified from variation in prices conditional on both demand and cost. Estimating θ_{ij} can be thought of as estimating a customer–product level fixed effect that solves the moment condition above.

Moment condition (4) makes use of the demand estimates and the distributional assumption of the search cost shocks, $G \sim N(0,1)$. Demand estimates enter the price-optimality condition through $q_{ij}(p_{ijt})$, $q'_{ij}(p_{ijt})$, and $CS_{ij}(p_{ijt})$, where the functional form of demand implies that $CS_{ij}(p_{ijt}) = -\exp\{\alpha p_{ijt} + \xi_{ijt}\}/\alpha$. Inside the brackets of moment condition (4), I also include a structural error term, ν_{ijt}, which is assumed to be a component of the seller's marginal cost that is unobserved to the econometrician (i.e., marginal cost is assumed to be $c + \nu$, where c is the marginal cost reported in the data). This structural error term helps rationalize the fact that prices may not satisfy the first-order condition exactly. Because the estimation sample only includes completed transactions (i.e., prices that did not induce search), and the value of ν_{ijt} affects prices, the moment condition is conditional on the event of "no search" (NS). In practice, conditioning on the event of "no search" only imposes restrictions on the value of $E[\nu_{ijt}|NS]$, as all the other objects that enter the moment condition are either data or parameters.

Instead of imposing a parametric restriction on the distribution of ν, I approximate $E[\nu_{ijt}|NS]$ using the fitted values of ν. For every evaluation of the objective function, I approximate $E[\nu_{ijt}|NS]$ using the average value of ν across all transactions involving product j (i.e., the product that corresponds to customer–product combination ij).

Using an approximation is reasonable in this context for two reasons. First, the overall probability of breakup is small (hence, the selection problem is mild), and second, the estimated values of ν are small relative to the observed component of marginal cost. The approximation also has the benefit of not relying on parametric assumptions.

The second step of the two-step method makes use of the estimates of the equilibrium net value of matching with a new seller, θ_{ij}. Notice that given θ_{ij}, the price-optimality condition in equation (3) provides the seller's optimal price as a function of the marginal cost and demand shock.[23] Based on equation (3), one can thus simulate the equilibrium price distribution for each customer–product combination ij making use of the estimate $\hat{\theta}_{ij}$ as well as simulated cost and demand shock draws. Using the equilibrium distribution of price quotes of every customer–product combination ij, I then compute the expected value of matching with a new seller, $E[V_{ij}]$, using equation (1). Given $\hat{E}[V_{ij}]$ and $\hat{\theta}_{ij}$, the estimate for sc_{ij} for customer–product combination ij is given by $\hat{sc}_{ij} = \delta\hat{E}[V_{ij}] - \hat{\theta}_{ij}$ (where this equation makes use of the identity that defines the equilibrium net value of matching with a new seller, i.e., $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$).

Because I first estimate demand, and then estimate search costs given the demand estimates, I follow the procedure outlined in Cameron and Trivedi (2005, p. 200) to adjust standard errors for the sequential estimation of these objects. Confidence intervals for search costs and welfare objects of interest are computed using 100 bootstrap samples.

Identification, Econometric Challenges, and Parametric Assumptions

The identification of the demand function follows from standard arguments. The demand specification includes price and customer–product fixed effects. The instruments for price are the product-specific unit cost interacted with four restaurant-type dummies (i.e., pizza, seafood, meat, other). Equation (3) suggests that the unit cost is a relevant instrument, as the sellers' pricing rule depends on this cost—the F-statistic of the first stage is 31.84. The model also suggests that these instruments are valid, as the customers' decisions do not depend on the unit cost other than through the effect of the unit cost on price.

The identification of search costs follows from two key properties of the model (see

[23] In the model, firms are ex-ante symmetric, but they face different realized values of marginal cost. This ex-post asymmetry across firms gives rise to within-customer–product price heterogeneity that is captured by an equilibrium price distribution at the customer–product level.

Proposition 1). First, buyers with a greater propensity to search are quoted lower prices (all else equal), where a buyer's propensity to search can be summarized by the equilibrium net value of matching with a new seller. Second, there is a monotonically decreasing mapping between a buyer's propensity to search and the buyer's search cost (see proof of Proposition 1.(ii)). These predictions combined suggest that buyers with lower search costs face lower price quotes in equilibrium (all else equal). Search costs are thus identified based on the prices paid by each buyer.

The main econometric challenge when estimating the demand functions and implementing the two-step method for the estimation of search costs is that the precision of the estimates depends on the availability of data on multiple transactions per customer–product combination. Because the sample period is only eight months and I aggregate the data to the customer–product–month level, the median and mean number of transactions per customer–product combination in the estimation sample is 6 and 5.8, respectively. As a way to alleviate the small-sample problem, I use the idea discussed in Section 2 of grouping customer–product combinations into clusters of customers that have similar observables. I then make the assumption that all the relevant primitives of the model are the same for all members of a cluster. As discussed in Section 2, the cluster fixed effects can explain a large fraction of the variation in price-cost margins that is systematic at the customer–product level. This suggests that replacing customer–product fixed effects with cluster fixed effects should not result in a significant loss in terms of model flexibility.

For each product, I cluster customers using five observable variables: the fraction of a customer's total expenditure throughout the sample period that was spent on the product, average quantity per transaction, an indicator for pizza restaurant, an indicator for meat restaurant, and an indicator for fish restaurant. As explained in Section 2, I make use of the k-medoid clustering algorithm using the Gower distance to measure dissimilarity across customers (Friedman et al., 2001), and compute $k = 10$ clusters of customers per product (see Appendix B for details).[24] After clustering, the median and mean number of transactions per cluster is 100 and 104.1, respectively, and the number of transactions per cluster ranges between 10 and 195 transactions (percentiles 1 and 99, respectively). Three clusters with fewer than three observations were dropped from the estimation sample. Henceforth, I use the term customer–product combination to refer to a cluster of customer–product combinations.

[24] I found that the choice of $k = 10$ clusters strikes a good balance between allowing the model to flexibly capture customer heterogeneity and guaranteeing that each cluster has a sufficient number of transactions.

The model relies on two main parametric assumptions: first, an assumption on the functional form of the demand function (q_{ij}), and second, an assumption on the distribution of search cost shocks, G. The empirical model assumes log-linear demand functions and a standard normal distribution for the search cost shocks. The assumption of log-linear demand functions also has implications for the customer–product specific distribution of demand shocks, H_{ij}, as H_{ij} is estimated based on the distribution of fitted demand residuals (i.e., $\hat{\xi}_{ij} + \Delta\hat{\xi}_{ijt}$). When discussing model estimates, I show that the estimates of the buyers' equilibrium net value of matching with a new seller do not significantly change when using an alternative parameterization, suggesting that the parametric choices in the baseline model are not driving the results.

The model also assumes that the demand shocks of a customer for a particular product are i.i.d. over time and across sellers. The fact that a customer draws a new demand shock every time it matches with a new seller (i.e., i.i.d. across sellers) simplifies estimation, as this assumption implies that $E[V_{ij}]$ (and hence $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$) does not depend on realizations of demand shocks. But by the same token, this assumption implies that imprecise estimates of the distribution of demand shocks of a customer–product combination affect the precision of the search cost estimates in the second step of the two-step method, as computing the value of $E[V_{ij}]$ requires solving an integral with respect to that distribution of demand shocks (see equation (1)).

The assumption that demand shocks are i.i.d. over time rules out within-customer–product correlation over time in demand shocks, which could exist due to seasonality of demand. Although this assumption could be relaxed to accommodate seasonality, the benefits of doing so are limited given that the estimation sample restricts attention to products that do not exhibit seasonality (i.e., flour, shortening, bleach, trash bags, plastic forks, napkins, and canned peppers).

No parametric assumption is imposed on the distribution of the buyer-specific average search costs, sc_{ij}—though there is an indirect relationship between the estimates for the average search costs and the other parametric assumptions in the model. Not imposing a parametric assumption on the distribution of buyer-specific average search costs has the benefit that the model is able to accommodate arbitrary patterns of correlation between demand shocks and search costs as well as arbitrary patterns of correlation in search costs within buyer across products.

Although the model is able to accommodate rich patterns of customer heterogeneity, the model does not allow for systematic seller heterogeneity (i.e., sellers are ex-ante identical). For example, the model does not allow for a seller-specific component of demand (e.g., capturing seller differentiation in service quality). In principle, the model could be extended to accommodate such patterns of seller heterogeneity. However, identifying such sources of seller heterogeneity would require having access to a richer dataset with data from multiple food distributors.

Estimation Results and Model Fit

The estimates of the model are reported in Table 5 and Figure 2. Table 5 shows that the estimate for the price coefficient of demand is -0.076. This estimate implies an average demand elasticity of -1.56, suggesting that buyers are responsive to prices even conditional on choosing to complete a transaction.[25] Figure 2 (Panel A) reports the cumulative distribution function of the customer–product level estimates of the buyers' equilibrium net value of matching with a new seller (or $\{\theta_{ij}\}$ in the notation of the previous subsections). The figure shows heterogeneity across customer–product combinations, with the net value of matching with a new seller being near zero in some cases and as large as 200 dollars in other cases, and taking an average value of 32.9 dollars (or 1.61 times the average price in the estimation sample). As discussed above, these estimates depend on two functional form assumptions: the demand function and the distribution of the buyer–seller specific component of search costs, $\varepsilon \sim G$. In Figure C3 in the Online Appendix, I show that the estimates of the buyers' equilibrium net value of matching with a new seller do not significantly change when using an alternative parameterization.

Estimates of the search costs—using the methodology described in the previous subsections—are reported in Figure 2 (Panel B). The mean and median search costs are 40.8 and 15.9 dollars, respectively. The search cost estimates are reasonable in that search is not implausibly costly (in absolute terms). The estimation results also show that the buyer–seller specific component of search costs, $\varepsilon \sim G$, represents less than 1 percent of the overall variance of search costs. That is, most of the variation in search costs is time-invariant variation across customer–product combinations.

With respect to model fit, Figure 3 shows the empirical distribution of prices as well as the distribution of prices predicted by the model for each data point. The figures show that the distributions of model predictions and actual data are similar, and there is a 96 percent correlation between actual and predicted prices.

The model also predicts that 69 percent of the variance of price–cost margins is systematic at the customer–product level. This number is lower than what is observed

[25]The lower and upper bounds of the 95-percent confidence interval for the mean elasticity are -1.88 and -0.82, respectively.

in the data (98.7 percent), but the model is still able to capture that most of the variation in margins is time-invariant variation across customer–product combinations. The model also replicates the fact that customers who have greater incentives to find a low price pay lower prices. Table C5 in the Online Appendix shows that the search costs estimates are strong predictors of the prices and price–cost margins paid by customers, with an estimated correlation between prices and search costs of 0.37.

Although the breakup data were not used in estimation, the model still presents a good fit along this dimension. The model predicts that an average of 5.8 percent of transactions will result in a breakup, which is similar to the 4.4 percent breakup rate in the data. Proposition 1 predicts that greater search costs should lead to a lower probability of search. To contrast this prediction with the model estimates, I decompose the search cost distribution between customer–product combinations with and without a breakup in the data. As predicted by the model, I find that the average search cost of customer–products with an observed breakup was 29.8 dollars, whereas it was 96.9 dollars for those without a breakup.

5 Price Discrimination and Market Efficiency

Using the estimates of the model, I analyze the impact of price discrimination on efficiency and welfare. The first step is computing the industry equilibrium when firms set a uniform price for all buyers (conditional on their marginal cost). In the uniform price case, the sellers' problem of choosing their optimal prices for product j becomes

$$\max_{p} \sum_{i} \int G(CS_{ij}(p;\xi_{ij}) - \theta_{ij}^{NoPD}) q_{ij}(p;\xi_{ij})(p-c) dH_{ij}(\xi_{ij}), \tag{5}$$

where $\theta_{ij}^{NoPD} \equiv \beta E[V_{ij}^{NoPD}] - sc_{ij}$ is the net value of matching with a new seller of customer–product combination ij in the uniform price equilibrium (see equation (1)), $H_{ij}(\xi_{ij})$ is the distribution of demand shocks of customer–product combination ij, and the sum is over every buyer i purchasing product j.[26] In this case, sellers cannot customize prices and thus maximize an expected profit function, where the expectation is with respect to the distribution of buyers as well as the search probability. As in the price discrimination case, price heterogeneity across sellers is driven by cost variation across sellers.

Given both the uniform price and price discrimination equilibria predicted by the

[26] Although equilibrium uniqueness is not theoretically guaranteed, the uniform price equilibrium I present below was robust to various initial points.

model estimates, I compare several outcomes to measure the effects of price discrimination. First, I compare the expected profit of serving each buyer under both equilibria. This comparison measures the profit gains of price discrimination from an ex-ante perspective (i.e., including transactions that were and were not completed). Next, I compare the expected value of a buyer (see equation (1)) as well as the expected customer surplus in a completed buyer–seller interaction under both equilibria. The expected value of a buyer measures the expected customer surplus accounting for costly search, whereas the customer surplus in a completed transaction excludes the resources spent searching and focuses on efficiency in completed transactions. Throughout, expectations are with respect to the equilibrium price distributions faced by each buyer.

Figure 4 shows the equilibrium price distributions for all buyers—organized by product—under both equilibria. As expected, some buyers are better off with price discrimination, whereas others are worse off. Below I empirically show that buyers with below (above) average search costs are better (worse) off with price discrimination, as they are on average quoted lower (higher) prices in the price discrimination case. These results are in line with Proposition 1.(ii).

Sellers

I first consider the impact of price discrimination on the sellers' expected profits. I do this in two steps. First, I compute the gains of practicing price discrimination when all other rivals behave according to the no price discrimination equilibrium. This comparison isolates the profit gains of price flexibility, as the exercise holds the rivals' behavior fixed. Second, I compare the equilibrium expected profits under each pricing regime. This second comparison incorporates the price flexibility effects as well as the competition intensity effect of price discrimination (or the equilibrium feedback effect).

To understand the competition intensity effect of price discrimination, let us start with the no price discrimination equilibrium. Consider the case of a seller who deviates from a no price discrimination regime and chooses to lower its price of product j for buyer i. If all other sellers also lower their prices for buyer i (and thus match the seller's price discrimination strategy), then buyer i is faced with greater incentives to search because of the lower prices. These greater search incentives lower a seller's probability of serving a customer at any given quoted price, which in turn lowers the seller's expected profit. Consequently, with price discrimination sellers enjoy the benefits of tailoring prices, but at the same time, price discrimination may boost search incentives, which may partially reverse these profit gains. This is the competition intensity effect of price discrimination. Conversely, price discrimination may also reduce a buyer's incentives to search, which will happen if instead all the sellers quote the buyer higher prices. In this alternative case, price discrimination lowers competition intensity, implying that the competition intensity effect magnifies (rather than reverses) the benefits of price flexibility.

Table 6 (Panel A) shows the results of analyzing the impact of price discrimination on the sellers' expected profits. The columns labeled NoPD and PD indicate the expected profit averaged across buyers under the uniform price and price discrimination equilibria, whereas the column labeled Dev indicates the expected profit when one (and only one) firm deviates from the uniform price equilibria and practices price discrimination. As discussed above, the difference between Dev and NoPD isolates the price flexibility profit gains, as all rivals' strategies are held fixed at the uniform price equilibrium strategies. The benefits of price flexibility range from 0.2 to 17.8 percent across the 7 products and suggest that price tailoring has a first order effect on profit. The difference between PD and NoPD measures the profit gains of price discrimination incorporating the price flexibility gains as well as the changes in competition intensity (i.e., the equilibrium feedback effect). The table shows that the price flexibility gains (Dev$-$NoPD) are on average magnified in equilibrium, suggesting that price discrimination on average softens competition; this is captured by the average decrease in the expected value of matching with a new seller, $E[V_{ij}]$, caused by price discrimination. The benefits of price discrimination in equilibrium range from 0.3 to 52.1 percent across the 7 products, with an overall profit increase of 5.5 percent (and a median profit increase across customer–product combinations of 12.1 percent). The largest profit gains are in the products with the greatest amount of variation in search costs (e.g., products 2 and 5).

Buyers

I next compare the benefits of price discrimination for buyers in two steps. I first analyze how price discrimination impacts three buyer-level market outcomes: price quotes, search intensity, and prices in completed transaction. Guided by these results, I then study how price discrimination impacts buyers' expected value of matching with a seller, $E[V_{ij}]$, as well as outcomes in completed transactions. The analysis on completed transactions will shed light on how price discrimination affects market efficiency.

Table 7 shows the impact of price discrimination on a series of market outcomes as a function of the buyers' (standardized) search costs. An observation in the table is a customer–product combination, and the outcome variables are also at the customer–product level, computed using data generated by the model. The first column shows that price discrimination on average increases the price quotes faced by buyers with above average search costs, whereas it decreases price quotes for buyers with below average search costs. The greater a buyer's search cost, the greater the price increases caused by price discrimination. The second column shows that buyers with above (below) average search costs are more likely (less likely) to search with price discrimination. Lastly, the third column shows that the changes in price paid by buyers as a consequence of price discrimination are increasing in search cost, though the slope is lower than in the first column. The lower slope in the third column can in part be explained by the fact that low search cost buyers on average pay the same price in both equilibria (see Figure C4 in the Online Appendix). Although buyers with below average search costs face greater price quotes in the uniform price equilibrium, it is cheap for them to compensate for these greater price quotes by intensifying their search activity. This greater search activity in the uniform price equilibrium—which drives the results in Table 7 (Column 2)—leads low search cost buyers to pay similar prices regardless of the pricing regime. This is in contrast to high search cost buyers, who pay higher prices in the price discrimination equilibrium both because they face higher price quotes and they find it costlier to search. As discussed below, these results have key implications for the efficiency effects of price discrimination.

Table 6 (Panel A) shows the impact of price discrimination on the buyers' expected value of matching with a new seller, $E[V_{ij}]$. The table shows that price discrimination on average reduces the value of matching with a new seller, with product-level effects ranging from -36.7 percent (product 2) to -1.1 percent (product 4) and an overall effect of -11.1 percent. This effect can be understood from the results in Table 7, which suggests that buyers on average pay lower prices in the uniform price equilibrium. That is, buyers with above average search costs face lower prices in the uniform price equilibrium, which implies that they are hurt by price discrimination. Buyers with below average search costs raise their search intensity in the uniform price equilibrium, implying that the prices they pay are not significantly impacted under either equilibria. The decrease in the expected value of matching with a new seller translates into a decrease in competition intensity in the price discrimination equilibrium, which in part explains the effects on profits discussed in the previous subsection.

Table 6 (Panel B) shows the efficiency effects of price discrimination in completed transactions. The middle columns show that price discrimination decreases customer surplus on average, with an overall decrease of 16.3 percent. This result is driven by buyers with above average search costs who pay higher prices with price discrimination. The gains to buyers with below average search costs, who face lower price quotes in the price discrimination equilibrium, do not compensate for the efficiency loss among buyers with higher search costs, as buyers with below average search costs pay similar prices in both equilibria (see previous paragraph). When examining total surplus, I find that price discrimination lowers total surplus, thus hurting market efficiency. The overall effect of price discrimination on total surplus is an average decrease of 11.6 percent, but the effect of price discrimination is heterogenous across products, ranging from -37.7 to -1.7 percent.

In summary, these results suggest that the seller's ability to practice price discrimination can increase expected profits in economically meaningful ways in the presence of sufficient customer heterogeneity, even in environments with search frictions. Price discrimination is found to significantly decrease market efficiency. The key effect explaining the efficiency result is that buyers with above average search costs—who face higher price quotes in the price discrimination equilibrium—find it too costly to increase their search intensity in the price discrimination equilibrium to compensate for the higher price quotes. These results suggest that price discrimination can magnify the efficiency effects of search costs.

6 Concluding Remarks

Many markets exhibit heterogeneity in the prices paid by buyers, where these price differences reflect buyer heterogeneity. Wholesale markets are an example. In these markets, sellers often choose not to commit to posted prices in order to customize prices, and buyers must incur costs to discover prices as a result. The practice of customizing prices has implications for market efficiency.

I study a wholesale market for commodity goods where some buyers may pay up to 70 percent more than others for the same product on the same day. These price differences are found to be systematic across buyers, and I present evidence supporting search costs as the source of market power. Search costs are a concern because they reduce efficiency by limiting customer mobility to low-price sellers. Aside from the allocative inefficiencies generated by price discrimination, price discrimination in the presence of search costs has an effect on competition intensity (i.e., search incentives) that may magnify the efficiency concerns created by search costs alone.

Inspired by this evidence, I estimate a search model with heterogeneous buyers to measure the welfare implications of price discrimination in the presence of information frictions. I find that price discrimination has a negative effect on total surplus relative to the case where sellers set uniform prices, and price discrimination is found to increase sellers' expected profits by up to 52.1 percent. These effects on total surplus and profits are partially explained by price discrimination on average decreasing search incentives, and hence reducing the intensity of price competition. These results illustrate how price discrimination may magnify the efficiency costs of search frictions.

References

Allen, Jason, Robert Clark, and Jean-François Houde (2014a) "Price Dispersion in Mortgage Markets," *The Journal of Industrial Economics*, Vol. 62, pp. 377–416.

——— (2014b) "Search Frictions and Market Power in Negotiated Price Markets," Working Paper 19883, National Bureau of Economic Research.

——— (2014c) "The Effect of Mergers in Search Markets: Evidence from the Canadian Mortgage Industry," *American Economic Review*, Vol. 104, pp. 3365–96.

Armstrong, Mark and John Vickers (2001) "Competitive Price Discrimination," *The RAND Journal of Economics*, Vol. 32, pp. 579–605.

Aurichio, Bruno and Luis Henrique Bertolino Braido (2018) "Dynamic price competition in auto-insurance brokerage."

Beckert, Walter, Howard Smith, and Yuya Takahashi (2015) "Competitive Price Discrimination in a Spatially Differentiated Intermediate Goods Market."

Burdett, Kenneth and Kenneth L. Judd (1983) "Equilibrium Price Dispersion," *Econometrica*, Vol. 51, pp. 955–969.

Busse, Meghan and Marc Rysman (2005) "Competition and price discrimination in Yellow Pages advertising," *The RAND Journal of Economics*, Vol. 36, pp. 378–390.

Cameron, A Colin and Pravin K Trivedi (2005) *Microeconometrics: methods and applications*: Cambridge university press.

Corts, Kenneth S. (1998) "Third-Degree Price Discrimination in Oligopoly: All-Out Competition and Strategic Commitment," *The RAND Journal of Economics*, Vol. 29, pp. 306–323.

Dafny, Leemore S. (2010) "Are Health Insurance Markets Competitive?" *The American Economic Review*, Vol. 100, pp. 1399–1431.

De los Santos, Babur, Ali Hortaçsu, and Matthijs R Wildenbeest (2017) "Search with learning for differentiated products: Evidence from e-commerce," *Journal of Business & Economic Statistics*, Vol. 35, pp. 626–641.

Diamond, Peter A. (1971) "A Model of Price Adjustment," *Journal of Economic Theory*, Vol. 3, pp. 156–168.

Flinn, Christopher and James Heckman (1982) "New methods for analyzing structural models of labor force dynamics," *Journal of Econometrics*, Vol. 18, pp. 115–168.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani (2001) *The elements of statistical learning*, Vol. 1: Springer series in statistics New York.

Goldberg, Pinelopi Koujianou (1996) "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*, Vol. 104, pp. 622–654.

Grennan, Matthew (2013) "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices," *The American Economic Review*, Vol. 103, pp. 145–177.

Hendel, Igal and Aviv Nevo (2013) "Intertemporal price discrimination in storable goods markets," *The American Economic Review*, Vol. 103, pp. 2722–2751.

Holmes, Thomas J. (1989) "The Effects of Third-Degree Price Discrimination in Oligopoly," *The American Economic Review*, Vol. 79, pp. 244–250.

Hong, Han and Matthew Shum (2006) "Using price distributions to estimate search costs," *The RAND Journal of Economics*, Vol. 37, pp. 257–275.

Hortaçsu, Ali and Chad Syverson (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*, Vol. 119, pp. 403–456.

Jindal, Pranav and Peter Newberry (2018) "To bargain or not to bargain: The role of fixed costs in price negotiations," *Journal of Marketing Research*, Vol. 55, pp. 832–851.

Katz, Michael L. (1984) "Price Discrimination and Monopolistic Competition," *Econometrica*, Vol. 52, pp. 1453–1471.

——— (1987) "The Welfare Effects of Third-Degree Price Discrimination in Intermediate Goods markets," *The American Economic Review*, Vol. 77, pp. 154–167.

Leslie, Phillip (2004) "Price Discrimination in Broadway Theater," *The RAND Journal of Economics*, Vol. 35, pp. 520–541.

Marshall, Guillermo (2015) "Hassle Costs and Price Discrimination: An Empirical Welfare Analysis," *American Economic Journal: Applied Economics*, Vol. 7, pp. 123–146.

Mortimer, Julie Holland (2007) "Price Discrimination, Copyright Law, and Technological Innovation: Evidence from the Introduction of DVDs," *Quarterly Journal of Economics*, Vol. 122, pp. 1307–1350.

Reinganum, Jennifer F. (1979) "A Simple Model of Equilibrium Price Dispersion," *Journal of Political Economy*, Vol. 87, pp. 851–858.

Salz, Tobias (2017) "Intermediation and competition in search markets: An empirical case study."

Schmalensee, Richard (1981) "Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination," *The American Economic Review*, Vol. 71, pp. 242–247.

Sorensen, Alan T. (2000) "Equilibrium Price Dispersion in Retail Markets for Prescription Drugs," *Journal of Political Economy*, Vol. 108, pp. 833–850.

Stahl, Dale O. (1989) "Oligopolistic Pricing with Sequential Consumer Search," *The American Economic Review*, Vol. 79, pp. 700–712.

Topkis, Donald M (1978) "Minimizing a submodular function on a lattice," *Operations research*, Vol. 26, pp. 305–321.

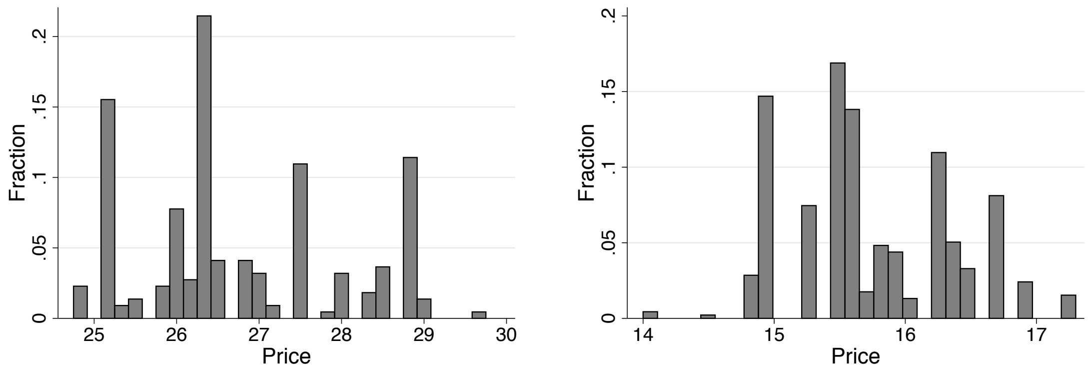
Varian, Hal R. (1980) "A Model of Sales," *The American Economic Review*, Vol. 70, pp. 651–659.

——— (1985) "Price Discrimination and Social Welfare," *The American Economic Review*, Vol. 75, pp. 870–875.

Villas-Boas, Sofia Berto (2009) "An empirical investigation of the welfare effects of banning wholesale price discrimination," *The RAND Journal of Economics*, Vol. 40, pp. 20–46.

Yoshida, Yoshihiro (2000) "The Effects of Third-Degree Price Discrimination in Oligopoly," *The American Economic Review*, Vol. 90, pp. 244–246.

Figure 1: Price variation across customers in a given day



Product A

Product B

Notes: An observation is a transaction (i.e., a customer–product–time combination).

Table 1: Summary statistics

	N	Mean	St. Dev.	p10	p90
A. Statistics at the product level					
Number of products	2,734	-	-	-	-
Transactions per product	2,734	70.420	157.979	1.000	180.000
B. Statistics at the product level (products with more than 250 transactions)					
Number of products	192	-	-	-	-
Transactions per product	192	521.026	322.149	270.000	914.000
Transactions per customer	192	13.817	7.942	5.558	26.618
Days between purchases	192	10.770	3.710	6.325	16.031
C. Statistics at the customer level					
Number of customers	553	-	-	-	-
Fraction of time present	553	0.672	0.348	0.048	0.968
Products purchased per week	8,588	9.472	9.191	2.000	21.000
Number of breakups per week	8,588	0.493	1.191	0.000	2.000
D. Statistics at the transaction level (i.e., customer–product–day)					
Price (per unit)	96,875	27.624	15.727	10.000	45.950
Quantity	96,875	5.590	42.501	0.500	6.000
Days between purchases	85,933	10.327	9.441	3.000	21.000
Price change (indicator)	96,875	0.069	0.253	0.000	0.000
Breakup (indicator)	96,875	0.044	0.204	0.000	0.000
E. Statistics at the transaction level (subsample of top 7 products)					
Price (per unit)	11,374	20.660	10.751	7.000	36.750
Quantity	11,374	2.802	12.067	1.000	6.000
Days between purchases	10,080	10.986	9.182	3.000	22.000
Price change (indicator)	11,374	0.147	0.354	0.000	1.000
Breakup (indicator)	11,374	0.045	0.206	0.000	0.000

Notes: Panels C and D report statistics for products with more than 250 transactions.

Table 2: Decomposing the variance of price–cost margins

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full sample Margin			Top 7 products Margin			
Quantity percentile			-1.302***			-0.786***	
			(0.039)			(0.062)	
Product FE	Yes	No	Yes	Yes	No	Yes	No
ZIP Code FE	No	No	Yes	No	No	Yes	No
Customer–Product FE	No	Yes	No	No	Yes	No	No
Cluster FE	No	No	No	No	No	No	Yes
N	96,875	95,549	94,665	11,374	11,268	10,982	11,268
R^2	0.644	0.962	0.672	0.730	0.987	0.789	0.922

Notes: An observation is a transaction (i.e., a customer–product–day combination). Margin is defined as margin = price − marginal cost. Quantity percentile is the customer's percentile rank in the distribution of average quantity transacted across customers for the product.

Table 3: Prices and the benefits of paying a low price: OLS regressions

	(1)	(2)	(3)	(4)
	Full sample Average price		Top 7 products Average price	
Share of total budget	-2.308***	-1.057**	-2.431**	-2.585**
	(0.444)	(0.460)	(1.057)	(1.030)
Quantity percentile		-1.553***		0.148
		(0.220)		(0.392)
Observations	8,582	8,582	879	879
R^2	0.948	0.949	0.981	0.981

Notes: Standard errors clustered at the customer level in parenthesis. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. An observation is a customer–product combination. All specifications include customer and product fixed effects. Average price (quantity) is the average price (quantity) that the customer paid (purchased) across all transactions. Quantity percentile is the customer's percentile rank in the distribution of average quantity transacted across customers for the product.

Table 4: Breakups and measures of price changes: OLS regressions

	(1)	(2)	(3)	(4)
	Sample:		Sample:	
	Full	Top 7	Full	Top 7
	Breakup		Time to next purchase (in logs)	
Counterfactual price increase:	0.052^{***}	0.023^{***}	0.430^{***}	0.371^{***}
$1\{p_{n+1} - p_n > 0\}$	(0.003)	(0.005)	(0.021)	(0.037)
Observations	96,841	11,351	48,833	5,668
R^2	0.159	0.236	0.368	0.431

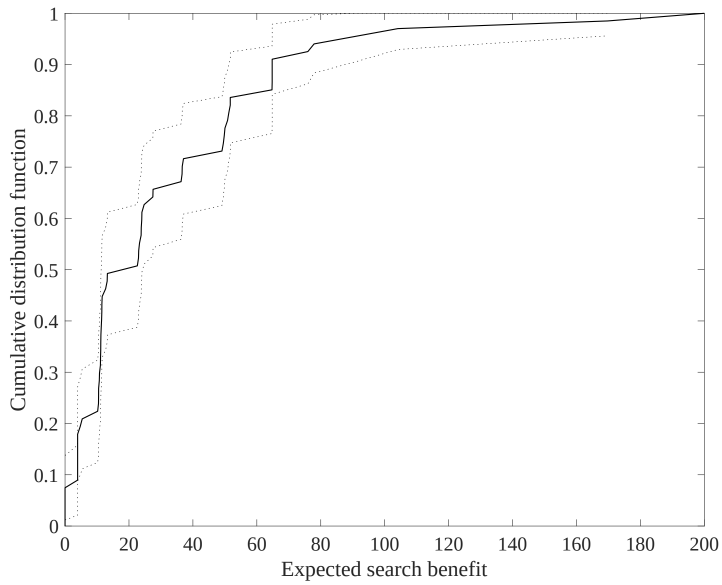
Notes: An observation is a transaction (i.e., a customer–product–day combination). All specifications include customer and product fixed effects. Columns 3 and 4 restrict the sample to transactions that happened in the first half of the sample period. The counterfactual price change at time t is measured to be the difference between the marginal cost at the future time when the customer was expected to return to purchase the good and the marginal cost at time t (see text for more details). The counterfactual price change is calculated for all transactions regardless of whether there was a breakup.

Table 5: Structural parameters

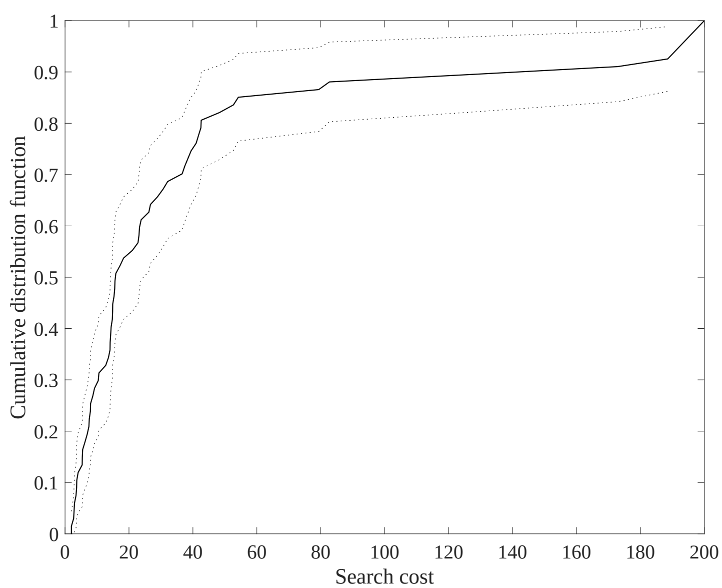
	Estimate
A. Demand	
Price (α)	-0.076^{**}
	(0.031)
B. Net value of matching with a new seller ($\{\theta_{ij}\}$):	
Summary statistics	
Mean	32.816
p10	3.987
p25	10.850
p50	22.853
p75	49.842
p90	65.391
(See Figure 2 (Panel A) for full distribution)	
Observations	4,479

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. An observation is a customer–product–month combination. The identity matrix is used as the weighting matrix, W. *Panel A:* Standard errors, clustered at the customer level, in parentheses. The instruments for price are the marginal cost interacted with indicators for type of restaurant (e.g., pizza, meat, fish, other). The F-statistic of the first stage is 31.84. *Panel B:* Standard errors are adjusted for the two-step estimation. See Cameron and Trivedi (2005, p. 200) for details.

Figure 2: Estimates of expected net search benefits and search costs



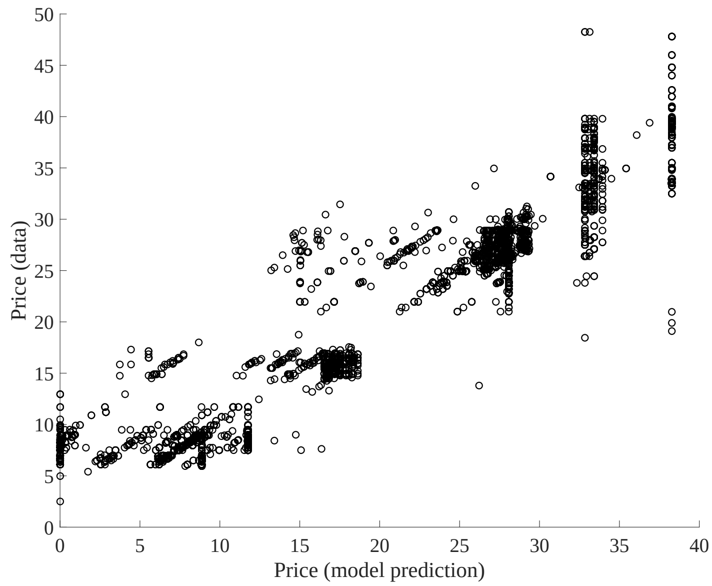
Panel A: Distribution of expected net search benefits (i.e., $\hat{\theta}_{ij}$, $\forall ij$)



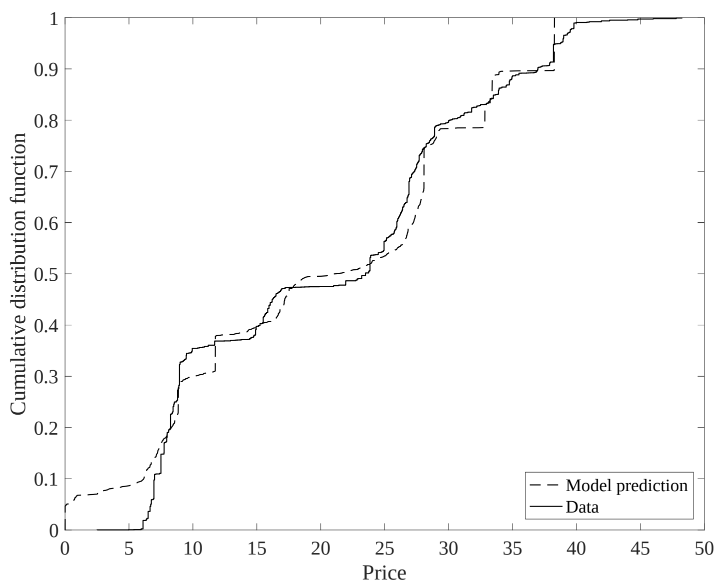
Panel B: Distribution of search costs (i.e., $\hat{sc}_{ij}$, $\forall ij$)

Notes: An observation is a customer–product combination. 95-percent confidence intervals constructed based on Greenwood's formula.

Figure 3: Empirical distribution of prices versus model predictions



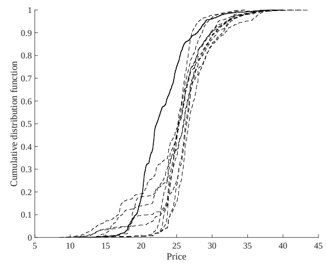
Panel A: Scatter plot



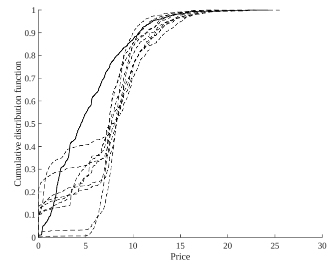
Panel B: Cumulative distribution functions

Notes: An observation is a customer–product–month combination.

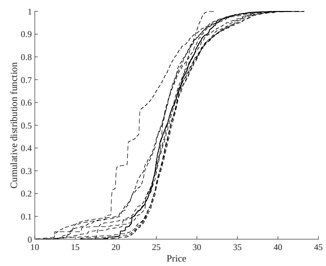
Figure 4: Equilibrium price distributions in price discrimination and uniform price equilibria



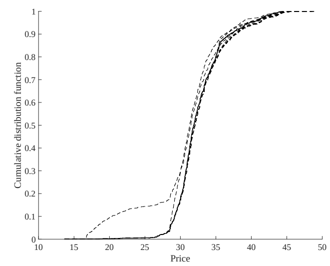
Product 1



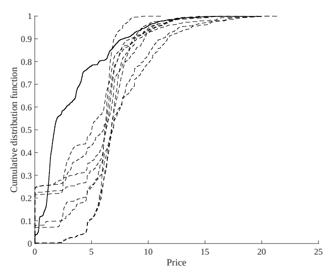
Product 2



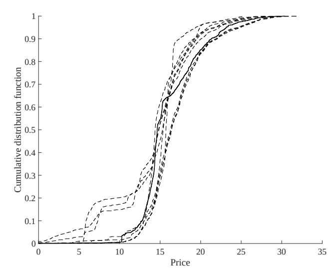
Product 3



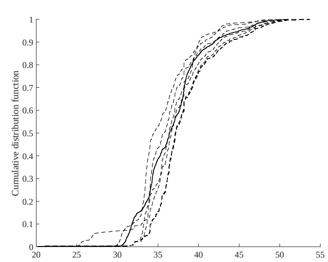
Product 4



Product 5



Product 6



Product 7

Notes: In each plot, the bold line shows the equilibrium price distribution in the uniform price case. The lighter gray lines in each plot indicate the equilibrium price distribution for each buyer in the price discrimination case.

Table 6: Equilibrium market outcomes with and without price discrimination

Panel A: Expected Values

	Expected Value Customers			Expected Profit					
	NoPD	PD	CI (Difference)	No PD	Dev	PD	CI (Dev - NoPD)	CI (PD - Dev)	CI (PD - NoPD)
Product 1	44.73	30.74	(-13.95,-0.02)	20.69	23.98	26.78	(0.01,3.29)	(0.01,6.26)	(0.00,3.31)
Product 2	50.21	31.76	(-20.19,-0.85)	15.12	17.82	23.00	(0.34,3.61)	(0.35,8.74)	(0.01,5.36)
Product 3	206.62	193.00	(-40.31,0.00)	179.23	184.54	185.26	(0.00,8.59)	(0.00,15.99)	(0.00,8.23)
Product 4	64.28	63.57	(-7.68,0.00)	62.27	62.41	62.42	(0.00,1.78)	(0.00,3.34)	(-0.00,1.60)
Product 5	70.30	47.36	(-23.11,-2.99)	28.81	32.38	38.22	(0.57,4.19)	(0.58,9.57)	(0.01,6.46)
Product 6	393.98	370.42	(-25.04,-0.94)	357.32	363.56	364.60	(0.09,7.73)	(0.09,9.23)	(0.00,2.03)
Product 7	26.67	24.35	(-2.54,0.00)	22.52	23.30	23.45	(0.00,0.79)	(0.00,1.08)	(-0.03,0.29)
Mean	120.20	106.90	(-14.18,-0.52)	96.39	99.50	101.66	(0.14,3.19)	(0.14,6.24)	(0.00,3.16)

Panel B: Outcomes in Completed Transactions

	Profit			Consumer Surplus			Total Surplus		
	NoPD	PD	CI (Difference)	NoPD	PD	CI (Difference)	NoPD	PD	CI (Difference)
Product 1	32.85	28.43	(-4.54,-0.11)	52.20	31.45	(-20.71,-0.18)	85.05	59.88	(-25.13,-0.29)
Product 2	34.70	27.44	(-7.87,-1.55)	62.79	33.28	(-31.26,-2.04)	97.48	60.72	(-38.81,-3.59)
Product 3	206.14	190.59	(-16.39,0.00)	222.38	196.11	(-57.91,0.00)	428.52	386.69	(-74.21,0.00)
Product 4	66.04	65.18	(-6.65,0.00)	66.92	65.50	(-15.69,0.00)	132.96	130.68	(-22.34,0.00)
Product 5	47.22	44.49	(-5.94,-2.03)	79.49	49.28	(-30.62,-3.86)	126.71	93.76	(-34.59,-4.48)
Product 6	383.04	367.38	(-27.03,-0.38)	410.53	371.97	(-42.27,-1.53)	793.57	739.35	(-63.43,-1.91)
Product 7	26.29	23.77	(-2.93,0.00)	29.35	24.53	(-4.95,0.00)	55.64	48.29	(-7.77,0.00)
Mean	111.91	104.98	(-9.51,-0.56)	129.61	108.46	(-23.49,-0.81)	241.52	213.43	(-32.66,-1.37)

Notes: CI stands for 95-percent (bootstrapped) confidence intervals. Expected profit (Dev) is the expected profit when the firm deviates from the uniform price equilibrium and practices price discrimination.

Table 7: Comparing market outcomes under price discrimination and uniform prices: OLS regressions

	(1)	(2)	(3)
	$P^{PD}_{quote} - P^{NoPD}_{quote}$	$\text{Search}^{PD} - \text{Search}^{NoPD}$	$P^{PD}_{paid} - P^{NoPD}_{paid}$
Search cost	0.956***	0.092***	0.619***
(standardized)	(0.201)	(0.015)	(0.159)
Observations	67	67	67
R^2	0.746	0.775	0.907

Notes: Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. An observation is a customer–product combination. All specifications include customer fixed effects, product fixed effects, and a control for the average demand shifter (at the customer–product level). Search cost (standardized) is the standardized search cost (i.e., mean zero and standard deviation one). $P^{PD}_{quote} - P^{NoPD}_{quote}$ is the average difference in price quotes. $P^{PD}_{paid} - P^{NoPD}_{paid}$ is the average difference in prices paid by customers. $\text{Search}^{PD} - \text{Search}^{NoPD}$ is the average difference in the (unconditional) probability of search. Search cost and mean demand shifter are defined at the customer–product level.

A Omitted Proofs

Lemma 1.

$$0 < \frac{dCS_{ij}}{dp}\frac{dp}{d\theta_{ij}} < 1,$$

where p is the solution to the firm's pricing problem (see equation (2)) and $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$.

Proof. See Online Appendix. $\square$

Proof of Proposition 1

i) The result follows from the observation that the profit function in equation (2) is log-supermodular in $(p, -(\delta E[V_{ij}] - sc_{ij}))$,

$$\frac{\partial^2 \log \pi}{\partial p \partial \theta_{ij}} = q(p)\frac{G(CS_{ij}(p_{ij}) - \theta_{ij})g'(CS_{ij}(p_{ij}) - \theta_{ij}) - g(CS_{ij}(p_{ij}) - \theta_{ij})^2}{G(CS_{ij}(p_{ij}) - \theta_{ij})^2} < 0,$$

where $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$ and the inequality makes use of the fact that G is log-concave. The result therefore follows from Topkis's univariate monotonicity theorem (Topkis, 1978).

ii) To establish the result, I first analyze how $E[V_{ij}]$ changes with sc_{ij}. $E[V_{ij}]$ is given by

$$E[V_{ij}] = \int \left(CS_{ij}(p(c)) + \int_{CS_{ij}(p(c))-\theta_{ij}}^{\infty} (\theta_{ij} - CS_{ij}(p(c)) + t)g(t)dt \right) h(c)dc,$$

where $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$ and g and h are the density functions of the search cost shock and marginal cost, respectively.[27] Using implicit differentiation and rearranging terms I obtain

$$\begin{aligned}\frac{dE[V_{ij}]}{dsc_{ij}} = \int \Bigg(& G(CS_{ij}(p(c)) - \delta E[V_{ij}] + sc_{ij})\frac{dCS_{ij}}{dp}\frac{dp}{d\theta_{ij}}\left(\delta\frac{dE[V_{ij}]}{dsc_{ij}} - 1\right) \\ & + (1 - G(CS_{ij}(p(c)) - \delta E[V_{ij}] + sc_{ij}))\left(\delta\frac{dE[V_{ij}]}{dsc_{ij}} - 1\right)\Bigg) h(v)dv\end{aligned}$$

or

$$\frac{dE[V_{ij}]}{dsc_{ij}} = -\frac{\int \left(G(CS_{ij}(p(c)) - \theta_{ij})\frac{dCS_{ij}}{dp}\frac{dp}{\theta_{ij}} + (1 - G(CS_{ij}(p(c)) - \theta_{ij}))\right) h(c)dc}{1 - \delta \int \left(G(CS_{ij}(p(c)) - \theta_{ij})\frac{dCS_{ij}}{dp}\frac{dp}{\theta_{ij}} + (1 - G(CS_{ij}(p(c)) - \theta_{ij}))\right) h(c)dc} < 0.$$

[27] Demand shocks are ignored for ease of notation.

The fact that $\frac{CS_{ij}}{dp}\frac{dp}{d\theta_{ij}} \in (0,1)$ (see Lemma 1) guarantees that the denominator is positive. The result then follows from combining $dE[V_{ij}]/dsc_{ij} < 0$ and Proposition 1(i).
iii) The result can be established by noting that

$$\frac{d\Pr(Search)}{dsc_{ij}} = \frac{d\int(1-G(CS(p(c))-\delta E[V_{ij}]+sc_{ij}))h(c)dc}{dsc_{ij}}$$
$$= \int -g(CS(p(c))-\delta E[V_{ij}]+sc_{ij})\left(\frac{dCS_{ij}}{dp}\frac{dp}{d\theta_{ij}}-1\right)\left(\delta\frac{dE[V_{ij}]}{dsc_{ij}}-1\right)h(c)dc < 0,$$

where $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$. The inequality makes use of $\left(\frac{dCS_{ij}}{dp}\frac{dp}{d\theta_{ij}}-1\right) < 0$ (see Lemma 1) and $\frac{dE[V_{ij}]}{dsc_{ij}} < 0$ (see proof of Proposition 1(ii)).

B Estimation Procedure

In what follows, I provide step-by-step details on the procedure that I use to estimate the empirical model.

1. *(Clusters of customer–product combinations)* For each product, I generate $k = 10$ clusters of customers using the k-medoid clustering algorithm. Clustering is based on five variables: the fraction of a customer's total expenditure throughout the sample period that was spent on the product, average quantity per transaction, an indicator for pizza restaurant, an indicator for meat restaurant, and an indicator for fish restaurant. Because I group customers based on both continuous and discrete variables, I use the Gower distance to measure dissimilarity across customers.
2. *(Demand Function)* The customer–product level demand function is given by $q_{ij}(p_{ijt}, \xi_{ijt}) = \exp\{\alpha p_{ijt} + \xi_{ijt}\}$ (i.e., subscript ij denotes a customer–product combination, and subscript t denotes a particular transaction), where $\xi_{ijt} = \xi_{ij} + \Delta\xi_{ijt}$. I assume that the innovation in demand shocks satisfy the moment condition: $E[z_{ijtl}\Delta\xi_{ijt}] = 0, \forall l \in \{1 \ldots L\}$. I estimate the model using a two-stage least squares estimator. The demand specification includes price and customer–product fixed effects. The instruments for price are the product-specific unit cost interacted with four restaurant-type dummies (i.e., pizza, seafood, meat, other).
3. *(Distribution of Demand Shocks, H_{ij})* Based on the demand estimates in step 1, I compute the predicted demand shocks for every transaction of every customer–product combination ij: $\hat{\xi}_{ij} + \Delta\hat{\xi}_{ijt}$. I use the distribution of predicted demand shocks for customer–product combination ij as the estimate of the distribution of demand shocks for that particular customer–product combination, $\hat{H}_{ij}$.
4. *(Search cost, sc_{ij})* I estimate the average search cost of each customer–product combination, sc_{ij}, using a two-step method proposed by Flinn and Heckman (1982). In the two-step method, one first uses a GMM estimator to estimate the equilibrium net value of matching with a new seller for every customer–product combination ij, $\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$. Given the estimate for θ_{ij}, one can then recover sc_{ij} using the structure of the model in a second step. I discuss both steps in what follows.

   In the first step, I estimate the equilibrium net value of matching with a new seller for a customer–product combination ($\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$) using the price-optimality condition given by equation (3). Specifically, for every customer–

product combination ij, I form the moment condition

$$E\left[p_{ijt} - c_{ijt} - \nu_{ijt} + \frac{G(CS_{ij}(p_{ijt}) - \theta_{ij})q_{ij}(p_{ijt})}{G(CS_{ij}(p_{ijt}) - \theta_{ij})q'_{ij}(p_{ijt}) - g(CS_{ij}(p_{ijt}) - \theta_{ij})q_{ij}(p_{ijt})^2}|NS\right] = 0,$$

where subscript t is used to enumerate the transactions of customer–product combination ij. These moment conditions require that the observed prices be optimal, and are a monotonic function of θ_{ij} (all else equal). These moment conditions depend on data (i.e., price and cost), the assumption that the distribution of search cost shocks is standard normal, $G \sim N(0, 1)$, and the demand estimates in three ways: q_{ij}, q'_{ij}, and $CS_{ij}(p_{ijt}) = -\exp\{\alpha p_{ijt} + \xi_{ijt}\}/\alpha$ (the formula for customer surplus in the case of log-linear demand).

Inside the brackets of the moment condition, I include the error term ν_{ijt}, which is assumed to be a component of the seller's marginal cost that is unobserved to the econometrician. Because the estimation sample only includes completed transactions (i.e., prices that did not induce search), and the value of ν_{ijt} affects prices, the moment condition is conditional on the event of "no search" (NS). Instead of imposing a parametric restriction on the distribution of ν, I approximate $E[\nu_{ijt}|NS]$ using the values of ν predicted by the model. Specifically, given θ_{ij}, one can compute the ν_{ijt} of every transaction t of customer–product combination ij using the price-optimality condition given by equation (3) (as well as the data, parametric assumptions, and demand estimates discussed above). Using the predicted values of ν for all transactions of all customer–product combinations, I approximate $E[\nu_{ijt}|NS]$ using the average value of ν for all transactions involving product j (i.e., the product that corresponds to customer–product combination ij),

$$\hat{E}[\nu_{ijt}|NS] = \frac{\sum_{i',t'} \nu_{i'jt'}}{NT_j},$$

where NT_j is the total number of transactions involving product j.

The estimates for $\{\theta_{ij}\}$ (i.e., the vector of equilibrium net values of matching with a new seller for every customer–product combination) are obtained using a GMM estimator. The GMM estimator is based on the moment conditions above (i.e., one moment condition per customer–product combination), with the identity matrix as the weighting matrix.

In the second step of the two-step method, I back out sc_{ij} using the estimates for the equilibrium net values of matching with a new seller ($\theta_{ij} \equiv \delta E[V_{ij}] - sc_{ij}$)

in conjunction with the structure of the model. In practice, I use the structure of the model to compute $E[V_{ij}]$ given $\hat{\theta}_{ij}$. With estimates for $E[V_{ij}]$ and θ_{ij} in hand, I can recover sc_{ij} using the identity that defines the equilibrium net value of matching with a new seller, $\theta_{ij} = \delta E[V_{ij}] - sc_{ij}$: $\hat{sc}_{ij} = \delta \hat{E}[V_{ij}] - \hat{\theta}_{ij}$. Computing $E[V_{ij}]$ requires solving the equilibrium of the game for customer–product combination ij, which I can do via simulation. Notice that given θ_{ij}, the price-optimality condition in equation (3) provides the seller's optimal price as a function of the marginal cost, demand shock, demand function, and distribution of search cost shocks. Given $\hat{\theta}_{ij}$, I thus simulate the equilibrium distribution of prices for customer–product combination ij using the optimal pricing rule given by equation (3) and $NS = 2,000$ draws of demand shocks and marginal costs. That is, I complete the following steps.

a) I take $NS = 2,000$ draws from i) the distribution of demand shocks, $\hat{H}_{ij}$; ii) the empirical distribution of the observed component of the marginal cost (c_{ijt}); iii) the distribution of fitted values of ν of transactions involving the product of customer–product ij. When sampling the values of ν, I use sampling weights that correct for the fact that lower values of ν may be oversampled in the distribution of predicted values of ν (i.e., the sample includes completed transactions only). The weight of predicted value ν_{ijt} is based on the predicted probability of the customer completing that transaction, $G(\hat{CS}_{ij}(p_{ijt}) - \hat{\theta}_{ij})$.

b) For every tuple of simulated draws s, $(\xi_{ijs}, c_{ijs}, \nu_{ijs})$, I use the optimal pricing rule given by equation (3) to find the optimal price, p_{ijs}. This step makes use of the estimates of demand, θ_{ij}, and the assumption that $G \sim N(0,1)$. I then approximate the equilibrium price distribution of customer–product combination ij, $F_{ij}(p)$, using the distribution of the simulated optimal prices, $\{p_{ijs}\}_s$, and assuming that sellers are ex-ante symmetric.

Using the estimates for the equilibrium price distribution, $F_{ij}(p)$, I then estimate the equilibrium value of matching with a new seller, $E[V_{ij}]$, using the definition

$$E[\hat{V}_{ij}] = \int \max\left\{CS_{ij}(p;\xi_{ij}),\ \hat{\theta}_{ij} + \varepsilon\right\} dG(\varepsilon) d\hat{H}_{ij}(\xi_{ij}) d\hat{F}_{ij}(p).$$

This step makes use of the assumption that $G \sim N(0,1)$ and the estimates of θ_{ij}, $F_{ij}(p)$, q_{ij}, and buyer i's distribution of demand shocks for product j, $H_{ij}(\xi)$.

5. (*Standard errors*) I follow the procedure outlined in Cameron and Trivedi (2005, p. 200) to ensure consistency of standard errors despite the sequential estima-

tion of q_{ij} and θ_{ij}. Confidence intervals for the welfare measures in Table 6 are constructed using parametric bootstrapping. That is, I make use of the joint asymptotic distribution of demand parameters and net value of matching with a new seller estimates to simulate 100 vectors of parameters. For each of these simulated vectors of parameters, I compute search costs and the relevant welfare objects in Table 6. Confidence intervals are then constructed using the bootstrapped distribution of welfare outcomes.