Intermediation and Competition in Search Markets: An Empirical Case Study*

Job Market Paper

Tobias Salz (NYU)

December 14, 2015

For latest version click here.

Abstract

In many decentralized markets buyers rely on intermediaries to find sellers. This paper argues that intermediaries can affect buyer welfare both *directly* by reducing expenses of buyers with high search cost but also *indirectly* through a search externality that affects the prices paid by those buyers that do *not* use intermediaries. To investigate these two effects this project uses data from the New York City trade-waste market in which all businesses in the city contract individually with private waste carters to arrange for their waste disposal. Search in this market is costly for buyers because of the large number of sellers and the idiosyncratic nature of the contractual arrangements. Buyers can either search and haggle by themselves or through a waste-broker. Combining elements from the empirical search and procurement-auction literature, I construct and estimate a model for such a decentralized market setting. Results from the model show that buyers both in the broker market and in the search market benefit significantly from the activity of intermediaries. Intermediaries also improve overall welfare by reducing the cost of price discovery and the mis-allocation of sales to higher cost sellers.

JEL classification: L11, L15, L80

Keywords: Market Structure, Decentralized Markets, Search Cost, Auctions, Intermediation

*I am indebted to my advisors John Asker, Guillaume Frechétte, Kei Kawai, and Alessandro Lizzeri for their guidance and support. I am extremely grateful to Isabelle Perrigne and Quang Vuong for helpful discussions. Luis Cabral, Michael Dickstein, Myrto Kalouptsidi, John Lazarev, Konrad Menzel, Ariel Pakes, Maryam Saeedi, Maher Said, Kevin Thom, Allan Collard-Wexler, Larry White, Jidong Zhou as well as seminar participants at the *NYU Econometric lunch* and the *NYU Stern Applied Micro Seminar* provided invaluable feedback. I also want to thank Klaus-Peter Hellwig, Yunmi Kong, Joseph Mullins and Sevgi Yuksel for helpful conversations. The research in this paper was conducted with data provided by New York City's Business Integrity Commission, which has been modified to preserve the anonymity of all market participants. In addition, results have been reviewed to ensure that no confidential information is disclosed. Research results and conclusions expressed are those of the author and do not necessarily reflect the views of the Business Integrity Commission. All remaining errors are mine.

1. INTRODUCTION

This paper empirically investigates how match-making intermediaries affect competition and welfare in a decentralized market. I introduce a model that combines elements from the search and procurement-auction literature. The model is applied to a new and detailed dataset from the New York City trade-waste market. Search in this market is costly for buyers because of the large number of sellers and the idiosyncratic nature of the contractual arrangements. Buyers have the option to either search by themselves (in the "search market") or delegate the search to waste brokers (the "broker market"), which are specialized intermediaries. The self-selection of buyers with high search cost into the broker market changes the composition of buyers in the search market and makes it more competitive for sellers. Intermediaries therefore create a search externality (Salop and Stiglitz (1977)) that benefits buyers in the search market through lower prices. Intermediation also mitigates the inefficiencies arising in search markets by reducing search cost and re-allocating contracts to lower-cost sellers.

The empirical literature in industrial organization has mostly explored settings in which firms post prices and compete for fully informed consumers. In many markets, however, buyers have only limited information about prices and the cost to provide the product, and therefore prices, are customer-specific.[1] This is especially true for retail service markets and those for investment goods. One common market response to these frictions is the emergence of intermediaries that specialize in the provision of information. Spulber (1996b) estimates that 25% of the U.S. economy is composed of intermediation.[2] The theoretical literature has acknowledged the importance of intermediaries and discusses their possible effect on market outcomes. And yet there is little empirical evidence on this important type of firm. This paper provides one of the first empirical assessments of the activity of intermediaries in concentrated markets and their impact on prices and welfare.[3] To that end the paper formulates an empirical model for decentralized markets that is applicable beyond this particular setting.

For historical reasons, the trade waste industry in New York provides a unique opportunity to gain insight into a decentralized market. As a measure against the historical entrenchment of organized crime in New York's waste carting industry, the city has established a regulator that monitors the industry and collects data about many operational aspects of carters' activity. This study uses an anonymized panel of all the bilateral agreements between those private carters and the more than 100,000 businesses (on the buyer side) that fall under the regulation. The trade waste market in New York features mostly smaller suppliers (carters) and in the majority of zip-codes buyers have the choice among more than 20 different waste carters. Buyers have the option to procure the contract through a waste broker.

These brokers can affect buyer-welfare in the market both *directly*, because their service allows buyers with high search cost to reduce their search expenses, but also *indirectly* through an effect on prices

[1] The empirical literature on bargaining, however, has dealt with transaction specific prices. See for example Grennan (2012) and Lee and Fong (2013). The former uses observed prices in the estimation of primitives and the latter provides a way of recovering unobserved transfers (prices).

[2] Examples beyond housing and real estate are: textile industry (http://www.economist.com/news/business/21657375-story-indias-biggest-maker-towels-and-their-journey-cotton-field-big-box (last accessed on 07/16/15)), advertising brokers, freight brokers (http://www.dat.com/~/media/files/dat/resources/dat_broker_start-up_guide_rnd2.pdf (last accessed on 08/29/15)), literary agents (http://aaronline.org/ (last accessed on 08/29/15)), energy brokers (https://en.wikipedia.org/wiki/Energy_broker (last accessed on 08/29/15)).

[3] Related empirical work with a different focus is presented in Hendel et al. (2009) and Gavazza (2011) and discussed in more detail below.

in the search market. The latter occurs under the plausible assumption that sellers do not observe an individual buyer's search cost. Intermediaries are used by buyers with relatively higher search cost. This self-selection therefore changes the composition of buyers in the search market and forces sellers to quote lower prices relative to a scenario where all buyers search by themselves. A descriptive look at the data indicates that both channels could be important in this setting. Firstly, a comparison of brokered contracts and those in the search market suggests that customers with relatively higher search cost are using brokers. While brokers are able to negotiate cheaper rates, final prices after commissions are higher than in the search market, which points at the selection of customers. Additional regressions show that areas with more intermediated contracts feature lower prices in the search market.

The data is analyzed through a model which combines elements from the empirical search and auction literature. It will map the above data facts into economically interpretable estimates and allow me to quantify the *direct* effect on buyers in the broker market and the *indirect* effect due to the search externality. As in an empirical procurement model, the cost of service provision in the model will be unique for a buyer-seller relationship. Previous search models, which typically apply to retail settings, are often identified from price and quantity data or even price data alone. Such models typically rely on the assumption that sellers' cost are constant across buyers and that the data is generated by a mixed strategy pricing equilibrium. While such an approach is sensible in a posted price setting where goods are not customer-specific, it is less appropriate in a market for a highly individualized service.

Using the available data, the service cost distribution can be estimated non-parametrically and the distribution of search cost semi-parametrically. The identification of the model is complicated by the fact that one needs to recover two primitives from observed prices in the market − the search cost for buyers and the cost of service-provision for sellers. Key to the identification strategy is the institutional fact that brokers award contracts through a *competitive bidding process* resembling a first-price auction as well as the availability of both brokered and un-brokered contracts in the data. Building on this, the identification follows an intuitive argument: observed contract prices in the brokered market are the lowest offered price amongst N bidders. Following standard arguments from the empirical literature on auctions, one can identify the cost of sellers from this data. In the search market we do not know how many competitors sellers take into consideration when making their price offers to customers. The number of competitors is a result of buyers' optimal search strategy. Both the equilibrium number of price inquiries and the search-cost are unobserved. But since the cost of sellers is known from the brokered market, one can infer the number of price quotes buyers must have asked for to rationalize the price distribution under these cost distributions. From this one can in turn identify the distribution of search cost for sellers under an exclusion restriction. Varying amounts of corrugated cardboard generated by customers, which is an important resource for carters and traded like a commodity, serves as such an exclusion restriction.

My estimates suggest that search costs make up a large percentage of buyers' total expenses, ranging from about 29% to 49%. Most of the price difference between the brokered market and the search market can be attributed to the fact that brokers entertain relationships with firms that persistently offer lower rates rather than the fact they they ask for additional price quotes.

In the main counter-factual of interest the ability to contract through brokers is removed. This alternative market scenario reveals that both the *direct* and *indirect* effects are large. Expenses for buyers that were using brokers rise on average by 17.0% if they have to contract directly through the search market. Prices in the search market are rising because the average buyer now compares fewer prices, which reduces the competitive pressure on sellers. Expenses for buyers that were already searching by themselves therefore rise on average by 13.25%. The search externality through intermediaries has therefore strong implications for the distribution of rents in the market: Buyers that never use an intermediary benefit almost as much from their activity as those buyers that do. Without intermediation, profits for sellers rise, and this increase is larger for sellers that have persistently higher cost.

Taking everything into consideration, one can bound the total annual decrease in welfare when intermediation is not possible. It lies between $18.6 and $43.6 million, or between 5.1% and 12.2% of the annual market volume.[4] Total welfare is determined by the total search and service cost that the market incurs to provide the service for buyers.[5]

An extension at the end provides a discussion of merger guidelines in decentralized markets, which builds on the conceptual framework introduced here. Using the estimates from this particular market it is shown that there are cases in which a merger criterion on prices alone would lead to a different recommendation than one that also accounts for the search cost of customers. This highlights the importance to account for the specific frictions that are present in decentralized markets and how it can inform standard practice in competition policy.

RELATED LITERATURE

The work here contributes to the empirical literature on search and intermediation and the model builds on tools in the empirical auction literature. In the following I discuss related papers in these areas.

INTERMEDIATION

The literature on middlemen is quite fragmented, reflecting the many different aspects of intermediation. Here the focus is on intermediaries that help buyers to *search* for sellers, but intermediaries also function as guarantors of quality and liquidity and act as market makers.[6] There are a few other empir-

[4] Welfare can only be bounded because total change in welfare depends on the fixed cost of broker services, for which I have no estimate.

[5] There is virtually no extensive margin as all buyers are required to contract with a carter. An exception are self-haulers. Self-haulers need to register with the regulator to do so. The registration fee for a two year term is $1000 and $400 for each utilized vehicle, see http://www.nyc.gov/html/bic/html/trade_waste/what_self-carting.shtml Decomposing the welfare change in this way I find that about 53.8% are due to an increase in search cost and the remainder due to a reallocation of contracts to firms with higher service-cost.

[6] Spulber (1996a) provides a taxonomy that classifies intermediaries according to four different functions: price setting and market clearing, the provision of liquidity and immediacy, guaranteeing and monitoring, and the facilitation of matching and searching. The theoretical studies on match-making intermediaries is rooted in the literature on dynamic search. Rubinstein and Wolinsky (1987) describe a model in which a certain amount of trade is (endogenously) conducted through agents that have no intrinsic valuation of the good. This result arises because waiting for a trade is costly and it can therefore be profitable for intermediaries to build up inventory and sell it to searching buyers. The study is inconclusive about the welfare effects of intermediation. Gehrig (1993) and Rust and Hall (2003) explore the competition between decentralized forms of trading and central clearing houses with posted prices. The latter is an extension of Spulber (1996b). Moraga Gonzalez et al. (2014) extends the framework of Rubinstein and Wolinsky (1987) to include heterogeneous search cost. Middlemen in their setting can exist because of unobserved quality of a good. Hellwig (2013) provides a setting in which middlemen are active and collect rents in the market without an exogenous advantage in search and screening technology. These models have in common that they

ical studies on intermediation that are broadly related. The study most closely related is Gavazza (2012) (henceforth G2012), which examines the effect of intermediaries in the secondary market for business aircrafts using a dynamic search and bargaining framework that is brought to data. My paper differs in conceptual ways, as well as the model, application, and findings. While G2012 is interested in intertemporal trading frictions, the focus here is on pricing incentives of a finite number of firms with market power, which is exacerbated by customers search costs. G2012 therefore only tries to explain aggregate state dependent prices in the market, while in this setting the unobserved sources of heterogeneity are both crucial for buyers strategic choice to delegate to an intermediary as well as in explaining the full distribution of transaction specific prices. The modeling choices reflect these conceptual differences. One important distinction is the fact that in my paper buyers make the choice of whether to use an intermediary, whereas in G2012, in any given period, buyers are matched exogenously with a seller or an intermediary. G2012 builds on the dynamic macro search literate while the model here builds on static search and auction models, which are more common in the industrial organization literature. G2012 finds that intermediary reduce welfare, whereas I find a welfare improvement. The source of the welfare improvement is a key contribution of my paper.

Hendel et al. (2009) compare the performance of a realtor listing service with a platform in which house owners sell their homes directly and delineate the welfare consequences for the two alternative forms of market organization. They find that, conditional on observables and after various techniques to control for selection of buyers across platforms, real estate agents are not able to obtain higher prices for buyers than the platform. Real estate agents do, however, sell homes faster.[7] Gavazza (2011) explores the role that commercial aircraft lessors play in reducing trading frictions.

Literature on Search Cost

The broader theme of this study are markets in which consumers lack full information about prices. Stigler (1961) was the first one to discuss equilibrium pricing behavior of firms and consumers' search effort under such conditions. Consumers' lack of information about prices is an explanation for the empirically pervasive phenomenon of price dispersion. Such models have therefore become an important

look at the dynamic incentives of agents in the market to enter and trade and explore the market structure resulting from these incentives. Here, I focus on static pricing and search incentives. The model is therefore much more closely related to the auction literature than to the aforementioned studies. A wide body of theoretical literature has emphasized other functions of intermedairies: An example of the monitoring role of intermediaries is provided in Lizzeri (1999) who explores strategic information disclosure of certification intermediaries and the effects of competition amongst such intermediaries. In a similar vein Inderst and Ottaviani (2012) study the role of kick-backs by suppliers to intermediaries for their recommendations to customers. Unlike Lizzeri (1999) the information provided to buyers is horizontal. In Biglaiser (1993) monopolistic middlemen have a higher incentive to become experts to screen the quality of the good because they transact larger quantities.

[7] There are other papers that consider the realtor market with a focus on different economic questions. Levitt and Syverson (2008) are interested in the agency issue that arises from realtors' superior information and argue that they sell houses too fast. Barwick and Pathak (2011) are empirically explore the question of excess entry in the Boston real estate market. They find that too many agents enter the market from a social welfare perspective since additional entrants do not lead to lower commission rates for consumers. Bar-Isaac and Gavazza (2013) document the contractual arrangements of real estate agents in the New York rental market. Han and Strange (2014) provide an overview of this literature. There are a number of papers that highlight the importance of intermediaries in the context of international trade: Akerman et al. (2010) explain wholeseller intermediation with returns to scope, Blum et al. (2009) describe intermediation in exports from Chile to Colombia, Jensen et al. (2010) look at how wholesale intermediation varies by product and country characteristics.

part of the literature in industrial organization.[8]

This study explores the effect of intermediaries and in particular the idea that intermediaries create a search externality by changing the composition of buyer types in the search market. The theoretical literature has explored such externalities, which arise if firms cannot tell consumers of different types apart. For example in Salop and Stiglitz (1977) or more recently in Armstrong (2014). A more general discussion is given in Stiglitz (1989). This is to my knowledge the first empirical paper that addresses the importance of such an externality and the effects of intermediaries on competition and welfare more broadly.

The literature distinguishes between sequential and non-sequential search. Part of the model presented here builds on the literature of non-sequential search. Under sequential search (McCall (1970)) people ask for one price quote at a time, weighing a potential improvement in their offer against the cost of waiting and searching. Non-sequential search (Stigler (1961)) on the other hand, is static, and the decision maker *ex ante* commits to a certain number of draws from the price distribution.[9] For the purposes of this study the assumption of non-sequential search has several advantages. Firstly, under this assumption the firm's problem against the searching consumer (in the search market) is equivalent to a first-price auction with an unknown number of competitors. This makes the problem tractable and allows me to rely on the empirical tools that have been developed in the auction literature (see below for an overview). It also increases transparency of market comparability and the identification. In terms of the realism of this assumption, several studies have found that non-sequential search better explains peoples' actual shopping behavior. De los Santos et al. (2012) use data on whole records of sequential retail-shopping behavior to test the assumption of non-sequential against sequential search. They note that non-sequential search is more in line with the data. Honka and Chintagunta (2014) also reject sequential in favor of simultaneous search using data from the automobile insurance industry.

The goal of most empirical studies in this literature is to identify the distribution of customer search costs. Firms' cost to service a customer is typically assumed to be constant across customers in this literature. Hortaçsu and Syverson (2003) document price dispersion in the mutual fund industry and estimate a search model that allows for product differentiation, using price and quantity data. Hong and Shum (2006) propose a procedure to estimate search cost from price data alone. In their underlying model firms are identical and the price distribution arises because firms engage in a mixed-strategy pricing equilibrium against a distribution of consumers with different search cost. The model is estimated using on-line prices for textbooks.[10]

In an extension I use the model estimates to perform a merger analysis in this particular setting of a decentralized market with search frictions. Such an analysis is closely related to the study by

[8] An early contributions from Burdett and Judd (1983) shows that a full continuous price distribution can arise from a mixed-strategy pricing equilibrium. Consumers in the model engage in non-sequential search, and both firms and consumers in the model are homogeneous with respect to their marginal cost and search cost. In Stahl (1989) consumers engage in sequential search and either have positive or no search cost. Firms play a mixed strategy equilibrium against these two types of consumers, which leads to dispersion in prices. As the fraction of consumers with positive search cost goes to one, the outcome converges to monopoly pricing, which is known as the Diamond Paradox (Diamond (1971)).

[9] Morgan and Manning (1985) formulate a general theory that combines both sequential and non-sequential search.

[10] Several studies build on this idea: Giulietti et al. (2014) estimate search cost in the British electricity retail market and allow the cost of the incumbent firm to be different from the cost of the entrants. Honka (2014) estimates search costs for the auto insurance industry and allows consumers to form consideration sets. Koulayev (2014) models the sequential search using data from on-line hotel booking, which provides him with the actual sequential decisions of the search process.

Allen et al. (2013), who use data from the Canadian mortgage industry along with quasi-experimental variation induced by a merger for a non-parametric procedure. The authors quantify how the merger alters pricing and search in the market and how customers with different search cost are differentially affected.

With the exception of Allen et al. (2013) all aforementioned papers assume homogeneous cost to service buyers.[11] The fact that these studies explore retail settings, where consumers purchase from firms, makes this a plausible assumption. In this setting, however, the buyers are firms themselves, and there is both observed and unobserved variation that determines how costly it is for sellers to service the buyer. Factors that determine the cost of service provision include the location, the quantity, the composition of the waste, the distance to the transfer station and many unobserved factors. I therefore do not follow the literature in the assumption that the observed price variation is coming from a mixed strategy pricing equilibrium and uniform cost and instead allow the cost to be customer-specific. The empirical model therefore has two distributions of unobservables: search-cost and service-cost.

EMPIRICAL LITERATURE ON AUCTIONS

The model in this study is related to the empirical literature on auctions. Brokers in this market explicitly use procurement auctions to allocate contracts to sellers, but competition in the search market can also be viewed through the lens of competitive bidding: A customer chooses a carter if he quotes the lowest price amongst the N carters that were called. The difference is that the carter does not know the number of competitors that he is bidding against, which depends on the search costs of the customer and the optimal search strategy against the known distribution of prices. The pricing sub-game of sellers in the search-market can therefore be viewed as a first price auction with an unknown number and composition of sellers. The identification of auction models has been discussed in Guerre et al. (2000) and further developed in Athey and Haile (2002) for the case of asymmetric auctions.[12]

ROADMAP

The remainder of the paper is organized as follows. Section 2 provides relevant industry facts and describes the data, section 3 establishes important descriptive facts, section 4 describes the model, section 5 the identification of the model, section 6 the estimation, section 7 the results, section 8 the counterfactual computations, section 9 an extension on merger analysis and section 10 concludes.

2. DATA AND INDUSTRY FACTS

This section first gives an overview of the data and then establishes several facts about the New York market and the activity of brokers. Along with the reduced form findings established in Section 3 these institutional details will guide the modeling choice. Two of these facts are crucial. First, the market

[11] However, due to the particular assumptions on the bargaining protocol and the information structure Allen et al. (2013) still only estimate a model of one-sided heterogeneity.

[12] The methods have been successfully used to investigate auctions with resale (Haile (2001)), entry into auctions (Li and Zheng (2009)), the inclusion of time incentives in highway procurement projects (Bajari and Lewis (2009)), collusion in auctions (Asker (2010)), bid preference programs (Krasnokutskaya and Seim (2011)) and many others.

supports a large number of suppliers in a geographic area, leaving buyers the choice among many different carters. Second, brokers procure contracts through a *Request for Proposals*, which is akin to a first price auction. The latter will be important in the identification of the model, which is discussed in detail in section 5.

2.1 Data

Trade waste industry is the official name for New York's private waste market.[13] To free the trade waste industry from the ties to organized crime, Mayor William Louis Giuliani established the *Trade Waste Commission* in 1995. This commission, which has subsequently been renamed the *Business Integrity Commission* (henceforth BIC), has a comprehensive oversight mandate over the industry. Private waste carters need to be licensed with the BIC.[14] The BIC monitors carters' financial and operational activity. As part of this effort carters need to report on many of the operational aspects of their business.

This project uses data that has been collected this way and modified to preserve the anonymity of customers and carters. It ranges from July 2009 to June 2014. The data includes all contracts in that period of time including the zip-code of the customer, the negotiated price (quoted either in terms of volume or weight) and the quantity of waste generated by the customer. Additional information for each contract includes the date at which it was signed, whether or not the contract was brokered, the type of waste and to which transfer station the waste is carted to. In total there are 1,227,024 panel observations at a half-yearly frequency.[15]

Table 1: Summary statistics

Variable	**Mean**	**Median**	**Std. dev**
Monthly charges ($)	224	94	460.2
Price	8.77	7.49	7.4
Monthly quantity (cu/yd)	26.44	8.66	57.31
Recyclables (yes/no)	0.46	0.0	-
Number of weekly pickups	5.26	5.0	3.69

Note: In total there are 1,227,024 observations. Weight based quantities have been converted to volume based quantity using weight to volume conversion rates for dry refuse. The quantity and pickup variables are winsorized at the 1% level to account for outliers.

Table 1 provides some summary statistics. The mean monthly charges of a business are about $224

[13]New York's residential waste disposal market is publicly administered by the Department of Sanitation. An exception are self-haulers. Self-haulers need to register with the BIC to do so. The registration fee for a two year term is $1000 and $400 for each utilized vehicle, see http://www.nyc.gov/html/bic/html/trade_waste/what_self-carting.shtml (accessed on 08/15/15).

[14]Beyond this, the BIC sets a rate cap for the market and sets rules about sub-contracting and merger applications. If measured in cubic yard this rate cap was $12.2 before 2008, $15.89 from 2008 to 06/2013 which is the relevant data period and $18.27 from 07/2013 onwards. Likewise, if measured in 100lbs. it was $8.00 before 2008, $10.41 from 2008 to 06/2013 and $11.98 thereafter. Another important restriction regards the length of contracts, which cannot exceed two years after which the customer has the option to sign with another carter.

[15]This excludes the contracts that only involve medical waste, shredding of paper and cardboard or grease haul.

with a very large standard deviation of 460.0 reflecting the tail of extremely large waste generators. The median number of pick-ups per week is five. Close to half of all businesses generate recyclables. Contracts differ in whether the price is charged per cu/yd or per lbs. Volume-based contracts are converted to weight-based contracts using conversion rates for municipal waste.[16] The data records both the monthly charges per unit as well as the total monthly billings and the quantity. For later analysis in the paper I construct an inclusive price variable by using monthly charges and dividing them by the reported monthly quantity. This assures that the resulting price will account for other charges such as the collection containers. In many cases the inclusive price that is derived this way is different from the quoted price per volume.[17] For contracts that were originally denoted in lbs. the converted quantity is used to perform this calculation. One data caveat is that the final charges of brokers to customers are only observed for the year 2014 and that they are not matched to the customer register contract data. I therefore run a hedonic regression of the broker charges on the set of observables that are available in the broker dataset as well as the customer register and impute the broker fee for a given contract as the predicted value from this regression. See Appendix A for details.

2.2 The New York City Market

With very few exceptions all businesses and other private institutions in New York City are required to have a contract with one of the private waste carters. On the buyer side this market therefore encompasses all commercial entities as well as other private institutions in New York and therefore a substantial fraction of the overall waste generated by the city.[18] Carter service consists mostly of the pickup and hauling of recyclable and non recyclable refuse.[19] After the pickup, the waste is hauled to one of the 61 transfer stations in or around New York (New York locations are shown as white triangles in Figure 1). This part of the waste market is typically referred to as the *direct haul* market.[20] Most of the carters operating in New York are not vertically integrated with the transfer station and therefore need to individually negotiate tipping fees with the transfer stations.[21] The haul from the transfer station to the final disposal facility is called the *ultimate disposal* market. Recyclables are typically transferred to one of the in-state processing facilities and non-recyclables predominantly exported to landfills in Pennsylvania or West Virginia.[22] This project considers the *direct haul* market.

The yearly market volume of the trade-waste industry in New York is about \$352 million, and it

[16]See, for example, the standard weight to volume conversion factors of the EPA, http://www.epa.gov/osw/conserve/tools/recmeas/docs/guide_b.pdf (accessed on 04.14.15)

[17]Asker and Cantillon (2008) theoretically study the case where buyers care about multiple dimensions of a sellers offer and show that this problem can be reduced to single-dimensional problem under a scoring rule. Bajari and Lewis (2009) provide an empirical investigation of highway procurements when the government evaluates sellers both with regard to their price and the time for completion of the project.

[18]According to 2011 estimates of the NYC Mayor's Office of Long Term Planning and Sustainability as seen in: http://www.earth.columbia.edu/sitefiles/file/education/capstone/Capstone-Final-Report.pdf

[19]The typical volume of one of the rear loaders used in New York City is about 20 to 30 cu/yd, and the median monthly quantity generated by a business in the city is about 8.66 cu/yd. I am not looking at the market for hazardous material, which is subject to much more stringent regulation.

[20]See for example the summary of the DOJ regulation of waste management services for the OECD: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf... (accessed on 05/09/2015)

[21]The fee to dispose waste at a processing facility.

[22]http://www.dec.ny.gov/docs/materials_minerals_pdf/frptbeyondwaste.pdf (accessed on 06.05.15). Before 2001 much of the waste was going to the Fresh Kills landfill, one of the largest on earth in that time frame (Royte (2007)).

accounts for 3.9 million tons of waste.[23] On average there are 94 active carters per reporting period (half-year) who serve 110,000 customers. To give a sense of the concentration of the industry: averaged over all reporting periods, the four biggest firms serve 37% of all customers, the seven biggest firms 48% of all customers and the ten biggest firms 55% of all customers.[24] Customers are the cross-section of businesses that operate in New York City and generate waste. A breakdown by business-type is provided in Table 9 in Appendix G.

A salient feature of the New York Market is the large number of suppliers that serve a geographic area. On average a zip-code is served by 20 carters (Figure 1), which is more than a fifth of the total number of operating firms. Most of the more densely populated zip-codes are served by at least 30 carters and some zip-codes are served by more than 45 carters, which is almost half the number of total firms that are active in the market at a given point in time. The fragmented supply is surprising, especially since carters services allow relatively little room for horizontal product differentiation. The following observation from an article in The New Yorker emphasizes this point:

"*When I recently walked down a four-block stretch of Broadway on the Upper West Side of Manhattan, I identified about forty businesses — restaurants, clothing shops, bodegas, banks. Licenses in windows listed the commercial-waste haulers they use — at least fourteen in all, by my count, for a stretch that covers only a fifth of a mile. If there was a pattern, I couldn't grasp it: the Starbucks at Ninety-third and Broadway uses a different commercial-waste company from the Starbucks at Ninety-fifth and Broadway.*" — **The New Yorker, 2009**[25]

The large number of suppliers in a zip code has also received attention from officials, and the city started to collect route information from carters to inform a decision about a change to a procurement system with exclusive territories.[26] Prices in the market are individually negotiated between customer and carters. Unlike in most retail settings, searching in this context therefore requires calling an individual carter and haggling. The large number of suppliers and the idiosyncratic nature of the arrangements suggests that this search process is costly for buyers.[27]

[23]The publicly administered market creates an additional 3.8 million tons. See Commission (2012). Average value of total yearly billings. Missing values for total billings have been replaced with the mean value of total billings. Medical waste and waste shredders are excluded.

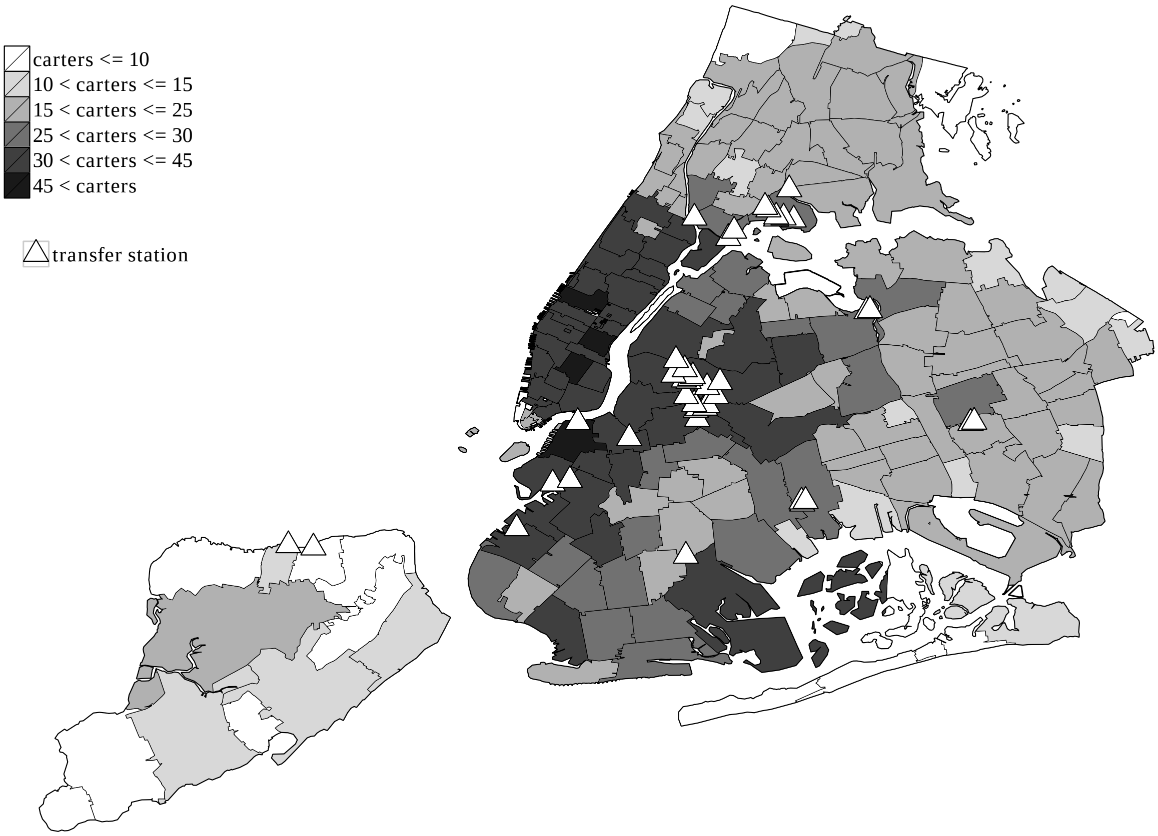
[24]Because of confidentiality reasons, a more complete picture of the firm-size distribution is not provided.

[25]http://www.newyorker.com/business/currency/a-better-way-to-take-out-the-garbage (accessed on 04/03/2015)

[26]See for example: http://citylimits.org/2015/05/19/city-weighs-reining-in-private-garbage-collectors/ (accessed on 07/11/2015)

[27]There are several potential explanations for the large number of local suppliers, which stands in contrast to the consolidation in other parts of the country. It is well known that historically the waste industry in New York was captured by organized crime (see for example Jacobs et al. (2001)). During the time in which the waste industry was consolidating on a national scale, New York City was still in the grip of a property-rights and racketeering system. According to Kelly (1999) this system was in place for more than 50 years and the New York District Attorney's office estimated over-billings of 30% to 40%, which was regarded as a "garbage tax" for doing business. In 1995 a New York City grand jury indicted 23 carters for price fixing, bid rigging, racketeering, corruption and the establishment of a property rights system. According to the District Attorney's office small business owners were paying $15,000 a year for the waste removal services and restaurants about $50,000. The article also mentions that while in other parts of the country, many firms were replaced by high-technology entrants, New York was still served by 600 labor-intensive small carters. This might explain why consolidation is delayed in the city and that there is still a large number of relatively small carters. Another potential explanation is the population density of the city. Anecdotally, route density is an important aspect for carters to reduce cost, leading to strong network effects. See also: Nguyen and Wilson (2010). New York City, however, is so densely populated that route density by itself might not be an important margin for overall cost reductions compared to other cities or rural areas.

Figure 1: Long-run average number of active carter per zip code



Note: This map shows the number of carters that are active in a zip-code, averaged over time. "Active" is defined as having at least one customer in the zip-code. Triangles show the location of the transfer stations.

2.3 BROKERS

Trade-waste brokers are a potential remedy for customers' search problem and their services are particularly valuable in markets like New York. Conversations with brokers reveal that their service allows customers, who often operate on a national scale, to have a one-stop shop to deal with the fragmented landscape of waste removal services. Since many carters in New York City are relatively small, they often are not able to offer the kind of easy access to their services that large companies can afford. Such access might be provided through a web-interface that allows customers to get an automatic price quote tailored to their size, location, etc. The burden of match-making is therefore predominantly on the customer side. Unlike customers, trade-waste brokers know which carters are available to service customers and have established contacts with a subset of available carters.[28] Conversations with brokers reveal that they award contracts through a *Request for Proposals* - a competitive bidding process

[28] Brokers arrange contracts between businesses and carters; they are not allowed to accept direct payments by carters, see http://www.nyc.gov/html/bic/downloads/pdf/regulations/tw_title17_chap_1.pdf (accessed on 08/15/15)

Table 2: Breakdown of Fraction of Contracts awarded through Broker

Variable	Brokered	Variable	Brokered
Types of business		**Borough**	
Retail non - food	.164	Bronx	0.169
Retail - food	.110	Brooklyn	0.151
Wholesale non - food	.147	Manhattan	0.141
Wholesale - food	.099	Queens	0.130
Restaurant/bar	.133	Staten Island	0.053
Hotel - small	.140	**Quantity (percentile)**	
Hotel - big	.173	$q < 25\%$	0.068
Medical offices	.350	$25\% < q < 50\%$	0.136
Automobile repair	.110	$50\% < q < 75\%$	0.155
Office building - small	.073	$q > 75\%$	0.198
Professional office	.066	**Recyclables**	
Office building - large	.101	No	0.132
Institution	.360	Yes	0.180

Note: This table shows the percentage of businesses using a broker conditional on several variables: by borough location, the percentile of the quantity of waste that the customer generates, the business type as well as recyclables.

akin to a first price auction - to these carters.[29] A recent article in the New York Times, which portraits one of the large trade-waste brokers, reiterates some of these points:

"[...] Two big national companies, Waste Management and Republic Services, dominate the market, owning fleets of trucks and hundreds of landfills. Thousands of smaller, regional trash haulers fill in the gaps. Rubicon, based in Atlanta, isn't in the business of hauling waste. It doesn't own a single truck or landfill. [...] It begins by holding an online bidding process for its clients' waste contracts, fostering competition among waste management businesses and bringing down their prices. [...] Through a combination of big data and online auctions for hauling contracts, Rubicon says it reduces clients' waste bills by 20 percent to 30 percent. [...]" — **The New York Times, 2015**[30]

The identification of the model will build on the fact that the price formation of brokered contracts works according to the well understood mechanism of a first price auction. This helps to pin down the supply-side cost distribution from the subset of brokered contracts. Once the cost of service provision is known, the variation in search cost can be identified from the search market.

The fraction of brokered contracts is relatively stable across different business types, locations and volume of buyers. About 13% percent of businesses indicate that they arranged contracts through a broker. Small waste generators in the first quartile have, at 6.8%, a lower propensity to use brokers

[29] As part of this research I have spoken to many brokers individually on the phone. All of them described that contracts are awarded through a competitive bidding procedure.

[30] www.nytimes.com/2014/10/26/business/dividing-and-conquering-the-trash... (accessed on 3.7.2015)

than the three remaining quartiles.[31] Table 2 provides a more detailed view showing the percentage of brokered contracts conditional on the Borough location, the types of businesses, the quantity of waste that businesses generate as well as whether they produce recyclables. We can see that the percentage of businesses in Manhattan using brokers is at 14.1%, about the same as Queens at 13.0% and Brooklyn at 15.1%. The Bronx has the highest percentage of brokered contracts at 16.9%. The percentage is significantly smaller in Staten Island at 5.3%. Regarding types of businesses, we see that for non-food retail and wholesale businesses, large hotels and institutions have a higher likelihood of using brokers.[32]

3. DESCRIPTIVE RESULTS

This section establishes several facts about the prices in the market. First it is shown that there is large variation in prices and residual prices, which are obtained by conditioning on a comprehensive set of observable price shifters as well as fixed effects. Such dispersion points at large expected returns to searching for customers, even in the case where sellers account for observable information in their prices. Second, a comparison of brokered prices and search-market prices provides evidence that brokers are used by those customers that have a relatively higher opportunity cost of searching. Lastly, I report direct correlations that are consistent with an externality of brokered contracts.

3.1 EVIDENCE OF PRICE DISPERSION

I document both the variation in raw prices p_{ijt} and, following Allen et al. (2013), the dispersion in residual prices $\tilde{p}_{ijt}$ to account for systematic pricing of observable information.

$$p_{ijt} = \mathbf{X} \cdot \boldsymbol{\beta} + \tilde{p}_{ijt}$$

The percentage of explained variation and the residual price dispersion are reported from two sets of regressions. The left hand side variable of these regressions is the price (per cu/yd) that carter j charges customer i during observation period t, which is a half-year. I only include the initial price of each contractual relationship. Both regressions include a common set of controls and differ in whether carter fixed effects are included:

[31] Note that a larger quantity of waste does not necessarily mean a larger business in terms of revenue, but one expects the two to be correlated. There might be multiple reasons for this pattern. It could be due to the differentially stronger discounts to large quantities in the broker market. In the following section I demonstrate that brokers do price observable information, such as the quantity of waste generated by a business, more systematically. A second reason could be that sales offices need to go to a "procurement procedure" through a broker to comply with audit rules.

[32] The model, which will be introduced later, assumes that the substitution between brokers and the search-market is solely based on search cost. But there might be other factors, such as service quality and reliability, which are driving the decision to search through a broker. One reason that speaks against this is the fact that virtually all carters that are active in the broker market also serving contracts in the regular market. It is hard to imagine that the workers that serve both types of contracts on a given route will know which locations have been brokered and treat them differently. In addition, one would expect that most customers keep a contract brokered if the broker service is mostly about service reliability as opposed to search frictions. **??** in Appendix E shows, however, that about half of the buyers that stay with the same carter drop the broker after three years.

$$\mathbf{X} = \{\text{business type FE, recyclables FE, time FE, zip code FE, transfer station FE}, q, \ldots, q^5, \\ \text{Number of Pickup FE}\} \times \{\text{No Carter FE, Carter FE}\}.$$

Both types of regressions are repeated separately for contracts in the brokered and in the search-market. For each of the regressions I document the percentage of variation that is left unexplained, i.e. $1 - R^2$, and the dispersion of the residual price $\tilde{p}_{ijt}$. Figure 2 shows the results of these regressions.

Figure 2: Documenting price dispersion

	First specification (no carter FE)		Second Specification (carter FE)	
	Not Brokered	**Brokered**	**Not Brokered**	**Brokered**
$1 - R^2$	0.59	0.34	0.54	0.32
$SD(p_{ijt})$	7.97	7.06	7.97	7.06
$SD(\tilde{p}_{ijt})$	6.00	4.08	5.75	3.95
$mean(p_{ijt})$	9.00	7.17	9.00	7.17

Note: The table shows the percentage of unexplained variation in observed prices p_{ijt} as well as the dispersion of rates and residual rates $\tilde{p}_{ijt}$. The the second specification is identical to the first one but also includes carter specific fixed effects. Included control variables are described in the text above.

Despite the fine-grained controls the percentage of unexplained price variation is substantial across specifications, and the standard deviation in observed and residual prices bear little difference. The inclusion of carter fixed effects does not reduce unexplained variation substantially, which shows that even within carter price variation is large. The mean and the standard deviation for brokered contracts are smaller. Brokered contracts therefore have smaller price dispersion to begin with and prices on these contracts are better explained by observable variation.

As a measure of price dispersion the literature often refers to the coefficient of variation which divides the sample standard deviation by the sample mean of the price distribution. The coefficient of variation is 0.88 for contracts in the search-market and 0.99 for brokered contracts when computed from the raw data. Instead, using the standard deviation of the residual rate, the coefficient of variation is 0.66 (search-market) and 0.57 (brokered market) without carter fixed effects and 0.64 (search-market) and 0.55 (brokered market) with carter fixed effects. As a comparison, in a well-known empirical study of price dispersion, Sorensen (2000) reports an average coefficient of variation of 0.22 in the retail market for prescription drugs.

The following calculations give a sense of the $-value of this dispersion: For a business that reports in cu/yd and generates the median quantity of waste (9.00 cu/yd), moving one standard deviation ($SD(\tilde{p}_{ijt})$) in the residual distribution of prices would imply an extra $1425.6 over the length of a contract, which is slightly longer than two years. The same calculation for a business with the mean quantity of waste (cu/yd 27.6) would imply an extra $4371.8 over the typical length of a contract, which is two years. This is after using all the observable information that is available to the econometrician, which in this particular market setting might be better than the information available to searching businesses.[33]

3.2 Comparing prices across Broker Market and Search Market

3.2.1 Evidence for the selection of Buyers

This section establishes that the following holds both for conditional (on relevant contract observables) and unconditional average prices: $\overline{Price}_{broker-market} + \overline{Commission} > \overline{Price}_{search-market} > \overline{Price}_{broker-market}$. This relationship calls for an unobserved variable that makes buyers who use broker services differentially inelastic to the prices charged in this part of the market and is therefore consistent with the interpretation that those buyers have higher search cost than buyers that contract in the search-market.

Figure 3 shows that this clearly holds when examining the raw CDFs of prices. We see that the distribution of brokered prices with commission stochastically dominates the distribution of non-brokered contracts, which in turn dominates the distribution of broker prices without commission. Brokers charge substantial markups over the rates that are obtained from carters in the auction. The median commission is 57.0%, and the mean commission is 70.0%.[34]

Results from two sets of regressions (shown in Table 3 and Table 4) confirm that this relationship also holds after accounting for the observable differences in contracts. The two sets differ only in the dependent variable. The first one includes broker prices without commissions and the second set with commission. The main variable of interest is a dummy that indicates whether the contract is brokered or not and I explore both the mean effect in an OLS specification as well the effect at different points of the distribution using quantile regressions. All of the regressions include the following controls: quantity of waste, transfer-station fixed effects, zip-code fixed effects, business-type fixed effects, length-of-contract fixed effects, recyclable materials fixed effects, reporting-date fixed effects, number of weekly pickups, and the zip-code HHI index.[35] The estimated linear effect of the broker dummy from the first specification in Table 3 is -1.233, which is about 12.14% of the mean price and shows that brokers obtain lower average prices from carters after observable information is taken into account. Results from the quantile regressions reveal that this difference is composed of a weaker difference in the lower tail of the distribution (25th percentile) with an effect of -0.206 and a stronger effect in the upper tail, which is -1.182 for the 50th percentile and an almost identical -1.184 for the 75th percentile.

[33] An additional way to put these numbers into perspective is to compare them to the revenues of a business in the city. Most businesses in New York City are small. According to the 2013 County Business Patterns provided by the Census Bureau, 60,856 out of 105,439 businesses have less than four employees, and 77,965 have less than 10 employees. BizBuySell.com provides the median sales price of a business in New York City, that have been sold over this platform, which was about $229,500 in 2013. Assuming an interest rate of 3% that would imply annual profits of $6885.0. The subset of businesses sold on this page is almost surely biased towards small companies, but it provides some reference point for the above calculations.

[34] Those commissions are decreasing in the quantity generated by customers: in the lowest quartile of generated quantity the median commission is 88.9%, the second lowest quartile 67.5%, in the second highest quartile of monthly billings it is 58.7% and in the highest quartile 65.8%.

[35] Goldberg (1996), for example, uses quantile regressions to re-examine experimental results on race-based price discrimination. The quantile regressions do not include controls for transfer stations. This is because it is computationally very costly to include that many controls. In the linear regressions the effect of the broker dummy is not affected much by whether transfer station fixed effects are included or not.

Figure 3: CDF for non-brokered and brokered prices (with and without commission)

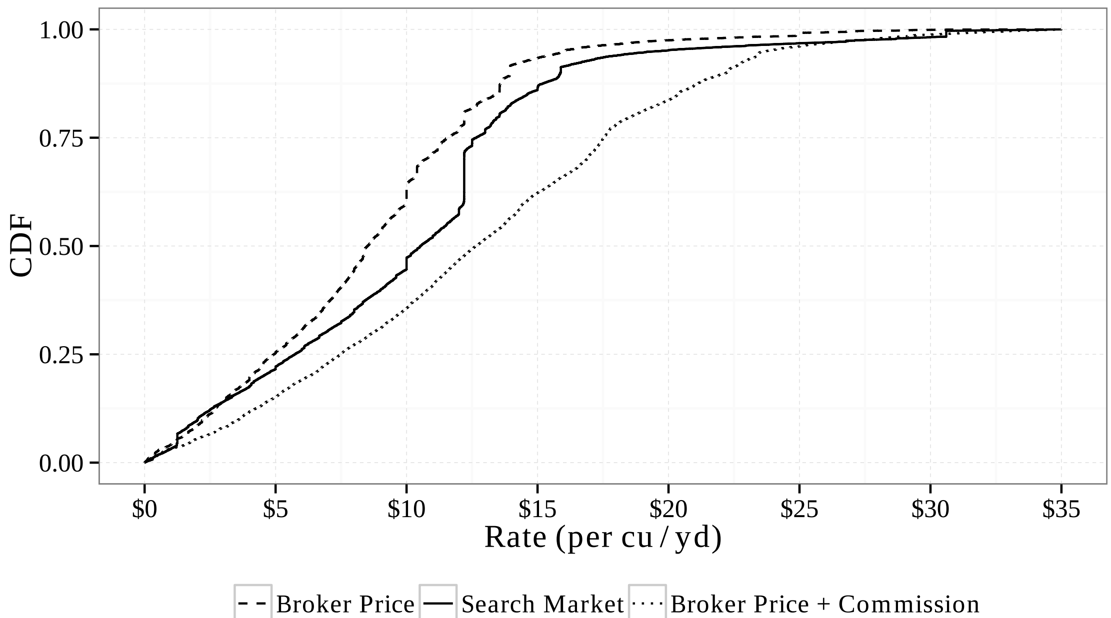


Table 3: Documenting differences between brokered and un-brokered contracts.

	(1)	(2)	(4)	(5)	(6)
	OLS	OLS	Quantile (0.25)	Quantile (0.50)	Quantile (0.75)
	p_{ijt} (rate)	p_{ijt} (rate)	p_{ijt} (rate)	p_{ijt} (rate)	p_{ijt} (rate)
Broker	-1.232**	-1.165**	-0.208**	-1.182**	-1.187**
	(0.117)	(0.110)	(0.0261)	(0.0333)	(0.0283)
Quantity	-0.0192**	-0.0190**	-0.0117**	-0.0132**	-0.0167**
	(0.000715)	(0.000701)	(0.000154)	(0.000206)	(0.000189)
Recyclables	-0.721**	-0.770**	-3.695**	-4.810**	-3.810**
	(0.178)	(0.170)	(0.0250)	(0.0317)	(0.0272)
Deals with Broker		-1.469**			
		(0.283)			
Observations	92245	92245	100942	100942	100942
Deals with Broker	No	Yes	No	No	No
Transfer FE	Yes	Yes	No	No	No
R^2	0.272	0.274			

Note: $+\ p < 0.10$, $*\ p < 0.05$, $**\ p < 0.01$. All specifications include the following set of controls: quantity of waste, transfer-station fixed effects, zip-code fixed effects, business-type fixed effects, length-of-contract fixed effects, recyclable materials fixed effects, reporting-date fixed effects, number of weekly pickups, and the HHI index. The aggregate regressions at the zip-code level include the average quantity at the zip-code level, the average number of pickups and the average number of customers that use recyclables. Standard errors are clustered at the zip-code level.

Table 4 shows the same set of regressions where broker prices include the commission.[36] The coefficient on the main independent variable of interest, the indicator whether a contract is brokered or not, is now reversed from negative and significant to positive and significant. The effect implies that final prices are about $3.35 higher for brokered compared to non-brokered contracts, holding other characteristics fixed. The difference is again larger at the high end of the distribution. This confirms the observation in the raw price distributions in Figure 3.

Table 4: Comparing final prices for brokered contracts with prices in non-brokered market

	(1)	(2)	(4)	(5)	(6)
	OLS	OLS	Quantile (0.25)	Quantile (0.50)	Quantile (0.75)
	p_{ijt} (rate)	p_{ijt} (rate)	p_{ijt} (rate)	p_{ijt} (rate)	p_{ijt} (rate)
Broker	3.139**	3.323**	1.946**	2.407**	4.333**
	(0.317)	(0.305)	(0.0297)	(0.0319)	(0.0312)
Quantity	-0.0233**	-0.0228**	-0.0141**	-0.0168**	-0.0200**
	(0.000868)	(0.000857)	(0.000173)	(0.000197)	(0.000210)
Recyclables	-0.696**	-0.693**	-3.851**	-4.854**	-3.750**
	(0.185)	(0.182)	(0.0284)	(0.0303)	(0.0297)
Deals with Broker		-2.361**			
		(0.178)			
Observations	92125	92125	100765	100765	100765
Deals with Broker	No	Yes	No	No	No
Transfer FE	Yes	Yes	No	No	No
R^2	0.279	0.285			

Note: $+\ p < 0.10$, $*\ p < 0.05$, $**\ p < 0.01$. All specifications include the following set of controls: quantity of waste, transfer-station fixed effects, zip-code fixed effects, business-type fixed effects, length-of-contract fixed effects, recyclable materials fixed effects, reporting-date fixed effects, number of weekly pickups, and the HHI index. The aggregate regressions at the zip-code level include the average quantity at the zip-code level, the average number of pickups and the average number of customers that use recyclables. Standard errors are clustered at the zip-code level.

3.2.2 Evidence for the selection of Sellers

The structural model, which is introduced later, identifies the cost of sellers from the broker market and then uses those cost distributions to back out search cost in the search market. For the validity of this approach it will be important to account for potential, systematic selection of sellers across the two markets. Specification two in Table 3 (and Table 4) provide evidence for such selection. Compared to specification one, this specification includes an additional dummy variable, which indicates whether a carter has obtain at least one contract through a broker (not necessarily the contract of this

[36] As I pointed out in the data section, broker prices are only available for a subset of the data period and at this time not linked to the data from the customer register. To match commissions I estimate a non-parametric regression of the commission on a fifth order polynomial in terms of the charges to carters as well as zip-code dummies. For the customer register I then impute the predicted values from this regression. I also conduct a robustness check in which I, instead of using the predicted commission, match the 5th percentile (19.45%) of commissions. In each case the coefficient of interest, the broker dummy, is still positive and significant. The results of these regressions are shown in Table 10 in Appendix G.

observation).[37] The inclusion of this additional control reduced the difference between brokered and non-brokered contracts to -1.165. More importantly, the additional dummy variable itself is negative and significant at -1.467. The comparison of the two specifications suggests that some of the differences across the two markets is due to the fact that brokers predominantly deal with carters that quote persistently lower prices than carters in the search market. The empirical model will take this selection into consideration and confirm that brokers obtain lower prices by contracting predominantly with persistently cheaper carters.

3.2.3 Direct Evidence of Externality

One key question raised in this study is whether the change of the buyer composition in the search-market through brokers has an effect on the prices in this part of the market. Figure 4 reports an additional set of regressions, which provide some direct evidence for such an externality. The independent variable of interest is now the fraction of brokered contracts at the zip code level. Specifications one and three have the individual price of customers as the dependent variable and specifications two and four the average price at the zip-code level. All regressions only include initial rates from the first year of a customer-carter relationship. Included controls are year and zip-code fixed effects, (average) recyclables, (average) quantity. In the individual level regression additional controls are carter fixed effects, business type fixed effects, and recyclable materials. Specification one includes all contracts and the average price in specification two is over all rates. The results show that a higher fraction of brokered contracts at the zip code level goes along with lower rates. Importantly, this is true even if all brokered contracts are dropped from the sample as in specifications three and four, which suggests that more brokered contracts can benefit buyers who are not directly contracting through brokers.

3.3 Summary and Interpretation of Descriptive Results

In this section I showed that residual variation in prices, after conditioning on many important observable aspects of a contract, suggests a high value of search to customers. Brokers negotiate cheaper prices with carters, especially at the higher end of the price distribution. Brokered contracts also account more systematically for observed variation in contracts, and prices are less dispersed. But because of high commissions the average price paid by customers who use brokers is higher, both in the raw data and after accounting for observables. These two facts together strongly suggest that buyers who are seeking the services of brokers have higher average search cost, which allows brokers to charge such high markups for their services. Lastly, regressions at the zip-code level show that a higher percentage of brokerage is associated with lower prices, even for businesses that do not use brokers.

While the last result is suggestive evidence in favor of an externality of brokered contracts, a direct causal interpretation of these results is problematic for obvious reasons. The percentage of brokered contracts is not randomly assigned and might be confounded by time varying unobservable variation that makes a zip-code cheaper to service and businesses more inclined to use brokers. To provide a more meaningful understanding of the quantitative effects of the activity of brokers, I will now turn to a model that captures customers' search strategy of as well as carters pricing strategy and makes explicit assumptions on the unobservables governing their behavior.

[37] To be precise, this indicator is one if at least one contract of the carter is arranged through a broker and zero otherwise.

Figure 4: Varying amount of brokerage and average prices in a zip code.

	(1 OLS) r_{ijt} (rate)	(2 OLS) $mean(r_{zt})$ (rate)	(3 OLS) r_{ijt} (rate)	(4 OLS) $mean(r_{zt})$ (rate)
Fraction of Brokered (ZIP)	-1.645**	-3.490**	-1.751**	-2.259+
	(0.485)	(0.715)	(0.619)	(1.233)
Quantity (Winsorized 1%)	-0.0191**	-0.0481**	-0.0186**	-0.0406**
	(0.000687)	(0.0107)	(0.000773)	(0.0135)
Recyclables		-1.948**		-1.883*
		(0.615)		(0.787)
Observations	310830	1045	269311	1045
Brokered Contracts Included	Yes	Yes	No	No
Year FE	Yes	Yes	Yes	Yes
ZIP Code FE	Yes	Yes	Yes	Yes
Carter FE	Yes	No	Yes	No
Business Type FE	Yes	No	Yes	No
R^2	0.275	0.622	0.284	0.584

Note: $+\ p < 0.10$, $*\ p < 0.05$, $**\ p < 0.01$. The individual level specifications include the following set of controls: quantity of waste, transfer-station fixed effects, zip-code fixed effects, business-type fixed effects, recyclable materials fixed effects, year fixed effects, number of weekly pickups. Standard errors are clustered at the zip-code level.

4. MODEL OVERVIEW

To establish a causal effect of the availability of intermediaries in the market, one would ideally compare markets with different and randomly assigned prices for broker contracts and test how markets fare under different conditions. However, such exogenous variation does not exist in this setting. For the evaluation of the effect of brokers, one needs to take into consideration that different observed fractions of brokered contracts might be due to different inherent market characteristics that affect both the propensity of brokerage and prices. The model will help to understand how the observed prices depend on customer search cost and the cost of service provision for sellers as well as account for the potential selection of customers and sellers response in pricing behavior.

The search of businesses for brokers is modeled as a sequential game between customers and carters. Brokers are non-strategic players. A customer j in the model is described by privately observed *iid* search cost κ_j, drawn from a continuous distribution $\mathcal{H}(\kappa|\mathbf{x}_j)$. These search costs are the marginal cost for an additional price inquiry. While I call κ_j *search* cost, one should interpret it to also capture the haggling cost to get the "best" price quote from a carter. Carters, on the other hand, are described by their cost, which are a function of observable and unobservable characteristics: $\mathcal{C}_{ij}(\mathbf{z}_j, c_{ij})$, where c_{ij} is an *iid* draw from a continuous distribution $c_{ij} \sim \mathcal{G}(.|\mathbf{z}_j)$. Variables $\mathbf{x}$, which affect search cost, and $\mathbf{z}$, which affect carter cost are customer-specific observable differences. The quantity q generated by the customer determines the marginal cost of searching directly as higher quantity customers benefit more from a given decrease in prices. The quantity might also determine the distribution of

$\mathcal{H}(\kappa|\mathbf{x})$ and is therefore included in $\mathbf{x}$. For the identification of the model it will be important that there is some variable that changes the price distribution only through its effect on carter cost while leaving the search expense distribution unaffected. In the estimation I will use the recyclables generated by customers as such an exclusion restriction.

Customers' search in the model is non-sequential. As I outlined in section 1 there are several reasons favoring the assumption of non-sequential search in this context. Firms' pricing decision under this assumption is equivalent to an auction with a stochastic number of competitors, which allows me to draw on the extensive set of empirical tools developed in this literature.[38] Empirical studies that test how people search in practice also favor non-sequential search (De los Santos et al. (2012) and Honka and Chintagunta (2014).

The timing of the game is as follows: *In period zero* customers decide whether to delegate search to a broker or not and if not how many prices $m \in \{1, \ldots, M\}$ to draw from the distribution, where M is the number of firms in the search market. *In period one* carters make their price quotes either in a first price auction when the contract is procured through a broker or offer a price quote to a customer in the search-market where they have uncertainty about the number of rival price quotes. The pricing is conditional on the customer-specific observable information $\mathbf{x}_j$ and $\mathbf{z}_j$. Such information, for example, includes the borough location, the type of business of the customer, the quantity q_j of waste generated by the customer as well as its composition.[39] Carters would, for example, anticipate that someone with higher quantity has larger incentives to search than someone with a low quantity.

4.1 Customer Search and Brokers

In the model I abstract from the competition between brokers. Instead, I assume that buyers make their decision to delegate the search based on the average brokered price (conditional on relevant observables). Brokers determine the seller through a competitive bidding process. I assume that searching costumers incur expenses for each additional price inquiry while I regard the broker infrastructure as fixed. A broker b will each time she receives a request by a customer hold an auction with N_b bidders. Let the expected price obtained through a broker be $\mathbb{E}[p^B|\mathbf{x},\mathbf{z}]$. This price is the average lowest bid over different broker auctions with varying number of competitors. In line with the request for proposal procedure I model these auctions as first price auctions. On the customer side each search expense type κ will optimally ask for $m(\kappa)$ price quotes, which yields an expected price $\mathbb{E}[p^{m(\bar{\kappa})}|\mathbf{x},\mathbf{z}]$. In the next section I will provide details on how $\mathbb{E}[p^B|\mathbf{x},\mathbf{z}]$ and $\mathbb{E}[p^{m(\bar{\kappa})}|\mathbf{x},\mathbf{z}]$ form. The markup charged by brokers is denoted as $\phi(\mathbf{x},\mathbf{z})$. For a given set of observables $\mathbf{x}$ and $\mathbf{z}$ there exists a cut-off search-cost type $\bar{\kappa}(\mathbf{x},\mathbf{z})$ who is indifferent between an arrangement with a broker and the expected cost of individual search under the optimal search policy:

$$q \cdot \mathbb{E}[p^B|\mathbf{x},\mathbf{z}] \cdot \phi(\mathbf{x},\mathbf{z}) = q \cdot \mathbb{E}[p^{m(\bar{\kappa})}|\mathbf{x},\mathbf{z}] + m(\bar{\kappa}(\mathbf{x},\mathbf{z})) \cdot \bar{\kappa}(\mathbf{x},\mathbf{z}) \tag{1}$$

Let $\mathcal{F}(p|\mathbf{x},\mathbf{z})$ be the equilibrium distribution of price offers and remember that $\kappa \sim \mathcal{H}(.|\mathbf{x})$ is the

[38] In an earlier stage of this project I also experimented with a sequential search model. For some parameter values it was not possible to solve for an equilibrium in these models. This finding along with the lack of general existence proofs for such an environment are other reasons to use the non-sequential search model, where such numerical problems never came up.

[39] The quantity as well as the composition of waste is assumed to be exogenous.

distribution of marginal search cost and $c \sim \mathcal{G}_k(.|\mathbf{z})$, $k \in \{L, H\}$ the distribution of cost draws for carters.[40] I will now discuss how the optimal search strategy of customers against such an equilibrium price offer function looks like. Denote $p^{(m)}$ the lowest price out of m sampled prices. The problem of a customer with search cost κ is to minimize his total cost over the number of searches $m \in \{1, \ldots, M\}$:

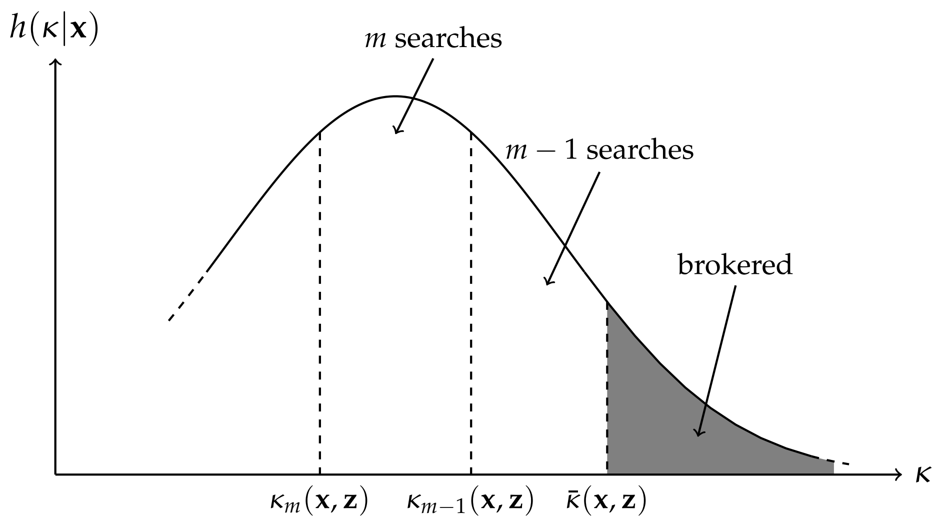
$$\min_m q_j \cdot \mathbb{E}[p^{(m)}|\mathbf{x}, \mathbf{z}] + m \cdot \kappa$$

using the distribution function of the lowest price in terms of the equilibrium price offer distribution $\mathcal{F}(p|\mathbf{x}, \mathbf{z})$, the expected cost for making m searches can be expressed as:

$$\min_{m \in \{1,\ldots,M\}} \int_0^{\bar{p}} m \cdot p \cdot q \cdot (1 - \mathcal{F}(p|\mathbf{x}, \mathbf{z}))^{m-1} \cdot f(p|\mathbf{x}, \mathbf{z}) dp + m \cdot \kappa$$

The following lemma says that buyers, for a given distribution of price offers, will sort themselves according to the optimal number of price inquiries that they want to make. Depending on $\mathbb{E}[p^B|\mathbf{x}, \mathbf{z}] \cdot \phi(\mathbf{x}, \mathbf{z})$ there is a marginal type $\bar{\kappa}(\mathbf{x}, \mathbf{z}) < \infty$ so that every type with higher search cost will delegate the search to a broker.

Figure 5: The sorting of buyers according to search cost



Notes: This figure provides a graphical illustration of the sorting of buyers according to their search cost. Higher search cost lead to fewer calls m. Types with search cost above $\bar{\kappa}(q, \mathbf{z})$ delegate the search to a broker.

Lemma 1. *For a combination of* $\mathbf{x}$ *and* $\mathbf{z}$ *there are marginal types* $0 \leq \kappa_M(\mathbf{x}, \mathbf{z}) < \ldots < \kappa_m(\mathbf{x}, \mathbf{z}) < \kappa_{m-1}(\mathbf{x}, \mathbf{z}) < \bar{\kappa}(\mathbf{x}, \mathbf{z}) \leq \infty$ *so that every type* $\kappa \in [\kappa_m(\mathbf{x}, \mathbf{z}), \kappa_{m-1}(\mathbf{x}, \mathbf{z})]$ *samples* m *firms and every type larger than* $\bar{\kappa}(\mathbf{x}, \mathbf{z})$ *delegates search to an intermediary.*

Proof: This Lemma follows from the fact that $\mathbb{E}[p^{(m)}|\mathbf{x}, \mathbf{z}]$ is concave and the search cost linear in the number of searches m. Note that these cut-off types are given by $\kappa_m(\mathbf{x}, \mathbf{z}) = q \cdot (\mathbb{E}[p^{(m)}|\mathbf{x}, \mathbf{z}] -$

[40]The distribution $\mathcal{F}(.|\mathbf{x}, \mathbf{z})$ itself is unobserved. The data only record the contract price which was the lowest price offered by the carters.

$\mathbb{E}[p^{(m+1)}|\mathbf{x},\mathbf{z}])$. Because of the sorting it must be true that if $q \cdot \mathbb{E}[p^B|\mathbf{x},\mathbf{z}] + \phi(\mathbf{x},\mathbf{z}) > \mathbb{E}[p^{(m)}|\mathbf{x},\mathbf{z}] - \kappa \cdot m$ that it must also be true that $q \cdot \mathbb{E}[p^B|\mathbf{x},\mathbf{z}] + \phi(\mathbf{x},\mathbf{z}) > q \cdot \mathbb{E}[p^{(m)}|\mathbf{x},\mathbf{z}] - \kappa \cdot (m+1)$ ■

Brokers provide "rents" to every type above $\bar{\kappa}(\mathbf{x},\mathbf{z})$. Figure 5 visualizes the sorting of buyers along the search-cost distribution into bins of types that want to make m searches and the selection into the broker market.

4.2 Carter Pricing

I now turn to carters' pricing decision. The empirical section showed that carters who deal with brokers are in general cheaper. To account for this selection, the model will allow for carters of type H and type L and their composition to vary across the brokered and search market. The number of firms in the search market are denoted N_H and N_L.

Denote $\widetilde{\mathcal{G}}(.|\mathbf{Z}) = 1 - \mathcal{G}(.|\mathbf{z})$. The optimal bidding functions $\beta_{b,k}(c|\mathbf{x},\mathbf{z})$, $k \in \{L, H\}$ in the procurement auction of Broker b with $N_b = N_{b,L} + N_{b,H}$ bidders are derived from the following objective function (for bidder of type H):

$$\max_p (p-c) \cdot \underbrace{\widetilde{\mathcal{G}}_L(\beta_{b,L}^{-1}(p)|\mathbf{z})^{N_{b,L}} \cdot \widetilde{\mathcal{G}}_H(\beta_{b,H}^{-1}(p)|\mathbf{z})^{N_{b,H}-1}}_{\substack{\text{Probability that } p \text{ is lower than} \\ \text{prices offered by } N_{b,L} \text{ and } N_{b,H} \text{ competitors}}} \tag{2}$$

In the search market carters make their price quotes not knowing the type κ of the consumer. They know they will be picked if they are the cheapest amongst m firms, where m is a multinomial random variable with probabilities $w_m(\mathbf{x},\mathbf{z}) = \mathcal{H}(\kappa_{m-1}(\mathbf{x},\mathbf{z})|\mathbf{x},\kappa < \bar{\kappa}) - \mathcal{H}(\kappa_m(\mathbf{x},\mathbf{z})|\mathbf{x},\kappa < \bar{\kappa})$. I will now describe how carters form their price offer functions $\beta_{S,k}(c|\mathbf{x},\mathbf{z})$, $k \in \{L, H\}$ in the sub-game where carters make price offers to consumers in the search market, planning with the correct vector of search weights $(w_1(\mathbf{x},\mathbf{z}) \dots w_M(\mathbf{x},\mathbf{z}))$ upon observing $\mathbf{x}$ and $\mathbf{z}$. These strategies map the customer-specific cost-draw to a price quote. Note, however, that the weights w_m depend on carters' price offers. Using this notation the maximization problem of a carter of type H for customer $(\mathbf{x},\mathbf{z})$ is:

$$\max_p (p-c) \cdot \left[\sum_{m=1}^{M-1} \underbrace{w_m(\mathbf{x},\mathbf{z})}_{\substack{\text{Customer} \\ \text{calls } m \text{ carter}}} \cdot \sum_{k=0}^{m} \cdot \underbrace{\frac{\binom{N_L}{k} \cdot \binom{N_H-1}{m-k}}{\binom{N_H+N_L-1}{m}}}_{\substack{\text{Probability that } k \\ \text{competitors are of type } L}} \cdot \underbrace{\widetilde{\mathcal{G}}_L(\beta_{S,L}^{-1}(p)|\mathbf{z})^k \cdot \widetilde{\mathcal{G}}_H(\beta_{S,H}^{-1}(p)|\mathbf{z})^{m-k}}_{\substack{k \text{ firms of type } L \text{ and } m-k \\ \text{firms of type } H \text{ bid above } p}} \right] \tag{3}$$

This maximization problem is akin to a first-price procurement auction with an unknown number and composition of competitors.[41] Putting together the behavior of carters and the customer search strategy, an equilibrium for the market can be formulated:

Definition 1. *An equilibrium in the decentralized market for customer type $(\mathbf{x},\mathbf{z})$ is a set of:*

[41] An auction where the number of bidders is uncertain and the cost functions symmetric turns out to have a closed-form bidding function, which can be derived using the revenue equivalence theorem as shown in Krishna (2009) for the case of a standard auction: $\beta(c) = \sum_{m=1}^{M} \left[\frac{w_m \cdot (1-\mathcal{G}(c))^{(m-1)}}{\sum_{k=1}^{M} w_k \cdot (1-\mathcal{G}(c))^{(k-1)}} \cdot \left(c + \frac{1}{(1-\mathcal{G}(c))^{(m-1)}} \int_c^{\bar{c}} (1-\mathcal{G}(u))^{(m-1)} du \right) \right]$

1. *Bidding strategies in the broker market:* $\beta_{b,k}(.|\mathbf{x},\mathbf{z})$, $k \in \{L, H\}, b \in \{1, \ldots, B\}$

2. *Bidding strategies in the search market:* $\beta_{S,k}(c|\mathbf{x},\mathbf{z})$, $k \in \{L, H\}$

3. *Customer search cost cut-off types* $\kappa_1, \ldots, \kappa_M$ *and* $\hat{\kappa}$.

Such that $\kappa_1, \ldots, \kappa_M$ *and* $\hat{\kappa}$ *result from customers optimal search behavior under the price distribution* $\mathcal{F}(.|\mathbf{x},\mathbf{z})$ *in the search market and* $\mathcal{F}^B(.|\mathbf{x},\mathbf{z})$ *in the broker market induced by* $\beta_{b,k}(.|\mathbf{x},\mathbf{z})$, $k \in \{L, H\}, b \in \{1, \ldots, B\}$ *and* $\beta_{S,k}(c|\mathbf{x},\mathbf{z})$, $k \in \{L, H\}$ *and* $\beta_{b,k}(.|\mathbf{x},\mathbf{z})$, $k \in \{L, H\}, b \in \{1, \ldots, B\}$ *are optimal given the number of bidders* N_b *for broker* b *and* $\beta_{S,k}(c|\mathbf{x},\mathbf{z})$, $k \in \{L, H\}$ *is optimal given the distribution of price inquiries resulting from* $\kappa_1, \ldots, \kappa_M$ *and* $\hat{\kappa}$.

5. IDENTIFICATION

This section discusses the identification of the primitives of the model, which are $\mathcal{H}(.|\mathbf{x})$ and $\mathcal{G}_k(.|\mathbf{z})$, $k \in \{L, H\}$ from the observables, which are the contract prices p, which carter serves a contract and whether a contract was arranged through a broker or not. Additional available information are the conditioning variables $\mathbf{z}$ and $\mathbf{x}$. Lastly, for a broker b, we can see the number of bidders N_b that the broker procures to as well as the commission charges $\phi(\mathbf{x},\mathbf{z})$.

Let a model $(\mathcal{S}, \Gamma)$ be a pair of a set $\mathcal{S}$ of distribution functions over unobserved variables and Γ be a set of models that map the set $\mathcal{S}$ into the set of distribution functions over observed random variables $\mathcal{K}$.

Definition 2. *A model* $(\mathcal{S}, \Gamma)$ *is identified if and only if for every* $(S, \hat{S})$, $\gamma(S) = \hat{\gamma}(S)$ *implies* $(S, \gamma) = (\hat{S}, \hat{\gamma})$.[42]

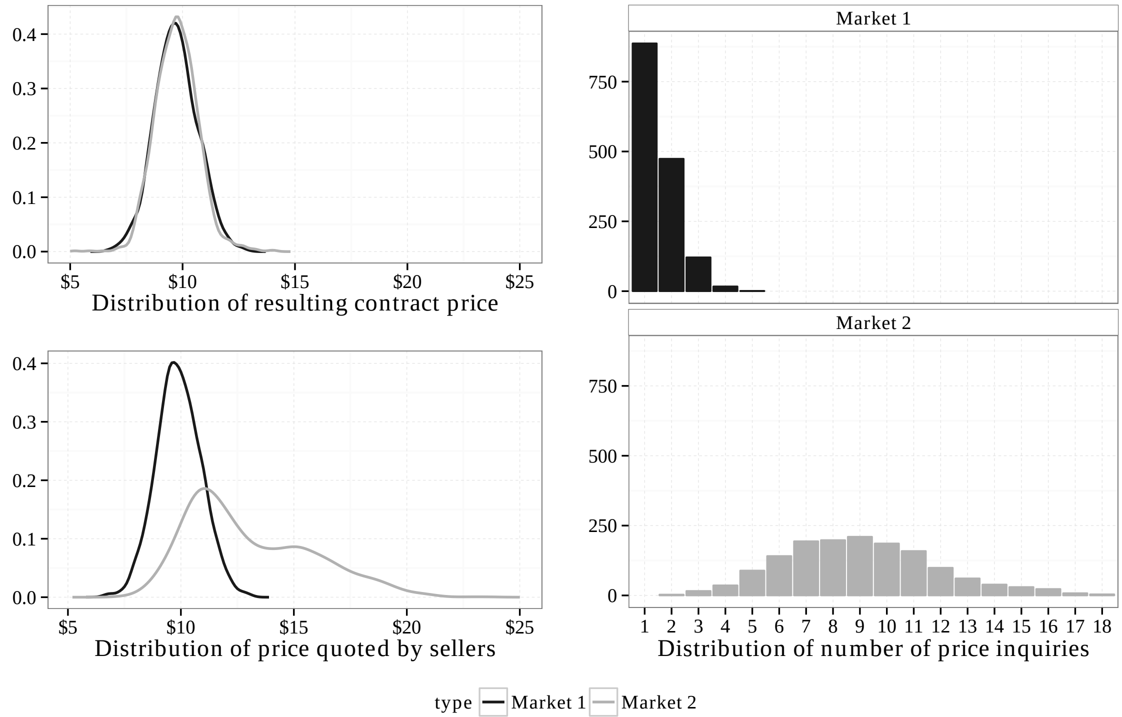
Figure 6 illustrates why prices in the search-market alone would not be enough to tell search cost and carters' service cost apart. It shows simulated data from two markets. The distribution of contract prices, shown in the top left corner, are the same, and the two markets are therefore observationally equivalent. The underlying primitives that give rise to these distributions are, however, very different. In market 1 buyers have higher search cost and ask for fewer price quotes relative to market 2. Market 1, on the other hand, is served by firms that have lower cost and quote lower prices. This example shows that it will not be enough to observe just one price distribution to tell apart the search cost for buyers and the cost of service provision for carters. For the evaluation of policy scenarios it will, however, be important to be able to distinguish between these two scenarios.

It is therefore crucial that the data records both the contracts arranged in the brokered market and those in the search market. The fact that brokers procure through auction process along with a proxy for the number of bidders in these auctions allows me to pin down carters' cost from this portion of the market. Known cost functions for carters can then be used with the remaining bilaterally negotiated contracts to identify the search cost distribution.

One confound to a direct application of this identification strategy is that only a selected set of carters bids in the broker market. Results in the descriptive section show that those carters are also

[42] See Athey and Haile (2002) or similarly Matzkin (2007).

Figure 6: High cost market with low search cost and low cost market with high search cost.



Notes: The four panels show data from two *simulated* markets. The top left panel is the contract price distribution observed by the econometrician and the bottom left panel the price offer distributions. The two right panels show the distribution of price inquiries for the two markets.

systematically cheaper in the search market. In the model, I will account for this selection by adjusting for the number of firms of type L and H in both markets.[43] Since the cost distributions are recovered from the brokered market, I need to observe at least one of each type of firm in this part of the market. In addition to the setup of the game I make the following assumptions:

Assumption 1. *At least one firm of each type $k \in \{L, H\}$ is present in the brokered market.*

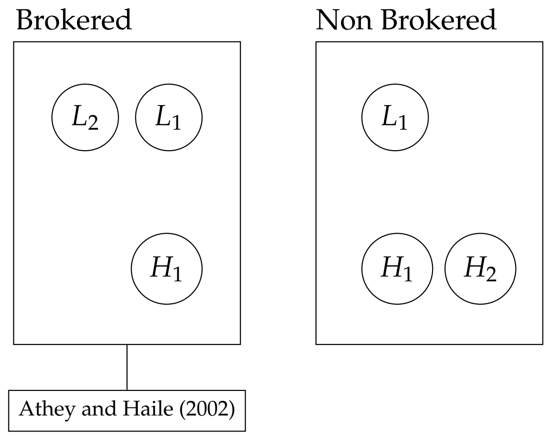
Assumption 2. *A decentralized market equilibrium exists and is unique in the data.*

I argue that under these conditions the distribution functions of carter cost $\mathcal{G}_k(.|\mathbf{z}), k \in \{L, H\}$ are identified and the search-expense distribution $\mathcal{H}(.|\mathbf{x})$ partially identified with known probabilities between the marginal types $\mathcal{H}(\kappa_M(\mathbf{x}, \mathbf{z})|\mathbf{x}, \kappa < \bar{\kappa})$,..., $\mathcal{H}(\kappa_{m-1}(\mathbf{x}, \mathbf{z})|\mathbf{x}, \kappa < \bar{\kappa}) - \mathcal{H}(\kappa_m(\mathbf{x}, \mathbf{z})|\mathbf{x}, \kappa < \bar{\kappa})$,..., $\mathcal{H}(\kappa_{\hat{m}}(\mathbf{x}, \mathbf{z})|\mathbf{x}, \kappa < \bar{\kappa}) - \mathcal{H}(\bar{\kappa}(\mathbf{x}, \mathbf{z})|\mathbf{x}, \kappa < \bar{\kappa})$ and $1 - \mathcal{H}(\bar{\kappa}(\mathbf{x}, \mathbf{z})|\mathbf{x})$.

Carter cost are identified due to **Theorem 6** in Athey and Haile (2002), which states that the distributions of valuations in asymmetric auctions can be identified from the transaction price and the identity of the winner, which is an application of the identification result in competing risk models presented in Meilijson (1981). We can therefore treat $\mathcal{G}_k(.|\mathbf{z}), k \in \{L, H\}$ as known objects for all firms that bid

[43] I am not modeling how the links between carters and brokers are formed, although this would certainly be an interesting question. One could, for example, allow for a stage before the auction where brokers and carters bargain for the slots in the auction.

Figure 7: Selection in The Market



Notes: This figure provides a graphical illustration of the type of selection that is allowed on the seller side. Firms L and H are both observed in the brokered markets. Due to arguments presented by Athey and Haile (2002) the cost are identified from observed contract prices in the brokered market. This in turn allows us to assign the already known cost functions from the broker market to firms of type L and H in the search market.

in the brokered market. Because we know the identity of the firms and due to the support assumption **A1** this at the same time identifies the cost function of firms that do not bid in the brokered market, which is illustrated in Figure 7. The optimal bidding functions in the search market $\beta_{S,L}(c|\mathbf{x},\mathbf{z},w(\mathbf{x},\mathbf{z}))$, $\beta_{S,H}(c|\mathbf{x},\mathbf{z},w(\mathbf{x},\mathbf{z}))$ are known up to the finite dimensional parameter vector $w(\mathbf{x},\mathbf{z})$. They can be inverted due to monotonicity. The expression for the observed equilibrium price distribution in the search market can therefore be written as:

$$\mathcal{F}^O(p|\mathbf{x},\mathbf{z}) = \sum_{m=1}^{M} w_m(\mathbf{x},\mathbf{z}) \sum_{k=0}^{m} \frac{\binom{N_L}{k}\cdot\binom{N_H}{m-k}}{\binom{N_H+N_L}{m}} \cdot \\ \left(1-\widetilde{\mathcal{G}}_L(\beta_{S,L}^{-1}(p|\mathbf{x},\mathbf{z},w(\mathbf{x},\mathbf{z}))|\mathbf{z})^k \cdot \widetilde{\mathcal{G}}_H(\beta_{S,H}^{-1}(p|\mathbf{x},\mathbf{z},w(\mathbf{x},\mathbf{z}))|\mathbf{z})^{m-k}\right). \quad (4)$$

For every observed p this functional relates the known value $\mathcal{F}^O(p|\mathbf{x},\mathbf{z})$ on the LHS to the RHS, which is known up to the M-dimensional vector $w(\mathbf{x},\mathbf{z})$. The search-weight vector is therefore over-identified from the above equation since it holds for all p in the continuous set of prices. The estimation section explains a minimum distance estimator based on Equation 4 and monte carlo results, which are shown in Appendix C show that one can recover $w(\mathbf{x},\mathbf{z})$ well using this estimator.

Once the weight vector $w(\mathbf{x},\mathbf{z})$ is known the unobserved distribution of price offers is also known as

$$\mathcal{F}(p|\mathbf{x},\mathbf{z},m) = \sum_{k=0}^{m} \frac{\binom{N_L}{k}\cdot\binom{N_H}{m-k}}{\binom{N_H+N_L}{m}} \cdot \left(1-\widetilde{\mathcal{G}}_L(\beta_{S,L}^{-1}(p|\mathbf{x},\mathbf{z},w(\mathbf{x},\mathbf{z}))|\mathbf{z})^k \cdot \widetilde{\mathcal{G}}_H(\beta_{S,H}^{-1}(p|\mathbf{x},\mathbf{z},w(\mathbf{x},\mathbf{z}))|\mathbf{z})^{m-k}\right).$$

Knowing the CDF one can compute the PDFs and from this compute the marginal types of buyers in the search market:

$$\kappa_m(\mathbf{x},\mathbf{z}) = \int_0^{\bar{p}} p \cdot f(p|\mathbf{x},\mathbf{z},m)dp - \int_0^{\bar{p}} p \cdot f(p|\mathbf{x},\mathbf{z},m+1)dp \ \forall m.$$

Lastly, $\mathbb{E}[p^B|\mathbf{x},\mathbf{z}]$ and $\phi(\mathbf{x},\mathbf{z})$ are observed in the data, which determines the cut-off search cost type who is indifferent between searching and delegating to a broker. Let $\hat{m}$ denote the number of searches of the cut-off type.

$$\bar{\kappa}(\mathbf{x},\mathbf{z}) = q \cdot (\mathbb{E}[p^B|\mathbf{x},\mathbf{z}] \cdot \phi(\mathbf{x},\mathbf{z}) - \int_0^{\bar{p}} \hat{m} \cdot pf(p|\mathbf{x},\mathbf{z},m)dp)/\hat{m}.$$

The weights w already provide significant information about the search-cost distribution. In general they will suffice to fit a parametric function for the distribution of search-expenses. To allow more flexibility in the estimation of $\mathcal{H}(.|\mathbf{x})$ one needs a variable that induces variation in the cut-off types in $\kappa_1, \ldots, \kappa_M, \kappa_B$ while leaving the distribution of search cost unaffected.

Assumption 3. *There exists a variable $z \in \mathbf{z}$ so that $\mathcal{H}(.|\mathbf{x},z) = \mathcal{H}(.|\mathbf{x})$.*

The variation in z will be used to obtain more than one set of cut-offs $\kappa_1, \ldots, \kappa_M, \kappa_B$ and allow me to estimate $\mathcal{H}(.|\mathbf{x})$ flexibly without imposing a parametric assumption.

6. ESTIMATION

In the first step of the estimation the cost distribution of carters is recovered from the brokered market where the underlying distribution of price offers is inferred from the distribution of winning bids as well as the number of bidders in the auction. These semi-parametric estimates of the bid distribution are then mapped into cost draws using the asymmetric first order conditions of carters. Once the cost functions are known, they are used in a moment based procedure to back out the values of $w(\mathbf{x},\mathbf{z})$ for different values of $\mathbf{x}$ and $\mathbf{z}$. The last step consists of a semi-parametric estimation of the function $\mathcal{H}(.|\mathbf{x})$, matching the vector $w(\mathbf{x},\mathbf{z})$. Part of the challenge of the estimation of the search cost is that the equilibrium bid distributions, which depend on the equilibrium search weights $w(\mathbf{x},\mathbf{z})$, have to be computed repeatedly for each new guess of $w(\mathbf{x},\mathbf{z})$ and different values of $\mathbf{x}$ and $\mathbf{z}$. The estimation of the unobserved equilibrium search weights therefore involves a nested fixed point computation in which the inner loop consists of a procedure to solve for the bidding functions, as suggested in Bajari (2001). Before going into the full details of the estimation, the following list provides a summary:

1. Estimate (unobserved) semi-parametric price offer distribution from observed contracts.
2. Use estimates from previous step to simulate cost draws via the first order conditions, shown in Appendix D.
3. Use simulated cost distributions from previous step to obtain a kernel estimate of the densities.
4. Use the cost estimates along with the price distribution in the brokered market to estimate buyers equilibium behavior w_m, $m \in \{1, \ldots, M\}$ in a minimum distance procedure.
    - For each new guess, recompute bidding strategies $\beta_{S,H}(.|w)$ and $\beta_{S,L}(.|w)$.

5. Obtain a semi-parametric estimate of $\mathcal{H}(.|\mathbf{x})$ by matching the probabilities implied by w_m, $m \in \{1, \ldots, M\}$.

   - Use exclusion restriction.

6.1 Estimation of the Carter Cost Distribution

In the first step of the estimation I obtain the cost distribution of carters using the subset of contracts awarded through brokers. To accommodate the remaining auction level heterogeneity on the seller side I make the assumption that the cost of carters is additive with the linear index of customer-specific observables $\mathbf{z}_j$ as well as the unobservable c_{ij}.[44] Note that $\mathbf{z}_j$ can be interpreted as auction-specific observable heterogeneity.

$$C_{ij}(\mathbf{z}_j, c_{ij}) = c_{ij} + \gamma \cdot \mathbf{z}_j \tag{5}$$

The advantage of this assumption is that the optimal price offer in a procurement auction when $\gamma \cdot \mathbf{z}_j \neq 0$ is just $\gamma \cdot \mathbf{z}_j + p_{ij}$, where p_{ij} is the price that would be offered if $\gamma \cdot \mathbf{z}_j = 0$. This has been pointed out in Haile et al. (2003). This means that we can obtain γ as well as a residual price $\tilde{p}_{ij}$ through a hedonic OLS regression. We can then perform the structural estimation to obtain the privately observed part of the cost c_{ij} with the residual price and later obtain rescaled final prices for the conditional cost case by adding $\gamma \cdot \mathbf{z}_j$ back in.[45]

After this initial regression step I obtain a semi-nonparametric estimate of the residual price distribution based on the observed residual contract prices $\tilde{p}_{ij}$, where I follow a modified version of the approach proposed in Gallant and Nychka (1987):

$$\hat{f}^{GN}(v_{ij}; \omega) = \Big[\sum_{k=1}^{K} \omega_k \cdot T_k(v_{ij})\Big]^2 + \epsilon_0 \cdot \phi(v_{ij}), \quad \sum_{k=1}^{K} \omega_k^2 + \epsilon_0 = 1$$

where $T_k(.)$ are the normalized hermite polynomials of degree k, all of which integrate to one and where K determines the smoothness.[46] This procedure can be numerically unstable when the support is not restricted. I therefore use a version of this estimator with a support truncated to $[\underline{x}, \overline{x}]$, which has been suggested in il Kim and Lee (2014):

$$\hat{f}^{TGN}(v_{ij}; \omega) = \frac{\hat{f}^{GN}(v_{ij}; \omega)}{\int_{\underline{v}}^{\overline{v}} \hat{f}^{GN}(v_{ij}; \omega) dv}, \quad v_{ij} = \frac{p_{ij} - \mu}{\sigma}$$

Since I assume two types L and H, I have to estimate a separate price offer functions for both types of firms. The set of parameters in this step is therefore $\theta_L = (\omega_L, \sigma_L, \mu_L)$ and $\theta_H = (\omega_H, \sigma_H, \mu_H)$.

[44] A similar assumption is for example made in Asker (2010).

[45] In many circumstances it might be important to control for auction level heterogeneity that is unobservable to the researcher. Krasnokutskaya (2011) shows how to deal with such sources of heterogeneity. The methods are based on deconvolution results in the statistical literature, which can be found in Li and Vuong (1998).

[46] Hermite polynomials are recursively defined as: $T_1(x) = ((\sqrt{2\pi})^{(-1/2)}) \cdot \exp((-x^2)/4)$ for $k = 1$, $T_2(x) = x \cdot T_1(x)$ and $T_k(x) = ((x \cdot T_{(k-1)}) - \sqrt{k-2} \cdot T_{(k-2)})/\sqrt{k-1}$. I use $K = 5$.

Putting these things together a log-likelihood contribution of a contract with winner of type k and N_k as well as N_{-k} bidders is:[47]

$$l_k(p_i;\theta_k,\theta_{-k}) = N_k \cdot \left[1 - F_k^{TGN}(p_i;\theta_k)\right]^{(N_k-1)} \cdot \left[1 - F_{-k}^{TGN}(p_i;\theta_{-k})\right]^{N_{-k}} \cdot f_k^{TGN}(p_i;\theta_k)$$

The estimated bid distributions are used to generate 15,000 pseudo cost draws according to the relationship in Equation 17 in Appendix D, which expresses the cost draws in terms of the observable price distribution based on the first order condition of firms. I then use a kernel estimator to obtain PDF values from this distribution with a normal kernel and a bandwidth based on silverman's rule (Silverman (1986)). The CDF of the cost function is obtained by gauss-hermite integration of the estimated kernel based on 45 nodes. For the estimation of the search cost, I have to be able to evaluate the estimated cost functions quickly and a kernel will in general be too slow. I therefore create a grid both for the PDF of the cost function as well as the CDF, where I use 320 grid points. This allows for a quick evaluation of both functions without the need to perform the kernel evaluation or numerical integration repeatedly.

6.2 Estimation of the Search Expense Distribution

The estimation of the search cost proceeds in two steps. First, the equilibrium proportions of buyers that make m searches are recovered from the data (w_m, $m \in \{0,\ldots,M\}$). With these in hand, one can compute the cut-off types and therefore match a probability distribution of search cost.

I employ a minimum distance estimator. The main challenge in this step of the estimation is the fact that, unlike in the symmetric case, there is no closed form solution for the potentially asymmetric bidding functions $\beta_{S,L}(c|\mathbf{x},\mathbf{z})$ and $\beta_{S,H}(c|\mathbf{x},\mathbf{z})$ and we need to repeatedly evaluate the inverse bidding function of carters to form the objective function value. Here I instead use collocation methods following Bajari (2001), the idea of which is to approximate the inverse bidding function with a high order polynomial and choose the parameters of the polynomial to minimize the deviations from the conditions implied by the FOC.[48] The first order conditions, based on the profit maximization problem described in Equation 3, are shown below. The dependence on observables is suppressed. For carters of type H:

[47] The probability that a firm of type k wins under N_k as well as N_{-k} bidder and the price is lower than p is given by $P(p,k\ wins) = \int_0^p N_k \cdot (1-F_k(u))^{(N_k-1)} \cdot (1-F_{-k}(u))^{N_{-k}} \cdot f_k(u)du$. Taking the derivative and applying Leibnitz's rule the density is given by: $f_k^O(p) = N_k \cdot (1-F_k(p))^{(N_k-1)} \cdot (1-F_{-k}(p))^{N_{-k}} \cdot f_k(p)$.

[48] What complicates the computation of such equilibria is the fact that one of the boundary conditions is not known. In this case of a procurement auction, this is the lowest support point of the price distribution. The literature has established several procedures to solve such asymmetric auctions. An alternative approach are backwards-shooting algorithms (Marshall et al. (1994), Bajari (2001), Li and Riley (2007)), which guess the lower-support point and use it to solve the differential equations given this guess. This procedure is repeated until the upper known boundary conditions (the highest cost type bids his cost) are met with some tolerance. This approach has the downside that is is computationally relatively slow and can be quite unstable.

$$\sum_{m=1}^{M-1} w_m \cdot \sum_{k=0}^{m} \frac{\binom{N_L}{k} \cdot \binom{N_H-1}{m-k}}{\binom{N_H+N_L-1}{m}} \cdot \Bigg[\widetilde{\mathcal{G}}_L(\beta_{S,L}^{-1}(p))^k \cdot \widetilde{\mathcal{G}}_H(\beta_{S,H}^{-1}(p))^{m-k} \left(1 - \frac{(p_i - c) \cdot k \cdot g_L(\beta_{S,L}^{-1}(p)))}{\beta'_{S,L}(\beta_{S,L}^{-1}(p)) \cdot \widetilde{\mathcal{G}}_L(\beta_{S,L}^{-1}(p))} - \frac{(p_i - c) \cdot (m-k) \cdot g_H(\beta_{S,H}^{-1}(p)))}{\beta'_{S,H}(\beta_{S,H}^{-1}(p)) \cdot \widetilde{\mathcal{G}}_H(\beta_{S,H}^{-1}(p))}\right) \Bigg] = 0 \quad (6)$$

and carters of type L:

$$\sum_{m=1}^{M-1} w_m \cdot \sum_{k=0}^{m} \frac{\binom{N_H}{k} \cdot \binom{N_L-1}{m-k}}{\binom{N_L+N_H-1}{m}} \cdot \Bigg[\widetilde{\mathcal{G}}_H(\beta_{S,H}^{-1}(p))^k \cdot \widetilde{\mathcal{G}}_L(\beta_{S,L}^{-1}(p))^{m-k} \left(1 - \frac{(p_i - c) \cdot k \cdot g_H(\beta_{S,H}^{-1}(p)))}{\beta'_{S,H}(\beta_{S,H}^{-1}(p)) \cdot \widetilde{\mathcal{G}}_H(\beta_{S,H}^{-1}(p))} - \frac{(p_i - c) \cdot (m-k) \cdot g_L(\beta_{S,L}^{-1}(p)))}{\beta'_{S,L}(\beta_{S,L}^{-1}(p)) \cdot \widetilde{\mathcal{G}}_L(\beta_{S,L}^{-1}(p))}\right) \Bigg] = 0 \quad (7)$$

The idea of solving for an equilibrium via collocation methods is to approximate $\beta_{S,L}^{-1}(.)$ and $\beta_{S,H}^{-1}(.)$ with two functions $\gamma(.;\boldsymbol{\alpha}_L, \underline{p})$, $\gamma(.;\boldsymbol{\alpha}_H, \underline{p})$ so that the above equations are satisfied as closely as possible over a grid of prices. The variables of this minimization problem are $\boldsymbol{\alpha}_L$, $\boldsymbol{\alpha}_H$ and $\underline{p}$, where the $\boldsymbol{\alpha}$ are the coefficients of the polynomial and $\underline{p}$ the lowest price charged in equilibrium. The grid over which the approximation is computed is therefore endogenous and range from $\underline{p}$ to $\bar{c}$, where the latter is upper support point of the cost distributions. Instead of a uniform grid as suggested by Bajari (2001) I found that an unequally spaced grid that allows for more points at the lower end of the bid function (where most of the density of cost is placed) is more stable and therefore compute $p_k = \underline{p} + (((k-1) \cdot (\bar{c} - \underline{p}))^2)/((n_k - 1)^2)$ for n_k grid points.

Similar to Bajari (2001) denote $\Psi(p_k^L, \boldsymbol{\alpha}_L, \underline{p})$ and $\Psi(p_k^H, \boldsymbol{\alpha}_H, \underline{p})$ the quadratic deviations from Equation 6 and Equation 7. The FOC do not imply that the boundary conditions hold. Bajari (2001) therefore suggests to solve for an equilibrium by using standard polynomials and solving the following:

$$\min_{\alpha_L, \alpha_H, \underline{p}} \sum_k \Psi(p_k^L, \boldsymbol{\alpha}_L, \underline{p}) + \sum_k \Psi(p_k^H, \boldsymbol{\alpha}_H, \underline{p}) + (\gamma(\bar{c}; \boldsymbol{\alpha_L}, \underline{p}) - \bar{c})^2 + (\gamma(\bar{c}; \boldsymbol{\alpha_H}, \underline{p}) - \bar{c})^2 + (\gamma(\underline{c}; \boldsymbol{\alpha_L}, \underline{p}) - \underline{p})^2 + (\gamma(\underline{c}; \boldsymbol{\alpha_H}, \underline{p}) - \underline{p})^2$$

Instead of standard polynomials, I rely on Bernstein polynomials, which allow me to impose the boundary conditions as constraints that have to be satisfied exactly and also provide an easy way to impose monotonicity on the solution. The terms of a Bernstein polynomial of degree n are given by:

$$t_{j,n}(x) = \binom{n}{j} \cdot x^j \cdot (1-x)^{n-j} \quad j \in \{0, \ldots, n\} \quad (8)$$

After normalizing the price to accommodate the fact that these polynomials are defined for values between zero and one the approximations of the inverse bid functions for a firm of type k is then given by:

$$\gamma(p;\boldsymbol{\alpha}_k,\underline{p})=\sum_{j=0}^{n}\alpha_j^k\cdot t_{j,n}\Big(\frac{p-\underline{p}}{\underline{c}-\underline{p}}\Big) \tag{9}$$

Practically, I choose a $n = 9$ for a tenth order polynomial and set $\alpha_0^k = \underline{p}$ for the lower boundary and $\alpha_n^k = \bar{c}$ for the upper boundary as well as $\alpha_0^k \leq \ldots \leq \alpha_n^k$ for monotonicity.[49]

This strategy to solve for the bidding function is used as the inner loop of a minimum distance procedure to estimate the search cost of buyers. I target the first and second moment of the price distribution as well as J quantiles corresponding to the percentiles from 0.15 to 0.85 in increments of 0.025:

$$dev_1(\mathbf{x},\mathbf{z};w)=\begin{bmatrix}\hat{E}(p|\mathbf{x},\mathbf{z})-E(p|\mathbf{x},\mathbf{z};w)\\ \hat{V}(p|\mathbf{x},\mathbf{z})-V(p|\mathbf{x},\mathbf{z};w)\\ \hat{Q}c_1(p|\mathbf{x},\mathbf{z})-Qc_1(p|\mathbf{x},\mathbf{z};w)\\ \vdots\\ \hat{Q}c_J(p|\mathbf{x},\mathbf{z})-Qc_J(p|\mathbf{x},\mathbf{z};w)\end{bmatrix} \tag{10}$$

and minimize:

$$\min_{w} dev_1(\mathbf{x},\mathbf{z};w)^{\top}\cdot W(\mathbf{x},\mathbf{z})\cdot dev_1(\mathbf{x},\mathbf{z};w)$$

for each $(\mathbf{x},\mathbf{z})\in\mathbf{z}\times\mathbf{x}$, where $W(\mathbf{x},\mathbf{z})$ is a diagonal matrix with the inverse of the sample standard deviation of the moments.[50] Remember that the w estimated in this step are coming from a truncated distribution $\mathcal{H}(.|\kappa<\bar{\kappa},\mathbf{x})$, the selected subset of buyers that search in the broker market. For a given guess of w one can compute the cut-off type and therefore the number of searches that will never be observed in the search-market (see Figure 5). This means that these weights should be equal to zero. To maintain the consistency with the next step in the estimation, where I estimate the full $\mathcal{H}(.|\mathbf{x})$ I already impose that all w cells below the broker cut-off type are zero. This means that for a given w I compute the bidding strategies of carters and then compute $\bar{\kappa}$ and set all weights w_m below $\bar{\kappa}$ to zero and re-normalize w so it adds up to one. In the objective function I then include a penalty if the original guess w deviates from the updated w.

The last step of the estimation procedure is to estimate a semi-nonparametric continuous distribution of the search cost $\mathcal{H}(.|\mathbf{x})$ that fits the w, which we backed out in the previous step. In this step the exclusion restriction is used and the search-expense distribution held fixed why I vary $z_j\in\mathbf{z}$. The

[49]I allow for 40 grid points. I have found that a smaller number of grid points tends to increase the average bid for a given cost, which means that less collocation nodes overestimate the predicted price.

[50]The diagonal elements of the weighting matrix are computed as $(\hat{\sigma}^2/n)^{(-1/2)}$ for the mean and $\hat{\mu}_4/n-(\hat{\sigma}^4\,(n-3))/(n\cdot(n-1)))^{(-1/2)}$ for the variance, where $\mu_4=\mathbb{E}[(X-\mu)^4]$. The standard deviation of the quantile c is computed as $(c\cdot(1-c))/(n\cdot K_p(p_c))^{(-1/2)}$, where $K_p(.)$ is the estimated kernel of the price distribution. The expectation $\mathbb{E}[p|\mathbf{x},\mathbf{z}]$ is a weighted average over the lowest order statistics of prices given m calls: $\mathbb{E}[p|\mathbf{x},\mathbf{z}]=\sum_m w(\mathbf{x},\mathbf{z})\cdot\mathbb{E}[p^m|\mathbf{x},\mathbf{z}]=\sum_m w(\mathbf{x},\mathbf{z})\cdot\int_0^{\bar{p}} m\cdot p\cdot q_j\cdot(1-\mathcal{F}(p|\mathbf{x},\mathbf{z}))^{m-1}f(p|\mathbf{x},\mathbf{z})dp$. The variance according to the variance decomposition formula: $Var[p|\mathbf{x},\mathbf{z}]=\sum_m w(\mathbf{x},\mathbf{z})\cdot(\mathbb{E}[p^m|\mathbf{x},\mathbf{z}]-\mathbb{E}[p|\mathbf{x},\mathbf{z}])^2+\sum_m w(\mathbf{x},\mathbf{z})\cdot Var[p^m|\mathbf{x},\mathbf{z}]$.

target both the weights that were backed out in the previous step and in addition average number of searches that were made:

$$dev_{2j}(\mathbf{x}, z_j; \theta) = \begin{bmatrix} w_M(\mathbf{x}, z_j) - \hat{\mathcal{H}}(\kappa_M(\mathbf{x}, z_j) | \kappa < \bar{\kappa}; \mathbf{x}, \theta) \\ \vdots \\ w_m(\mathbf{x}, z_j) - \left(\hat{\mathcal{H}}(\kappa_{m-1}(\mathbf{x}, z_j) | \kappa < \bar{\kappa}; \mathbf{x}, \theta) - \hat{\mathcal{H}}(\kappa_m(\mathbf{x}, z_j) | \kappa < \bar{\kappa}; \mathbf{x}, \theta) \right) \\ \vdots \\ w_b(\mathbf{x}, z_j) - \left(1 - \hat{\mathcal{H}}(\bar{\kappa}(\mathbf{x}, z_j) | \mathbf{x}; \theta) \right) \\ \sum_{m=1}^{M} m \cdot \left(w_m - \hat{\mathcal{H}}(\kappa_{m-1}(\mathbf{x}, z_j) | \kappa < \bar{\kappa}; \mathbf{x}, \theta) - \hat{\mathcal{H}}(\kappa_m(\mathbf{x}, z_j) | \kappa < \bar{\kappa}; \mathbf{x}, \theta) \right) \end{bmatrix} \tag{11}$$

Now for a given $\mathbf{x}$ and J different values of $\mathbf{z}$ the following criterion function is minimized:[51]

$$\min_{\theta} \begin{bmatrix} dev_{21}(\mathbf{x}, z_1; \theta) \\ \dots \\ dev_{2J}(\mathbf{x}, z_J; \theta) \end{bmatrix}^{\top} \cdot W(\mathbf{x}, \mathbf{z}) \cdot \begin{bmatrix} dev_{21}(\mathbf{x}, z_1; \theta) \\ \dots \\ dev_{2J}(\mathbf{x}, z_J; \theta) \end{bmatrix}$$

6.3 Practical considerations for Estimation

6.3.1 Contract Dynamics

The relationship between a carter and a business may persist over a longer time horizon, and price adjustments, while infrequent, may occur during this relationship. In this project I use only the initial price in the relationship between customers and carters. While I do see the length of the relationship between customers and carters, the data provides no information on the dates at which the contract terms have been renewed. For the estimate of the search cost it will be important to correctly specify the length of the time period over which buyers are bargaining. Within this static modeling approach there are two reasonable ways to specify the length of the contract. The first one is to base it on a regulations by the BIC, which mandates that customers must be able to switch after two years. Another way would be to compute from the data how long a contract lasts in expectation.[52] For the latter "lasts" is defined as the average time until one of the three events occurs: the customer obtains a new rate, goes out of business or switches. This average is at 2.18 and therefore close to the regulated maximum length of two years. Given that these numbers are close I assume that buyers expect to pay the initial rate for two years and omit any continuation value that might be derived from this initial price.[53]

[51] I am currently using the identity matrix as a weighting matrix in this step. Ones a boot-strapped distribution of the w is available I will replace diagonal elements by the inverses of the standard errors of the w's.

[52] http://www.nyc.gov/html/bic/html/trade_waste/customer_info_contracts.shtml (last accessed on 08/19/15)

[53] This might for example be violated if buyers can always go back to the initial price and treat it as a free draw from the price distribution. Note that $\kappa_m = (q \cdot (\mathbb{E}[p^{(m)}] - \mathbb{E}[p^{(m+1)}]))/\delta$, where δ is the discount factor. A change in th length of the contract will therefore merely scale the search cost distribution by a multiplicative factor. To the extend that customers bargain in fact for a longer time horizon because of hysteresis in contracts my approach leads to an under-estimate of the search-cost. Appendix C provides more details on the price development of contracts within a customer-carter relationship and shows that prices are relatively stable for two years and then trend slightly upwards, which provides some evidence that buyers can not treat the initial price as a reservation price for their next search.

6.3.2 Number of Bidders in Broker Market

The number of bidders on a broker contract is set to the number of carters that have won a contract through this broker in the borough of the observed contract. A concern of this assumption is that this systematically underestimates the number of bidders. However, an auction participant that does not show up in the data is one that never won a contract and is therefore likely an uncompetitive firm that should have indiscernible effects on the bidding strategies of other firms and therefore the estimates.

6.3.3 Accounting for Observed Variation

The second step of the estimation is time consuming because of the nested procedure and needs to be performed separately for a given combination of $\mathbf{x}$ and $\mathbf{z}$. Observable that are explicitly handled in the model therefore need to be restricted to a small set. The quantity of waste generated by a customer directly affects the marginal benefit of searching for a customer and therefore needs to be included in the model. In addition to that I need a variable $z \in \mathbf{z}$ that changes a carter's cost to service a customer while leaving a customer's ability to search unaffected.[54] A good choice is the amount of recyclables generated by the customer. One of the most common recyclables is old corrugated cardboard (OCC). Cardboard is included in the waste stream of about 70% of all customers that generate recyclables.[55] OCC has an average price between \$90 and \$200 per ton and is an important source of income for carters according to the following statement of a CEO of one of the New York companies, which is referring to the common theft of cardboard from the curbside of the customer location:[56]

"The daily theft of cardboard hurts our entire industry, from small family-owned hauling operators to larger firms who lose the revenue, and their customers that feel the loss in greater fees."[57]

In summary, I allow both the customer search-expense distribution and the carter service-cost distribution to depend on the quantity of waste generated by the customer. In addition to that, carter cost depend on the recyclables generated by a customer, while being excluded from customers search cost. Contracts are sorted into one of four groups $\{Q_{0,25}(q), Q_{25,50}(q), Q_{50,75}(q), Q_{75,100}(q)\}$, where $Q_{a,b}(q)$ are quantities in the range of the a'th and b'th quantile. Regarding recyclables, I allow for three levels: *No Recyclables* when customers have No Recyclables, *Low Recyclables* when recyclables make up less than 10% and *High Recyclables* above that.

To deal with observable variation that is not explicitly considered in the model I homogenize prices along several observable dimensions of a contract. In analogy to a price deflator I first run a regression of the log-price on a set of dummy variables: the type of business, the reporting period, the transfer station as well as the number of pick-ups per week. Based on the results of this regression I compute predicted values $\hat{y}_i$ and re-base the price, dividing through the exponential of the predicted value and multiplying by the exponential of the estimated intercept: $p_i \cdot exp(\hat{b}_0)/exp(\hat{y}_i)$. The observation is now interpreted as belonging to the joint set of observations defined by the excluded variables in each set of dummy variables. The choice on these excluded variables therefore determines to what category prices

[54] Note that customers *incentives* to search are still affected by the amount of recyclables to the extend that it changes the price in the market.

[55] Other recyclables are much less common, 5.7% of all customers generate glass, 6.1% plastic, 0.7% aluminum and 1.6% metal.

[56] Source: http://www.alibaba.com/showroom/occ-scrap.html (accessed on 08/03/2015)

[57] Source: http://www.recyclingtoday.com/cardboard-theft-new-york.aspx (accessed on 08/03/2015)

are re-normalized. I choose the following: the most recent year (2014), the most common business type category (retail non-food), the average number of picksups (five) and the largest transfer station. To give a sense of the variation in the final dataset constructed in this way, specification one and two in Figure 8 show regressions of the homogenized prices on the dummies for these different tiers of quantities, separately for brokered and search-market contracts.

6.3.4 Grouping Sellers and Spatial Scale Economies

The model introduced firms of type L and H in order to control for the selection of carters across the two markets. The identification argument, which builds on Athey and Haile (2002), requires that firms can be classified according to some observable variable. We therefore need an observable that can account for the persistent differences in prices. In the spirit of a "fixed effect" in a linear model, firms are classified according to their average price. For each carter I compute the fixed effect on the price and then call all carters with a fixed effect above the mean of fixed effects H and all remaining carters L. While this reduces differences in carters prices to a binary variable, controlling for it fully purges price differences between carters that deal with brokers and those that do not as can be seen in Figure 8 in specification six.

However, ideally one would want to classify carters according to a "more exogenous" criterion. I have explored two alternative variables that could potentially capture scale economies in servicing customers along established routes. These variables, however, do not explain variation in prices very well and unlike the aforementioned indicator can not purge price differences across the two markets in the reduced form. The first variable is simply firm size, based on the number of serviced customers. For the second one I collected, for a given carter-customer relationship, the marginal driving time from the zip-code centroid of a customer to the zip-code centroid of the next-closest customer.[58] I then average these over the entire set of customers within a given carter. The latter provides a measure of the density of the network. Both variables are based on the idea that more or more densely populated customers reduce the overall service cost. Specification three shows that the inclusion of the distance variable has a positive sign, which is unexpected. Moreover, the dummy variable that decodes whether a carter deals with a broker is still negative and highly significant. The same pattern emerges in specification four, which includes the route density measure.

[58]The data was collected via the Google distance matrix API. Exact travel time data as opposed to Euclidean distances have also been used in the labor matching literature, for example Harmon (2013).

Figure 8: Regressions accounting for differences across carter

Variable	(1) Price	(2) Price	(3) Price	(4) Price	(5) Price	(6) Price
$Q_{25,50}(q)$	-1.817**	-1.896**	-1.733**	-1.742**	-1.737**	-1.594**
	(0.107)	(0.0856)	(0.104)	(0.109)	(0.108)	(0.102)
$Q_{50,75}(q)$	-2.521**	-2.616**	-2.429**	-2.437**	-2.433**	-2.363**
	(0.0998)	(0.0844)	(0.0979)	(0.101)	(0.100)	(0.0967)
$Q_{75,100}(q)$	-3.800**	-3.331**	-3.658**	-3.664**	-3.660**	-3.522**
	(0.135)	(0.0836)	(0.136)	(0.139)	(0.139)	(0.131)
Low Recyclables	-0.449**	-1.347**	-0.364**	-0.388**	-0.369**	-0.202+
	(0.101)	(0.108)	(0.102)	(0.102)	(0.101)	(0.102)
High Recyclables	-0.594**	-1.471**	-0.550**	-0.565**	-0.552**	-0.553**
	(0.124)	(0.0601)	(0.122)	(0.122)	(0.121)	(0.0987)
Deals with Broker			-1.001**	-0.996**	-1.016**	-0.130
			(0.135)	(0.137)	(0.137)	(0.118)
Mean marginal distance			0.105**		0.108**	
			(0.0364)		(0.0389)	
Brokered			-0.740**	-0.710**	-0.754**	-0.470**
			(0.116)	(0.113)	(0.114)	(0.110)
Carter Size				-0.000241	0.000360	
				(0.00114)	(0.00117)	
Carter is L						-2.936**
						(0.134)
Observations	176417	25736	176417	176417	176417	176417
Only Broker	No	Yes	No	No	No	No
R^2	0.0747	0.104	0.0825	0.0820	0.0825	0.119

Note: $+\ p < 0.10$, $*\ p < 0.05$, $**\ p < 0.01$. These set of regressions provide an overview over the data that enters the structural modal and explores several variables along which carters could be divided in groups L and H. Specifications one and two show the successively lower price for customers with higher quantity and higher recyclables. Specifications three to five show that two variables, route density measured in the average marginal distance between customers as well as carter size, are not able to purge differences across carters that deal with brokers and those that do not. The last regressions includes a variable "Carter is L" that directly classifies carters according to their average price, which is able to purge these differences. All regressions have standard errors clustered at the zip-code level in accordance with previous regressions.

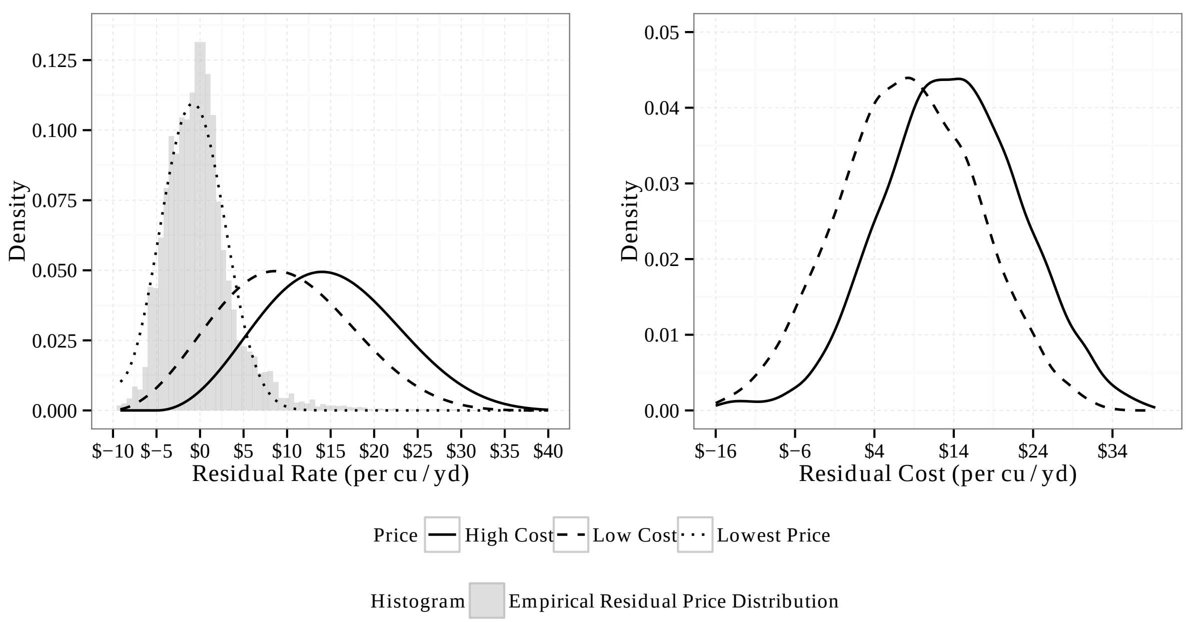
7. RESULTS

The empirical results reveal that search cost are a large fraction of buyers total expenses, ranging from 29% to 48.3%. Search cost also increasing in the quantity of waste that buyers dispose of, which might be due to increasing contractual complexity or because larger waste generators make the search and negotiation process a managerial task. The recovered equilibrium search strategy shows that buyers ask, on average, for more price quotes than there are bidders in the broker market.

7.1 Carter Cost and Equilibrium Bidding

Results of the first-stage cost estimation are shown in Figure 9. The contract price distribution results from the lowest price offers made by low cost bidders (L) and high cost bidders (H). These distributions are not observed but estimated in the first step of the estimation procedure. The dotted line describes the density of the estimated residual contract-price distribution in the brokered market. It closely fits the empirical density, which is shown as a histogram in the same graph. On average there are eight bidders of type L in a typical auction and two bidders of type H. The price offer function of firms of type H is only slightly higher than that of carters L.[59] The densities of low cost bidders and high cost

Figure 9: The residual bid/constract price distributions (left) and residual carter cost (right)



Notes: This figure provides a graphical illustration of the first-stage cost-estimation. The left part shows the semi-nonparametric distributions of bids in the brokered market. The right-hand side shows the resulting density of the residual cost distribution, which is derived using the first order condition as shown in Appendix D.

bidders are not too different, but given the number of solicited bids these small differences can still lead

[59] Draws below -17.0 and above 40.0 have been dropped, which is make up about 0.74% of all simulation draws. There are cost draws that go to $-\infty$ and so these need to be removed before the density is estimated to net let it be driven by such outliers. "Exploding" values are quite common in this two-step procedure.

to large differences in the average price offered by the two types of firms as well as the probability of winning in the auction.

Table 5: Expected cost (standard errors) for carter per cu/yd per month

	Low Cost Carter ($)			High Cost Carter ($)			Winner ($)		
	Recyclables			Recyclables			Recyclables		
Subset	No	Low	High	No	Low	High	No	Low	High
$Q_{0,25}$	17.58	16.23	16.11	23.03	21.67	21.56	6.04	4.67	4.74
	(0.24)	(0.27)	(0.25)	(0.3)	(0.31)	(0.32)	(0.13)	(0.17)	(0.15)
$Q_{25,50}$	15.68	14.33	14.21	21.12	19.77	19.65	4.04	2.72	2.7
	(0.22)	(0.24)	(0.22)	(0.32)	(0.32)	(0.33)	(0.1)	(0.12)	(0.1)
$Q_{50,75}$	14.96	13.6	13.49	20.4	19.05	18.93	3.38	2.05	1.85
	(0.24)	(0.25)	(0.23)	(0.31)	(0.31)	(0.32)	(0.13)	(0.14)	(0.13)
$Q_{75,100}$	14.25	12.89	12.78	19.69	18.34	18.22	2.76	1.33	1.16
	(0.22)	(0.24)	(0.22)	(0.3)	(0.3)	(0.31)	(0.12)	(0.13)	(0.13)

Note: The table shows the expected cost of the population of low cost and high cost bidders as well as the cost for carters conditional on winning. To compute the latter I draw repeatedly (1000 times) from the distribution of bidders. $Q_{a,b}$ refers to quantities from quantile a to b. Bootstrapped standard errors in parentheses (50 iterations).

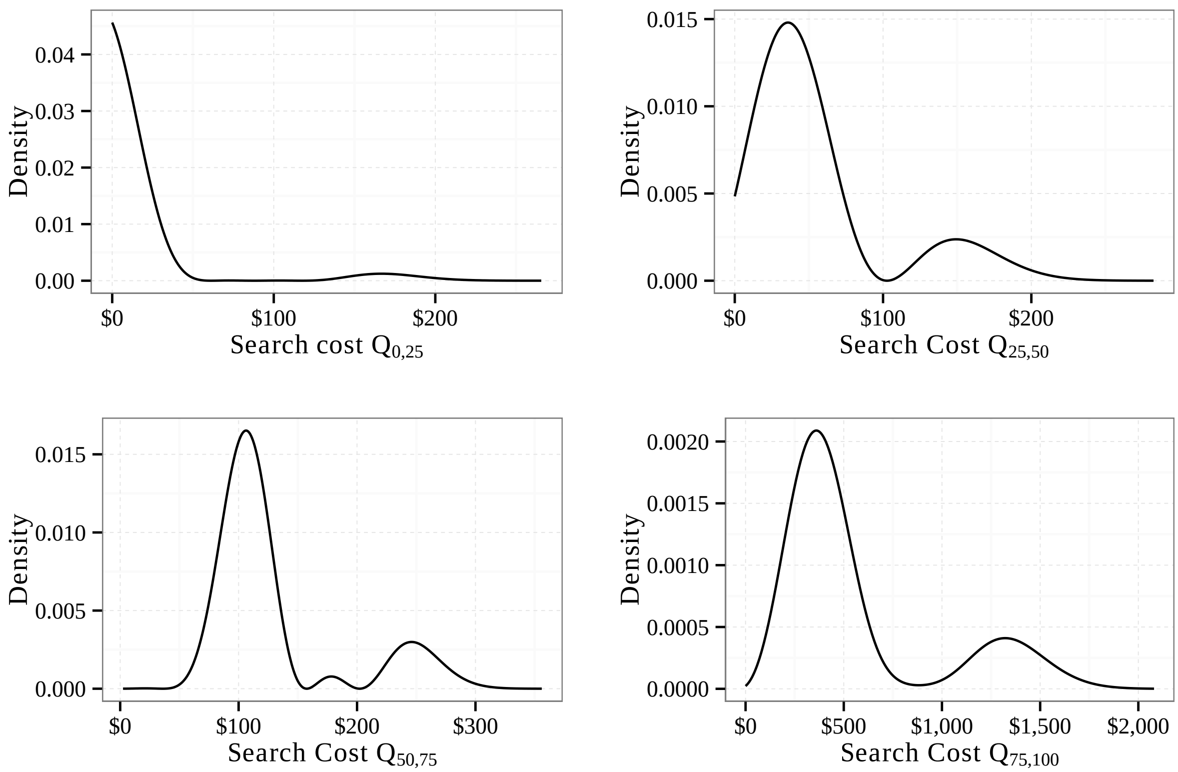
Table 5 shows the expected cost – per cu/yd – of low cost and high cost bidders for contracts along the two dimensions of observed heterogeneity that are included in the model, the quantity of waste generated by the customer and the recyclables. The last column shows the cost conditional on winning the contract under the average number of auction participants of both type. The average cost for a firm of type H is between 31% and 40% than that of a firm of type L. Note that the cost for customers in the upper quartile with High Recyclables are very low. Part of the cost distributions go into the negative domain, due to the value of recyclables, and the resulting cost are therefore an average over mostly positive and some negative values.

7.2 Search cost and Equilibrium Search Strategy

This section describes the estimated search cost of customers of different type, their corresponding equilibrium search strategy as well as the equilibrium pricing strategy of carters under this search rule. Table 6 gives a broad overview. The average search costs per price inquiry is strongly increasing in the quantity of waste generated by the customer. It is $24.24 in the lowest quartile, $56.84 for customers with quantities from the 25th to the 50th percentile, $129.3 for customers with quantities from the 50th to the 75th quartile, and $540.54 in the highest quartile.

What is behind this strong increase in the search cost for higher volume customers? One possible explanation is the increasing contractual complexity and length of negotiations for businesses that dis-

Figure 10: Search Cost Densities



Notes: This figure shows search cost densities. There are four different distributions, conditional on the four different quartiles in terms of the quantity of waste that buyers generate.

pose of large quantities of waste.[60] For large waste generators the disposal might be such an important part of the business that these negotiations become a managerial task whereas smaller waste generators delegate it to lower level employees. Another explanation, if one were to accept a more liberal interpretation of the search-cost, might be the loss of control that large institutions face over such "purchase" decisions. The responsible employees might, for example, have favorite vendors who minimize their hassle cost. Preventing this type of moral hazard becomes increasingly hard in larger institutions.

A separate look at average search cost for customers that use brokers and those that do not gives a sense of the selection into broker services. The search cost of customers that use brokers are between 221% and 724% higher than for customers that do not use brokers. The estimates therefore reflect the high mark-ups charged by brokers which generate this wedge between both markets. Figure 10 shows the four corresponding distributions. A characteristic feature of distributions is the additional mode to the right. This separation in the distribution mirrors the wedge between the observed search prices

[60] In terms of the variation in the data this means that the dispersion in the conditional price distribution for higher quantities is therefore not shrinking fast enough to be rationalized by a constant search cost. To explain this, it is useful to remember that the estimation of the search cost depends on the marginal types, which are computed as $\kappa_m = q \cdot (\mathbb{E}[p^{(m)}] - \mathbb{E}[p^{(m+1)}])$. For the search cost to stay constant as we move to higher quantity customers it would need to be true that this marginal benefit of searching remains constant. This in turn would require the difference in expectations from an additional inquiry to shrink proportional to the increase in the quantity. This difference would be smaller if the price distribution was less dispersed.

Table 6: Search cost per inquiry, Means and Standard Deviations

Subset	Expected Cost Per Inquiry ($) all κ	$\kappa < \Delta_B$	$\kappa > \Delta_B$	Standard Deviation ($) all κ	Average Searches $\kappa < \Delta_B$	Average Search Expenses ($) $\kappa < \Delta_B$
$\mathbf{Q_{0,25}}$						
No Recyclables	24.24 (1.29)	13.88	171.86	40.8 (2.05)	14.13	130.23
Low Recyclables	24.24 (1.29)	13.84	171.46	40.8 (2.05)	13.84	126.31
High Recyclables	24.24 (1.29)	13.84	171.43	40.8 (2.05)	13.74	125.03
$\mathbf{Q_{25,50}}$						
No Recyclables	56.84 (7.68)	40.98	164.17	47.68 (7.41)	12.57	354.12
Low Recyclables	56.84 (7.68)	38.93	158.46	47.68 (7.41)	12.86	372.04
High Recyclables	56.84 (7.68)	38.9	158.34	47.68 (7.41)	12.47	358.51
$\mathbf{Q_{50,75}}$						
No Recyclables	129.3 (3.85)	106.87	254.01	57.76 (6.06)	10.45	929.83
Low Recyclables	129.3 (3.85)	104.67	246.59	57.76 (6.06)	9.56	868.76
High Recyclables	129.3 (3.85)	104.34	244.64	57.76 (6.06)	10.82	892.37
$\mathbf{Q_{75,100}}$						
No Recyclables	564.76 (58.74)	373.45	1339.88	416.21 (54.74)	10.49	3046.71
Low Recyclables	564.76 (58.74)	352.0	1188.65	416.21 (54.74)	10.92	3109.94
High Recyclables	564.76 (58.74)	329.5	1045.55	416.21 (54.74)	11.59	2840.99

Note: This table shows expected search cost and the standard deviation for the four estimated distribution as well the eight different susbet of customer types. Both are computed by gauss-hermite integration. Bootstrapped standard errors in parantheses.

and the higher final broker prices in the descriptive section. There are several explanations for this additional smaller mode in the distribution. These could be buyers that face additional cost (or higher opportunity cost of searching) because they search out of town or because they are in the startup phase of the business.

Table 6 also gives an overview of the total cost that customers incur for searching – i.e. their cost per inquiry times the equilibrium number of price solicitations. For $\mathbf{Q_{0,25}}$ without recyclables, customers in the search market spend in expectation $130.23. To put these expenses into perspective, one can compare the expenses on search to the total expenses (search + contract expenses), which is 33.3% in this case. For the highest quartile without recyclables buyers spend in expectation $3045.01 on their search which is %45.02 of the total expenses.[61] Search cost are therefore a surprisingly large part of the total expenses.

Knowing the value of search cost is important for policy makers. For example, to evaluate the trade-off between a centralized versus decentralized market organization or, as is currently debated in New York, the trade-off between a decentralized market and a procurement system with exclusive

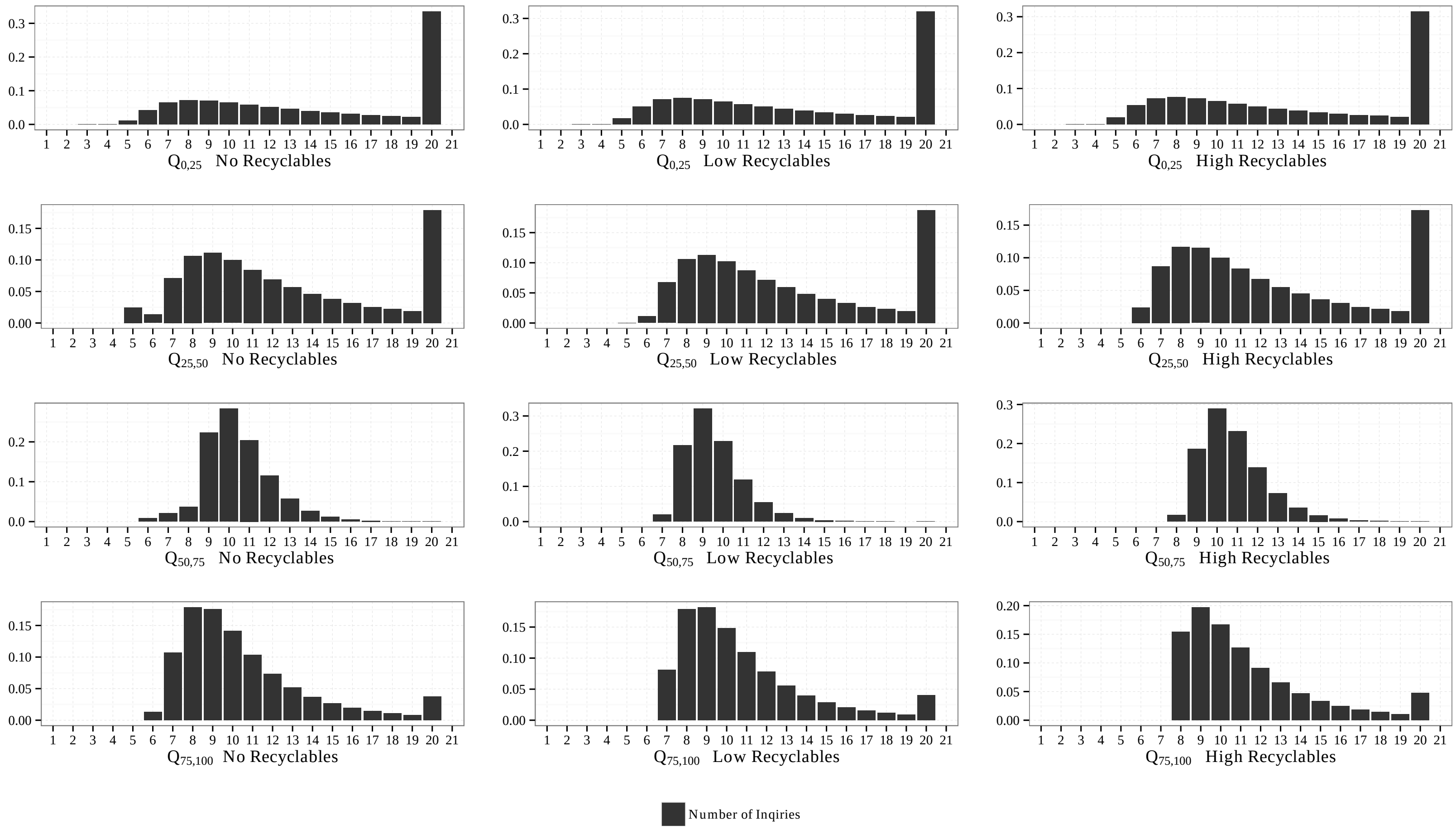
[61] For the remaining cases the numbers are as follows. For $\mathbf{Q_{0,25}}$ with low fraction of recyclables broker cost are $431.67 and search cost 33.3% and with high fraction of recyclables $425.51 and 29.9%. For $\mathbf{Q_{25,50}}$ without recyclables broker expenses are $843.1 and the search fraction is 37.4%, with low fraction of recyclables broker cost are $704.43 and search cost 41.7% and with high fraction of recyclables $691.27 and 37.7%. For the quartile range $\mathbf{Q_{50,75}}$ without recyclables broker spending is $1562.76 and fraction of search cost 41.2%, with Low Recyclables $1294.36 and 38.1% without recyclables respectively and high fraction of recyclables $1266.36 and 44.3%. For $\mathbf{Q_{75,100}}$, the remaining cases are $4521.37 and 47.6% with low fraction of recyclables and $4441.06 and 48.3% with high fraction.

territories. A system with exclusive territories forces buyers to contract with one carter and therefore removes the ability but also the cost of searching. An evaluation of all cost might lead to a different recommendation than one based on price alone.

Part of the estimation is to recover customers unobserved equilibrium search strategies. How many price solicitations do customers make? In the lowest quartile in terms of quantities customers make on average about 14 price inquiries, customers in $\mathbf{Q_{25,50}}$ make approximately 12.5 search inquiries, customers in $\mathbf{Q_{50,75}}$ between 9.5 and 10.8 and customers in the highest quartile on average between 10.5 and 11.3. This compares to an average number of 9.8 bidders in the brokered market. Surprisingly, firms in the search-market therefore solicit more prices than there are bidders in the brokered market. The full distributions of search inquiries are shown in Figure 19. The distributions reveal that in each case some fraction of customers call all available carters in the market and would therefore benefit from additional competition. The histogram also shows the cut-off at the left tail. In the lowest quartile no one makes fewer than three searches and in the upper quartiles, depending on the presence of recyclables the cut-off is between three and six. The customer that would have made fewer searches are contracting in the brokered market.

To give a sense of the uncertainty that customers with different optimal search strategies are facing, Figure 22 in Appendix E shows the price offer densities of two different types of customers in the subset $Q_{25,50}$ without recyclables. The two customers are taken from the extreme points of the search-cost types that are in the search market, the first one making five searches in equilibrium and the second one twenty searches. Customers that make twenty searches can essentially rule out prices above \$13 where the price density is close to zero. The price distribution for customers who maker five searches, on the other hand, has a right tail admitting prices up to \$22.

Figure 11: Number of inquiries made by different customers



Notes: This figure shows histograms of the number of inquiries that are made by buyers in equilibrium.

7.3 Prices, Profit and Market Share

Table 7 shows prices, profits, and market shares of carters. Profits for low cost carters are about 15% to 20% higher than for high cost carters. The percentage of profits are increasing as one moves to higher quantity customers and more recyclables. There is no obvious correlation between the average price and the average number of searches since this is being confounded by the variation in cost due to recyclables and quantities. The market share of low cost carters is about 83%, an over-proportionate amount of the market relative to their share amongst the population of sellers, which is 65%.

Table 7: Number of Search inquiries, Prices, Markups and Market Shares

	Average Price ($)			Average Profit ($)			Market Share	Customer
Subset	All	Low	High	All	Low	High	Low (%)	Searches
$Q_{0,25}$								
No Recyclables	9.3	9.23	9.72	4.87	4.9	4.75	84.26	14.13
Low Recyclables	7.94	7.85	8.43	4.8	4.84	4.6	84.72	13.84
High Recyclables	7.87	7.76	8.44	4.84	4.88	4.58	84.61	13.74
$Q_{25,50}$								
No Recyclables	7.85	7.74	8.45	4.96	5.01	4.7	84.32	12.57
Low Recyclables	5.93	5.83	6.48	4.59	4.63	4.38	84.91	12.86
High Recyclables	6.21	6.1	6.8	4.71	4.76	4.41	84.48	12.47
$Q_{50,75}$								
No Recyclables	7.38	7.26	8.02	4.59	4.63	4.36	84.33	10.45
Low Recyclables	6.38	6.27	6.98	4.63	4.7	4.24	83.54	9.56
High Recyclables	6.01	5.87	6.76	4.91	4.96	4.66	85.07	10.82
$Q_{75,100}$								
No Recyclables	6.89	6.77	7.5	4.78	4.85	4.42	83.45	10.54
Low Recyclables	5.39	5.26	6.05	4.71	4.79	4.33	83.68	10.77
High Recyclables	5.39	5.27	6.03	5.1	5.17	4.76	84.21	11.33

Note: This table shows the average price and markup for customers and carters of different type as well as the market share of contracts for high low cost firms. Lastly, the table shows the average number of price inquiries made by customers in the last column.

A graphical overview over the bidding functions of the two types $-$ L and H $-$ are given in **??**. The graphs show that high cost firms and low cost bid very similar.[62] Despite almost identical bidding behavior low cost firms are more expensive because they win overproportionally in those cases where all bidders had high cost draws and therefore quote high prices.

8. THE ROLE OF INTERMEDIARIES

This section discusses results from the main counter-factual of interest, in which the option to contract through a waste-broker is removed from the market. Before proceeding I ask the question how customers would fare in case they could use the same "search technology" as brokers, which highlights some important differences across the two markets.

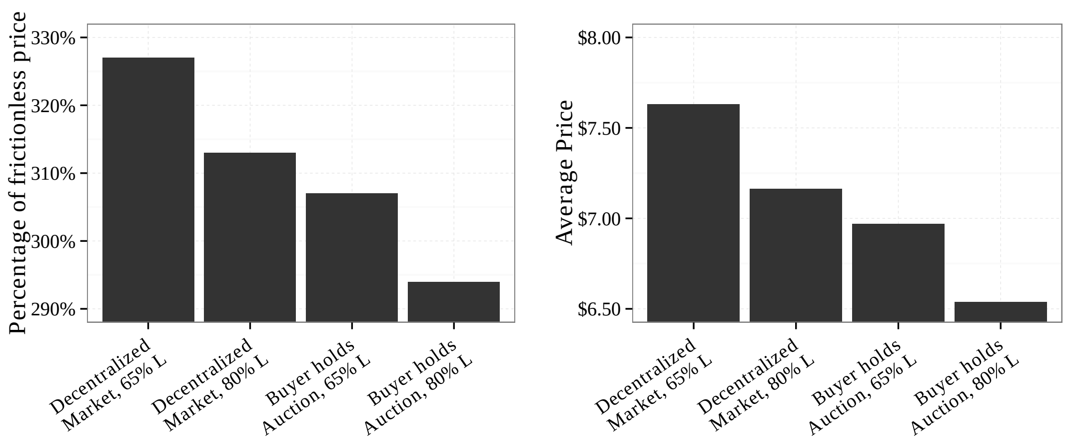
[62]While not being visible on the graph, high cost firms bid on average slightly more aggressive. Integrated over the price distribution, firms of type H have on average, across different buyer types cost that are 0.03$ to 0.05$ higher, which is between 0.8% and 17.8% of the average cost.

8.1 PRICE EFFECTS OF DIFFERENT SEARCH TECHNOLOGIES

Broker auctions have on average about ten participants. With one exception this is less than the average number of price inquiries made by customers, which lies between 9.6 and 14.1. Brokers do therefore not obtain lower prices because they simply "search" more. However, there are two other important differences between the two markets. First, brokers are eliciting prices from a different composition of sellers with a higher percentage of firms of type L. Second, brokers are holding auctions in which all carters know their competitors whereas in the search market carters bid against an uncertain number of competitors.

I now explore what happens when buyers in the search market have the same means at their disposal. In the "Buyer holds auction" scenario I assume that a buyer who wants to make m searches according to her optimal search strategy, instead holds a standard first price auction with m bidders. This isolates the price change due to the different market clearing method. I then interact this change with a change in the composition of bidders, which I move from 65% type L in the search market to 80% in the broker market. The changes in prices are set in relation to the frictionless scenario where buyers obtain a price quote from every firm in the market.[63] Figure 12 shows the average changes in prices across different buyers and Table 12 in Appendix E gives a full overview. What this shows is that

Figure 12: Price Decomposition

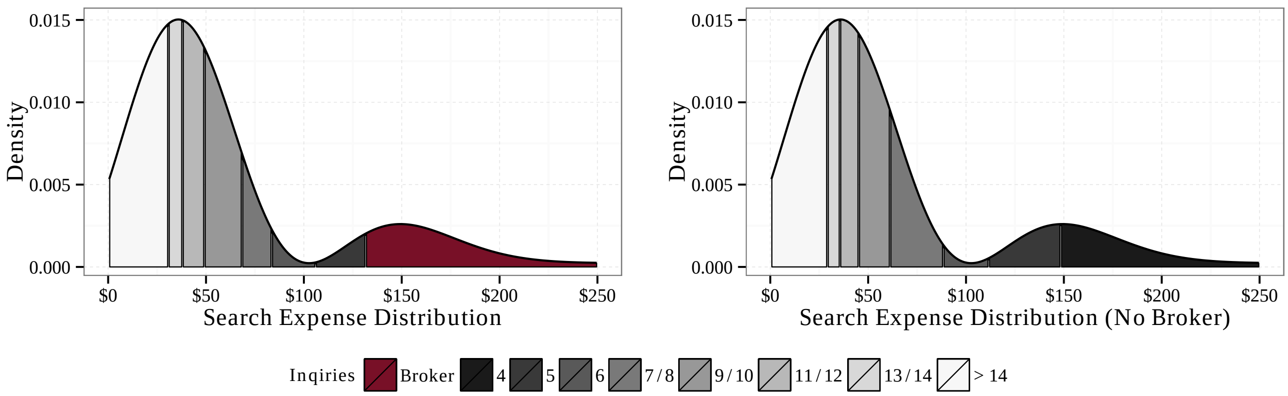


Notes: The figure compares the prices customers get under their optimal search strategy (Decentralized Market) with what they get if instead they were holding an auction where the number of competitors is given by the optimal number of search inquiries. In each case this is done computed for the carter composition in the search market (65% are of type L) as well as the carter composition in the broker auction (80% are of type L). The left panel shows the price relative to the frictionless price, where all carters are asked for price quotes and the right panel the absolute price. All numbers are weighted averages over all customer types.

both the different composition as well as the auction institution itself lead to marked decreases in the average price. Sellers' uncertainty about the number of bidders is therefore an important component in the price increase when going from the broker market to the search market. It would be interesting to explore how general this result is and whether and how it depends on the shapes of the two latent

[63] In his scenario I again also assume the broker market carter composition.

Figure 13: Search-expense distribution



Notes: The figure shows the equilibrium search strategies of customers in the original scenario (left), where a the high search-cost customers still use the broker and on the right where every customer searches in the search-market.

distributions of search cost and sellers service cost.

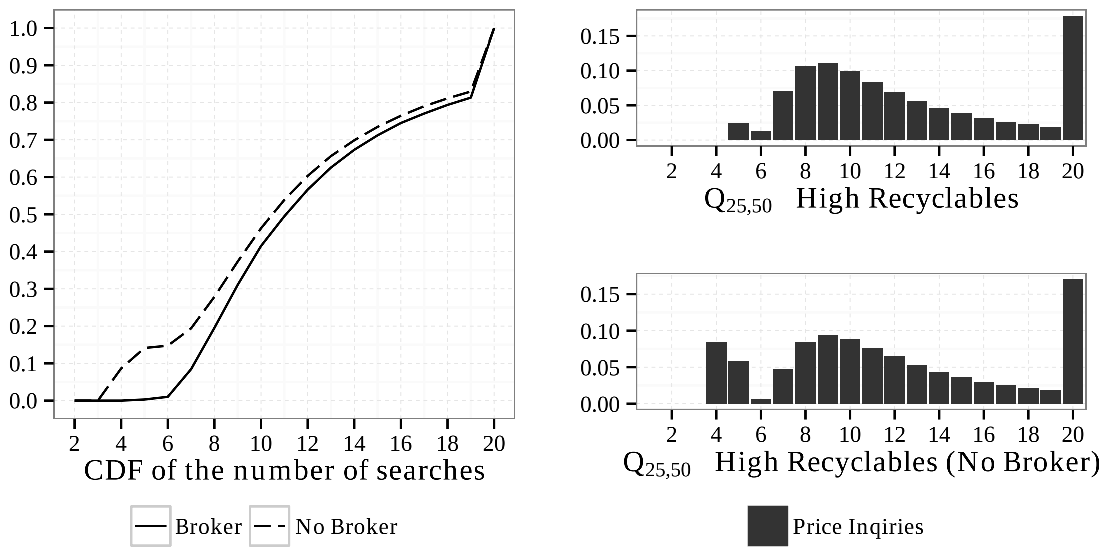
8.2 Market without Intermediaries

I now discuss the main counterfactual. The absence of broker services introduces several changes on both market sides. On the buyer side it changes the composition of search cost in the search market as well as the market price and therefore overall expenses for the contract. It also increases the number of buyers that have to search. Brokers are loosing the commission. For the carters expected profits change because buyers in the search market search less. Losses in welfare come from an increase in the overall expenses on search as well as the fact that contracts are now served at higher average cost.

8.2.1 An Overview

Before proceeding to a full description of the results, Figure 13 provides a graphical representation for the sorting of buyers with observable $\mathbf{Q_{25,50}}$ and High Recyclables (the qualitative insights are the same across buyer types and I therefore only show this example). The two graphs are the two empirical counterparts to Figure 5 in the model section and illustrate the main intuition of the changes on the buyer side. The red portion of the density in the distribution on the left panel are those customers that were using brokers in the original scenario. Applying the optimal search rule, these buyers are now situated at the left tail of the distribution of price inquiries. The fraction of customers that make four and five price inquiries expands as a result. This means that sellers are now faced with a different distribution of buyers, which lowers the expected number of competitors. Another observation from the graph is the intensive margin adjustments of searchers. For example, the probability mass of buyers that make five searches is moving to the right, which means the buyer who take this action now have higher average search cost. This is due to a re-evaluation of the marginal benefit and the marginal cost of searching. Prices in the market go up and buyers want to partially offset this effect by searching more. The resulting change becomes even more clear when comparing the two CDFs of the equilibrium search strategy of customers in Figure 14, which shows that the old distribution of search inquiries

Figure 14: CDF of search of equilibrium search strategies



Notes: The left panel shows the cumulative CDFs for the two original and the counter-factual search strategy and the right panel the two histograms.

strictly stochastically dominates the new one.

For all customer types, sellers raise their prices due to the changing composition of buyers and the decrease in the average number of searches, which ranges from 0.3% to 14.08%. The increase in prices ranges from 2.7% to 33.7%. Firms benefit substantially from the newly arriving high search cost customers with realized profits increasing between 3.8% and 43.7%. On average, across buyers of different type, profits increase by 20%. This effect is stronger for high cost firms, whose profits increase on average by 23.5% relative to low cost firms, whose profits increase by 19.8%. Surprisingly, however, there is almost no change in the market share that low cost firms capture. Customers' expenses rise because of the increase in prices and because they adjust by searching more. Buyer with high search cost, who were contracting through brokers, spent on average 17% more, whereas buyers in the search market pay about 13.25% more. For the latter, most of the increase is due to an increase in prices, which raises the contract cost by 14.46%. Search cost increase by 9.47%, which is due to customers re-evaluation of the optimal amount of searching. A complete overview over the changes in these variables is provided in Table 14 and Table 15 in Appendix E.

8.2.2 Welfare Calculations

The following section summarizes what these market changes mean for welfare. One important caveat, that needs to be dealt with, is that the fixed cost for the operation of broker services is not included in the calculations since the model does not produce an estimate for this. But it is possible to bound these fixed cost and therefore the welfare change. The upper bound (on the cost increase, i.e. welfare decrease) results from the assumption that the fixed cost are zero. The lower bound can be obtained by assuming that the total fixed cost are equal to the total observed variable profits, which equals the total commission payments.

The total change introduced by the in-availability of brokers can be composed into different parts. Let ce_t^s be customers' search cost in t when they were buying in the search market initially. This equals search cost plus contract cost, $ce_t^s = se_t^s + cc_t^s = se_t^s + q \cdot p_t^s$. Likewise, let $ce_t^b = se_t^b + cc_t^b = se_t^b + q \cdot p_t^b$ be the expenses in t for customers that were buying in the brokered market initially. Search cost in the broker market are defined as the fraction of expenses that are marked as fees, i.e. $se_t^b = \phi \cdot p_t^b \cdot q$. Let firms realized profits for customers that were originally in the search market be $\pi_t^s = q \cdot (p_t^s - c_t^s)$ and $\pi_t^b = q \cdot (p_t^b - c_t^b)$ for those originally in the brokered market. Finally, let η be the share of customers who contract in the search market. The upper bound on the total welfare change, ΔUBW, can then be computed as:

$$\begin{aligned}\Delta UBW = \eta \cdot &\Big[(q \cdot (p_1^s - c_1^s) - se_1^s - q \cdot p_1^s) - (q \cdot (p_0^s - c_0^s) - se_0^s - q \cdot p_0^s)\Big] \\ &+ (1-\eta) \cdot \Big[(q \cdot (p_1^b - c_1^b) - se_1^b - q \cdot p_1^b) - (q \cdot (p_0^b - c_0^b) - se_0^b - q \cdot p_0^b) - se_0^b\Big]\end{aligned}$$

In the last term an extra se_0^b is subtracted since search cost in the brokered market are just a transfer from buyers to brokers. Prices are eliminated from this expression to obtain the final expression for the upper bound on the welfare change:

$$\Delta UBW = \eta \cdot \Big[q \cdot c_0^s + se_0^s - q \cdot c_1^s - se_1^s\Big] + (1-\eta) \cdot \Big[q \cdot c_0^b - q \cdot c_1^b - se_1^b\Big]$$

By subtracting the broker variable profits the lower bound on the changes is obtained:

$$\Delta LBW = \Delta UBW - (1-\eta) \cdot q \cdot \phi \cdot p_0^b$$

Since every buyer needs to contract in this market, total welfare in the market can be measured in the total cost to provide the service plus the search-cost that the market incurs in order to produce matches between buyers and sellers. The results below show that both margins are important. Looking at the upper bound, one can obtain three different terms that allow for a straightforward interpretation of the sources of welfare improvements: (1) The change in search cost for customer that were already buying in the brokered market: $\eta \cdot (se_0^s - se_1^s)$, (2) the additional search cost for customers that were using brokers $-\eta \cdot se_1^b$ as well as (3) the change in the total service-cost, $\eta \cdot \Big[q \cdot c_0^s - q \cdot c_1^s\Big] + (1-\eta) \cdot \Big[q \cdot c_0^b - q \cdot c_1^b\Big]$.

The following are the changes in these three components of the upper bound. Changes of search cost of buyers that were already contracting in the search market (1) have a annual total value of $5.83 Million and makes up 13.3% of the total change, (2) the added search cost due to newly entering buyers $20.3 Million and are 46.5% of the total and the increase due to service cost (3) is $17.5 Million and makes up 40.1% of the total. This amounts to a total annual change of $43.6 million, which is approximately 12.2% of the annual market volume.[64] The lower bound is obtained by subtracting the full variable

[64] This decomposition of the total change is only provided for the upper bound. For the lower bound one would have to decide how to "label" the broker commissions. The most logical approach would be to account for them in the change of search cost for formerly brokered customers. According to this definition search cost for formerly brokered customers would rise by $ 9.65 million in the scenario without brokers.

profits of brokers which are \$25.05 million. The total change would then be \$18.6 million or 5.1% of the annual market volume.

Figure 15 gives a more in-depth overview about the relative cost-changes on both market sides. Table 13 in Appendix E an overview over the absolute changes. It reveals a number of consistent changes across rows. The total expenses for *both* market sides are increasing. In all cases the percentage change is sizable. The effect for buyers that were formerly contracting through a broker is stronger in percentage terms. But since there are many more buyers who were already searching the total contribution to the cost increase is larger from this market-side. Total expenses for buyers in the search market increase between 2.9% and 21.5%.

8.3 Summary and Discussion of Counterfactual Results

Intermediaries in this market redistribute rents from sellers to buyers. By keeping high search cost buyers away from the search market they make it more competitive for sellers. This lowers rents for sellers, especially those that have persistently higher cost. Buyers in the search market benefit through lower prices and lower search expenses. This in itself is an important insight for policy makers, such as city planners who are interested in increasing consumer welfare, which depends on the variety of local businesses rather than large rents for utility providers.

This study raised the question, to what extend intermediaries benefit buyers, who are *directly* involved in a transaction with an intermediary as well as buyers who search by themselves through a search externality that lowers prices (*indirect effect*). The results above show that buyers expenses in the search market increase in percentage terms almost as much as the expenses of buyers who were transacting with brokers. These counter-factual results therefore highlight the important *indirect effects* of intermediaries. This means that these services are in general under-provided since intermediaries create social returns that are not reflected in their private pay-offs. While this result provides a rationale for direct market interventions supporting intermediaries it is also an important insight for regulatory information provision and disclosure policies (for examples, see Jin and Leslie (2003)). It means that small penetration rates of such policies can lead to much larger effects through supply side adjustments.

A natural question is whether these results would translate to other markets with search frictions and buyer-specific cost. Quantitative results will necessarily depend on the exact primitives that govern behavior in other markets. However, one would expect that in many markets, that fall within this paradigm, search frictions are even more severe, which possibly increases the welfare effect created by intermediaries. In markets for investment goods, such as construction equipment or production lines, products are often ordered with extremely idiosyncratic specifications and sellers in these markets are often scattered over many countries, which increases the cost of search and negotiations.[65]

[65] Note, however, that in markets with extremely idiosyncratic features buyers typically organize their own "request for proposals". This is, for example, true for large architectural projects and capital equipment purchases.

Figure 15: Overview of Changes when Brokers are not available

	Total Customer Expenses (%)		Breakdown Externality (%)						
Subset	Formerly Broker	Formerly Search	Contract Expenses	Search Expenses	Market Search Cost UB (%)	Market Search Cost LB (%)	Total Service Cost (%)	Total Market Cost UB ($1000)	Total Market Cost LB ($1000)
$Q_{0,25}$									
No Recyclables	16.3	18.3	20.3	10.4	20.6	-6.4	10.8	1832.2	615.6
Low Recyclables	26.3	21.5	26.7	3.7	13.5	-8.6	11.1	14.6	2.6
High Recyclables	11.7	2.9	2.0	6.1	17.9	-4.7	-3.5	98.4	1.0
$Q_{25,50}$									
No Recyclables	13.8	14.3	14.6	13.3	39.0	1.4	21.8	6019.6	2769.1
Low Recyclables	18.5	15.0	17.8	8.1	35.7	-0.1	46.2	102.0	48.7
High Recyclables	19.6	8.8	9.8	6.1	31.4	-2.0	29.7	1075.8	375.0
$Q_{50,75}$									
No Recyclables	7.7	10.3	11.6	6.5	28.2	-1.8	22.4	10849.4	4866.7
Low Recyclables	15.6	6.1	6.3	5.5	31.0	-0.2	27.3	616.6	251.8
High Recyclables	17.5	15.0	18.8	6.5	34.4	2.1	68.9	4305.8	2404.5
$Q_{75,100}$									
No Recyclables	24.4	11.3	11.9	9.8	41.7	-1.5	27.4	37423.5	13831.9
Low Recyclables	31.0	10.8	12.4	7.4	43.1	2.3	72.4	8620.3	3621.1
High Recyclables	34.3	13.8	15.0	11.4	65.5	4.1	136.2	15579.7	6033.2

Note: This table gives an overview over percentage changes in the total cost (search cost + contract) for buyers that were already transacting in the search market and those who were previously transacting through broker. For buyers already in the search market, the table provides a further breakdown into the changes in search expenses and contract expenses. For buyers in the broker market such a breakdown does not make much sense since the initial broker expenses provide no logical separation into search and contract expenses. The broker commissions are in general higher than the search expenses in the new scenario so that search cost according to this definition would decrease for those buyers although their total expenses increase. The next four columns give the upper and lower bounds on the changes in total search cost that the market incurs for making matches as well as the total cost for servicing buyers. The last two columns show the upper and lower on total $-changes. All numbers are computed over the length of a contract, which is two years.

9. EXTENSION: MERGER ANALYSIS

Counter-factual merger analysis is arguably one of the most important applications of models in empirical industrial organization. Merger analysis in posted price markets is very developed mostly due to discrete choice demand models (usually paired with the assumption of differentiated product Bertrand price competition), foremost Berry (1994) and Berry et al. (1995). The strength of this literature is that by incorporating unobserved brand preference shocks it allows for realistic handling of firm heterogeneity and cross-substitution patterns, which is important to understand how prices evolve after the merger.[66]

Merger analysis is much less developed for markets where prices are arranged individually despite the fact that they are so common. The important aspects one needs to keep track of to understand how a merger will change prices are somewhat different in such a setting. They are typically outside the realm of retail and idiosyncratic tastes − like brand preferences − are not as important in such settings. But other sources of heterogeneity are important to keep track of, like in this setting the unobserved search cost for customers. Persistent cost differences on the firm side make the effect of the merger on consumer welfare ambiguous. Indeed, one of the most prominent arguments in favor of mergers are cost synergies that are passed on to consumers (see for example Farrell and Shapiro (2001)).

The literature has noticed this lack and there is now a growing interest to explore the trade-offs of mergers in negotiated price markets with search frictions. A recent example is Allen et al. (2013), who look at the Canadian mortgage market. The mortgage rates in the market are individually negotiated between the lenders and buyers. Like in the setting here buyers are heterogeneous in their search cost. A number of assumptions allow the authors to elegantly map the model into a a reduced form (changes-in-changes) estimator, which makes for a transparent identification strategy. But some of the assumptions are restrictive. According to the bargaining protocol buyers first obtain two "free" price quotes and then need to decide whether they want to exert effort to acquire additional price quotes in a second stage where with some probability they receive the maximum number of quotes n or otherwise none at all. The probability of n quotes is increasing in effort and effort cost depend on the search cost.[67] In addition to that, sellers observe the type of the customer and therefore offer them the reservation price. In equilibrium customers therefore never search.[68] This means that the price distribution only depends on customers type through firms optimal price offers and that every searching customer benefits from an additional firm in the market.

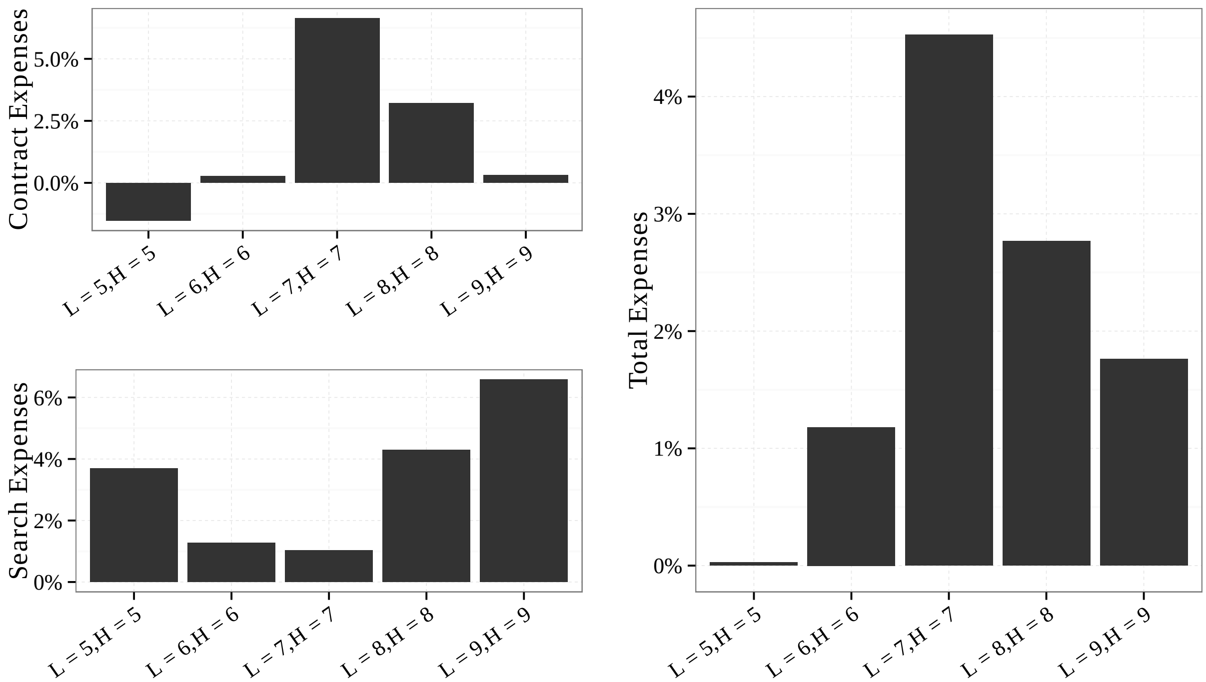
In my setting I can relax some of these assumptions. In reality, when search actually takes place, some customers might never exhaust all the options and an additional firm only matter through prices. Other customers with particular low search cost might also directly benefit from having an additional option. The cost composition of firms matter to all customers due to the uncertainty about which firms will be encountered. Since the distribution of firms' service cost and customers' search cost interact with each other and jointly give rise to observed prices, a change induced by a merger can have different effects for different initial cost compositions in the market. Most importantly, since search in the model actually takes place, a recommendation based on predicted prices alone can differ from the recommendation when all the cost to buyers, including search cost, are considered.

[66] Suggestions have been made (Moraga-González et al. (2015)) to extend this line of work, introducing search cost while still allowing for unobserved brand heterogeneity.

[67] Such a bargaining framework is more often used in the lab our literature. See Postel-Vinay and Robin (2002)

[68] This also implies that there is no social cost due to searching.

Figure 16: Expenses in different Merger Scenarios



Notes: These graphs show changes in contract expenses and search expenses (on the two left panels) and total expenses (on the right) for different merger scenarios.

Figure 16 provides an example, using the estimates for customers in the second quartile ($\mathbf{Q_{25,50}}$) with High Recyclables. For different initial numbers of firms of type L and H the graph shows what happens if two firms of different type merge and create a new firm that has an L cost distribution. In each case the graph documents the contract and search-expenses in the left panel and the final expenses on the right. According to a pure price criterion one would only allow the merger when there are initially five firms of each type in the market, since only then the cost composition effect outweighs the benefit for customers to be able to call an additional firm. The final price increase, however, is the result of buyers new equilibrium search strategy and contrary to the change in prices, total search cost rise in all scenarios for the average customer. Taking all expenses into consideration, the merger would always lead to an increase in total expenses for customers, if only slightly in the case of ten initial firms. While this is only one particular example, it shows that it is important to take search-cost into consideration.

10. CONCLUSION

This paper studies the competitive and welfare effects of intermediation in a decentralized market. Such a structure is common in retail services markets, wholesale trade markets, and markets for investment goods. I argue that intermediaries can give rise to search externalities that benefit buyers who do *not* contract through intermediaries. The self-selection of buyers with high search cost into the intermediated market changes the composition of buyers in the search market and makes it more competitive for sellers. To quantify these effects of intermediation the study makes use of a new and detailed dataset from from the New York City trade-waste industry, which provides a comprehensive and rare insight in a decentralized market.

Methodologically, this study contributes to the literature by introducing a new search model that takes into account the idiosyncratic cost of servicing buyers. This is accomplished by bringing together elements from the empirical search and procurement-auction literature. The identification of the model is challenging, since one needs to tell apart the distribution of sellers' service-cost from buyers' search-cost, both of which are unobserved. This stands in contrast to previous empirical studies in the search literature, which only estimates a full distribution of search cost under a uniform cost assumption and posted prices. The joint data from the brokered and the bilateral search market together identify search- and service cost. An important institutional detail is that prices in the brokered market are formed according to competitive bidding. This pins down the cost of carters. Known cost distributions can then be used in conjunction with prices in the search market to back out the distribution of search inquires. The latter plus an exclusion restriction allows to recover buyers' search cost.

The estimates reveal that search cost are an important cost factor for buyers that one needs to take into account when thinking about policy, such as merger guidelines or the trade-off between the prevalent decentralized market and a system of procurement of exclusive territories, as currently contemplated by the city.

Counterfactuals show that intermediaries redistribute a sizable portion of rents from sellers to buyers and improve overall welfare by reducing the search cost incurred by the market and by reallocating contracts to lower cost suppliers. The results highlight the importance of positive search externalities created by intermediaries, which means that their service is under-provided.

Several possible follow-up projects emerge from the analysis. The model currently abstracts from the competition between brokers, which might be important to consider for some applications. Another possible direction for further research is to extend the model to allow for differentiation in the quality of service provision or the dynamic aspects of buyers' choice amongst sellers.

REFERENCES

Akerman, Anders et al., "A theory on the role of wholesalers in international trade based on economies of scope," *Research Papers in Economics*, 2010, *1.*

Allen, Jason, Robert Clark, and Jean-François Houde, "The Effect of Mergers in Search Market: Evidence from the Canadian Mortgage Industry," Technical Report, National Bureau of Economic Research 2013.

Armstrong, Mark, "Search and ripoff externalities," 2014.

Asker, John, "A study of the internal organization of a bidding cartel," *The American Economic Review*, 2010, *100* (3), 724–762.

_ **and Estelle Cantillon**, "Properties of scoring auctions," *The RAND Journal of Economics*, 2008, *39* (1), 69–85.

Athey, Susan and Philip A Haile, "Identification of standard auction models," *Econometrica*, 2002, *70* (6), 2107–2140.

Bajari, Patrick, "Comparing competition and collusion: a numerical approach," *Economic Theory*, 2001, *18* (1), 187–205.

_ **and Gregory Lewis**, "Procurement contracting with time incentives: theory and evidence," Technical Report, National Bureau of Economic Research 2009.

Bar-Isaac, Heski and Alessandro Gavazza, "Brokers' contractual arrangements in the Manhattan residential rental market," 2013.

Barwick, Panle Jia and Parag A Pathak, "The costs of free entry: an empirical study of real estate agents in Greater Boston," Technical Report, National Bureau of Economic Research 2011.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 1995, pp. 841–890.

Berry, Steven T, "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 1994, pp. 242–262.

Biglaiser, Gary, "Middlemen as experts," *The RAND journal of Economics*, 1993, pp. 212–223.

Blum, Bernardo S, Sebastian Claro, and Ignatius Horstmann, "Intermediation and the nature of trade costs: Theory and evidence," *University of Toronto, mimeograph*, 2009.

Burdett, Kenneth and Kenneth L Judd, "Equilibrium price dispersion," *Econometrica: Journal of the Econometric Society*, 1983, pp. 955–969.

Commission, Citizen Budget, "Taxes In, Garbage Out The Need for Better Solid Waste Disposal Policies in New York City," 2012.

Diamond, Peter A, "A model of price adjustment," *Journal of economic theory*, 1971, *3* (2), 156–168.

Farrell, Joseph and Carl Shapiro, "Scale economies and synergies in horizontal merger analysis," *Antitrust Law Journal*, 2001, pp. 685–710.

Gallant, A Ronald and Douglas W Nychka, "Semi-nonparametric maximum likelihood estimation," *Econometrica: Journal of the Econometric Society*, 1987, pp. 363–390.

Gavazza, Alessandro, "Leasing and secondary markets: theory and vvidence from commercial aircraft," *Journal of Political Economy*, 2011, *119* (2), 325–377.

— , "An empirical equilibrium model of a decentralized asset market," *Available at SSRN 1601883*, 2012.

Gehrig, Thomas, "Intermediation in search markets," *Journal of Economics & Management Strategy*, 1993, *2* (1), 97–120.

Giulietti, Monica, Michael Waterson, and Matthijs Wildenbeest, "Estimation of search frictions in the British electricity market," *The Journal of Industrial Economics*, 2014, *62* (4), 555–590.

Goldberg, Pinelopi Koujianou, "Dealer price discrimination in new car purchases: Evidence from the consumer expenditure survey," *Journal of Political Economy*, 1996, pp. 622–654.

Gonzalez, José Luis Moraga, Zsolt Sándor, and Matthijs R Wildenbeest, "Prices, Product Differentiation, and Heterogeneous Search Costs," 2014.

Grennan, Matthew, "Price discrimination and bargaining: Empirical evidence from medical devices," *Available at SSRN 1721145*, 2012.

Guerre, Emmanuel, Isabelle Perrigne, and Quang Vuong, "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 2000, *68* (3), 525–574.

Haile, Philip A, "Auctions with resale markets: an application to US forest service timber sales," *American Economic Review*, 2001, pp. 399–427.

— **, Han Hong, and Matthew Shum**, "Nonparametric tests for common values at first-price sealed-bid auctions," Technical Report, National Bureau of Economic Research 2003.

Han, Lu and WC Strange, "The microstructure of housing markets: Search, bargaining, and brokerage," *Handbook of Regional And Urban Economics*, 2014, *5.*

Harmon, Nikolaj A, "Are Workers Better Matched in Large Labor Markets?," *Job market paper, Princeton University*, 2013.

Hellwig, Klaus-Peter, "Intermediaries in a Decentralized Market with Information Frictions: Facilitators or Bottlenecks?," 2013.

Hendel, Igal, Aviv Nevo, and François Ortalo-Magné, "The Relative Performance of Real Estate Marketing Platforms: MLS versus FSBOMadison. com," *American Economic Review*, 2009, *99* (5), 1878–1898.

Hong, Han and Matthew Shum, "Using price distributions to estimate search costs," *The RAND Journal of Economics*, 2006, *37* (2), 257–275.

Honka, Elisabeth, "Quantifying search and switching costs in the US auto insurance industry," *The Rand Journal of Economics*, 2014, *45* (4), 847–884.

— **and Pradeep K Chintagunta**, "Simultaneous or sequential? search strategies in the us auto insurance industry," *Search Strategies in the US Auto Insurance Industry (April 29, 2014)*, 2014.

Hortaçsu, Ali and Chad Syverson, "Product differentiation, search costs, and competition in the mutual fund industry: a case study of the S&P 500 Index Funds," Technical Report, National Bureau of Economic Research 2003.

il Kim, Kyoo and Joonsuk Lee, "Nonparametric Estimation and Testing of the Symmetric IPV Framework with Unknown Number of Bidders," 2014.

Inderst, Roman and Marco Ottaviani, "Competition through commissions and kickbacks," *The American Economic Review*, 2012, pp. 780–809.

Jacobs, James B, Coleen Friel, and Robert Raddick, *Gotham unbound: How New York city was liberated from the grip of organized crime*, NYU Press, 2001.

Jensen, J Bradford, Stephen J Redding, Peter K Schott, and Andrew B Bernard, "Wholesalers and Retailers in US Trade," *Peterson Institute for International Economics Working Paper*, 2010, (10-10), 2011–89.

Jin, Ginger Zhe and Phillip Leslie, "THE EFFECT OF INFORMATION ON PRODUCT QUALITY: EVIDENCE FROM RESTAURANT HYGIENE GRADE CARDS.," *Quarterly Journal of Economics*, 2003, *118* (2).

Kelly, Robert J, "Dirty Dollars: The Carting and Waste Disposal Industries," in "The Upperworld and the Underworld," Springer, 1999, pp. 135–155.

Koulayev, Sergei, "Search for differentiated products: identification and estimation," *The RAND Journal of Economics*, 2014, *45* (3), 553–575.

Krasnokutskaya, Elena, "Identification and estimation of auction models with unobserved heterogeneity," *The Review of Economic Studies*, 2011, *78* (1), 293–327.

— **and Katja Seim**, "Bid preference programs and participation in highway procurement auctions," *The American Economic Review*, 2011, *101* (6), 2653–2686.

Krishna, Vijay, *Auction theory*, Academic press, 2009.

Lee, Robin S and Kyna Fong, "Markov perfect network formation: An applied framework for bilateral oligopoly and bargaining in buyer-seller networks," 2013.

Levitt, Steven D and Chad Syverson, "Market distortions when agents are better informed: The value of information in real estate transactions," *The Review of Economics and Statistics*, 2008, *90* (4), 599–611.

Li, Huagang and John G Riley, "Auction choice," *International Journal of Industrial Organization*, 2007, *25* (6), 1269–1298.

Li, Tong and Quang Vuong, "Nonparametric estimation of the measurement error model using multiple indicators," *Journal of Multivariate Analysis*, 1998, *65* (2), 139–165.

— **and Xiaoyong Zheng**, "Entry and competition effects in first-price auctions: theory and evidence from procurement auctions," *The Review of Economic Studies*, 2009, *76* (4), 1397–1429.

Lizzeri, Alessandro, "Information revelation and certification intermediaries," *The RAND Journal of Economics*, 1999, pp. 214–231.

los Santos, Babur De, Ali Hortaçsu, and Matthijs R Wildenbeest, "Testing models of consumer search using data on web browsing and purchasing behavior," *The American Economic Review*, 2012, *102* (6), 2955–2980.

Marshall, Robert C, Michael J Meurer, Jean-Francois Richard, and Walter Stromquist, "Numerical analysis of asymmetric first price auctions," *Games and Economic Behavior*, 1994, *7* (2), 193–220.

Matzkin, Rosa L, "Nonparametric identification," *Handbook of Econometrics*, 2007, *6*, 5307–5368.

McCall, John Joseph, "Economics of information and job search," *The Quarterly Journal of Economics*, 1970, pp. 113–126.

Meilijson, Isaac, "Estimation of the lifetime distribution of the parts from the autopsy statistics of the machine," *Journal of Applied Probability*, 1981, pp. 829–838.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R Wildenbeest, "Consumer search and prices in the automobile market," 2015.

Morgan, Peter and Richard Manning, "Optimal search," *Econometrica: Journal of the Econometric Society*, 1985, pp. 923–944.

Nguyen, Thuy TT and Bruce G Wilson, "Fuel consumption estimation for kerbside municipal solid waste (MSW) collection activities," *Waste Management & Research*, 2010, *28* (4), 289–297.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

Rothenberg, Thomas J, "Identification in parametric models," *Econometrica: Journal of the Econometric Society*, 1971, pp. 577–591.

Royte, Elizabeth, *Garbage land: On the secret trail of trash*, Back Bay Books, 2007.

Rubinstein, Ariel and Asher Wolinsky, "Middlemen," *The Quarterly Journal of Economics*, 1987, pp. 581–594.

Rust, John and George Hall, "Middlemen versus Market Makers: A Theory of Competitive Exchange," *Journal of Political Economy*, 2003, *111* (2).

Salop, Steven and Joseph Stiglitz, "Bargains and ripoffs: A model of monopolistically competitive price dispersion," *The Review of Economic Studies*, 1977, pp. 493–510.

Silverman, Bernard W, *Density estimation for statistics and data analysis*, Vol. 26, CRC press, 1986.

Sorensen, Alan T, "Equilibrium price dispersion in retail markets for prescription drugs," *Journal of Political Economy*, 2000, *108* (4), 833–850.

Spulber, Daniel F, "Market making by price-setting firms," *The Review of Economic Studies*, 1996, *63* (4), 559–580.

_ , "Market microstructure and intermediation," *The Journal of Economic Perspectives*, 1996, pp. 135–152.

Stahl, Dale O, "Oligopolistic pricing with sequential consumer search," *The American Economic Review*, 1989, pp. 700–712.

Stigler, George J, "The economics of information," *The journal of political economy*, 1961, pp. 213–225.

Stiglitz, Joseph E, "Imperfect information in the product market," *Handbook of industrial organization*, 1989, *1*, 769–847.

A. DATA CODING

A.1 Definition of Brokered Contract and Number of Bidders

There are two data fields from which a brokered contract can be inferred. The first one is a field that contains an indicator on how the contract was formed. A second field provides the name of the brokerage company. A contract is defined as brokered if either the first field indicates so or if the latter field contains an entry that does not clearly disqualify as a broker company. In some cases, for example, it is clear that the field does not contain the contact name of a broker but instead the customer name itself.

There is no numeric identifier for brokers. The broker identifier has therefore been hand coded from the broker-name field that carters have to fill out in the customer register. In most cases it is straightforward to assign a given observation to a recurring broker-name. In some cases however, the listed name might be an individual contact name and therefore does not allow me to match this observation to a broker company. The broker identifier is used to infer the number of bidders, defined as all carters that were awarded a contract during the observation period within the borough of the contract. If a broker name only appears once it mostly the case because a contact person as opposed to the actual company is listed and I therefore randomly assign the number of bidders from the remaining distribution. Sometimes a contract between a customer and a carter is indicated as "brokered" in the second reporting period (remember that reporting periods are half-years) although no other adjustments occured, including the price. In that case I assume that the initial contract was already brokered. The data also shows when a contract is closed after solicitation through a carter. I am currently ignoring this information although it could interesting in future work to separately identify the search cost from haggling cost

A.2 Matching Broker Commissions

The broker prices are not part of the original *customer register* data. The BIC has collected this information for the first time in 2015. In the absence of a business identifier I can only match the expected broker commission to a contract in the *customer register*. I am therefore using a flexible hedonic regression in terms of the observables that are available in both data-sets. I am using a regression of the commission on a 15th order polynomial of the customers total billings and also include zip-code fixed effects. The R^2 from this regression is 0.29.

B. ADDITIONAL DISCUSSION OF THE IDENTIFICATION OF UNOBSERVED CUSTOMER SEARCH STRATEGY

The arguments for identification of $\mathbf{w}_o$ follows those for general parametric model as provided in Rothenberg (1971). The general conditions guarantee local identification if the information matrix is non-singular. A parameter $\mathbf{w}_0$ is said to be locally identified if there exists an open neighborhood of $\mathbf{w}_0$ containing no other $\mathbf{w}$.

We start from equation Equation 4, which relates the observed CDF of prices to the known cost function primitives and the unknown $\mathbf{w}_0$, where for convenience the dependence on covariates is suppressed.

$$\mathcal{F}^O(p|\mathbf{w}) = \sum_{m=1}^{M} w_m \cdot \sum_{k=0}^{m} \frac{\binom{N_L}{k} \cdot \binom{N_H}{m-k}}{\binom{N_H+N_L}{m}} \cdot \left(1 - \widetilde{\mathcal{G}}_L(\beta_{S,L}^{-1}(p,\mathbf{w}))^k \cdot \widetilde{\mathcal{G}}_H(\beta_{S,H}^{-1}(p,\mathbf{w}))^{m-k}\right).$$

Define to express the above equation compactly as: $\mathcal{F}^O(p|\mathbf{w}) = \sum_{m=1}^{M} w_m \cdot \tilde{\mathcal{F}}^m(p|\mathbf{w})$ and taking derivatives w.r.t. p on both sides $f^O(p|\mathbf{w}) = \sum_{m=1}^{M} w_m \cdot \tilde{f}^m(p|\mathbf{w})$. Under the following assumptions:

1. The structural parameter space is an open set in R^m.
2. The function f^O is a proper density function for every $\mathbf{w}$ in $\mathcal{W}$. In particular, f^O is nonnegative and the equation $\int f^O(p,\mathbf{w})dp = 1$ holds for all $\mathbf{w}$ in $\mathcal{W}$.
3. The set B of p values for which f^O is strictly positive is the same for all $\mathbf{w}$ in $\mathcal{W}$. B is termed the sample space of P.
4. The function $f^O$is smooth in $\mathbf{w}$. Specifically, we assume that for all $\mathbf{w}$ in a convex set containig $\mathbf{W}$ and for all p in the sample space B the functions f^O and $log(f^O)$ are continously differentiable with respect to $\mathbf{w}$.

local identification of $\mathbf{w}$ is equivalent with the information matrix

$$R(\mathbf{w}_0) = \begin{bmatrix} E\left(\frac{\partial log(f^O(p|\mathbf{w}))}{\partial w_1} \cdot \frac{\partial log(f^O(p|\mathbf{w}))}{\partial w_1}\right) & \dots & E\left(\frac{\partial log(f^O(p|\mathbf{w}))}{\partial w_1} \cdot \frac{\partial log(f^O(p|\mathbf{w}))}{\partial w_M}\right) \\ \vdots & \ddots & \vdots \\ E\left(\frac{\partial log(f^O(p|\mathbf{w}))}{\partial w_M} \cdot \frac{\partial log(f^O(p|\mathbf{w}))}{\partial w_1}\right) & \dots & E\left(\frac{\partial log(f^O(p|\mathbf{w}))}{\partial w_M} \cdot \frac{\partial log(f^O(p|\mathbf{w}))}{\partial w_M}\right) \end{bmatrix}$$

being non-singular, where $\mathbf{w}_0$ is a regular point of the matrix, which means that there exists an open neighborhood around $\mathbf{w}_0$ in which R has constant rank.

Note that it is hard, if not impossible, to show analytically that the requirements for local identification discussed above are satisfied. For example, the optimal bidding strategies of carters depends on $\mathbf{w}$ and it is hard to derive comparative statics in the differential equations that describe players strategies. But since I implement a step-wise procedure to estimate w, it can be directly checked through a monte carlo study whether a set of weights can be estimated under the first stage carter cost estimates. For that I use the proposed minimum distance estimator (Equation 11) based on the functional Equation 4. I simulate 5000 and 20000 prices under two sets of (synthetic) search weights. Table 8 reports the true underlying weights, the average recovered weights as well as the standard deviation of the recovered weight. Due to computational time, these experiments are restricted to smaller weight vectors with six and eight firms. Each of the four monte-carlo studies is based on thirty repetitions. The results confirm that the unobserved weights can be recovered accurately and with small standard errors from the estimator that I propose.

Table 8: Monte Carlo Results

Vector Entry	True Weight S1	S2	Mean (standard deviation) Recovered $N = 5000$	$N = 20000$	$N = 5000$	$N = 20000$
w_1	0.3	0.3	0.306 (0.010)	0.304 (0.007)	0.300 (0.007)	0.300 (0.006)
w_2	0.0	0.0	0.020 (0.025)	0.013 (0.022)	0.006 (0.014)	0.004 (0.011)
w_3	0.6	0.0	0.560 (0.065)	0.575 (0.061)	0.029 (0.027)	0.030 (0.026)
w_4	0.0	0.3	0.020 (0.043)	0.013 (0.041)	0.270 (0.024)	0.272 (0.023)
w_5	0.1	0.1	0.097 (0.010)	0.096 (0.012)	0.081 (0.042)	0.090 (0.031)
w_6	-	0.15	-	-	0.168 (0.043)	0.153 (0.024)
w_7	-	0.15	-	-	0.147 (0.006)	0.148 (0.004)

Note: The table reports on a monte carlo study, which is based on the minimum distance estimator Equation 11. Results are shown for two different set of weights (S1 and S2) and two different sample sizes, 5000 and 20000. The table shows the true weights and the recovered weights as well as standard errors of the recovered weights in parentheses.

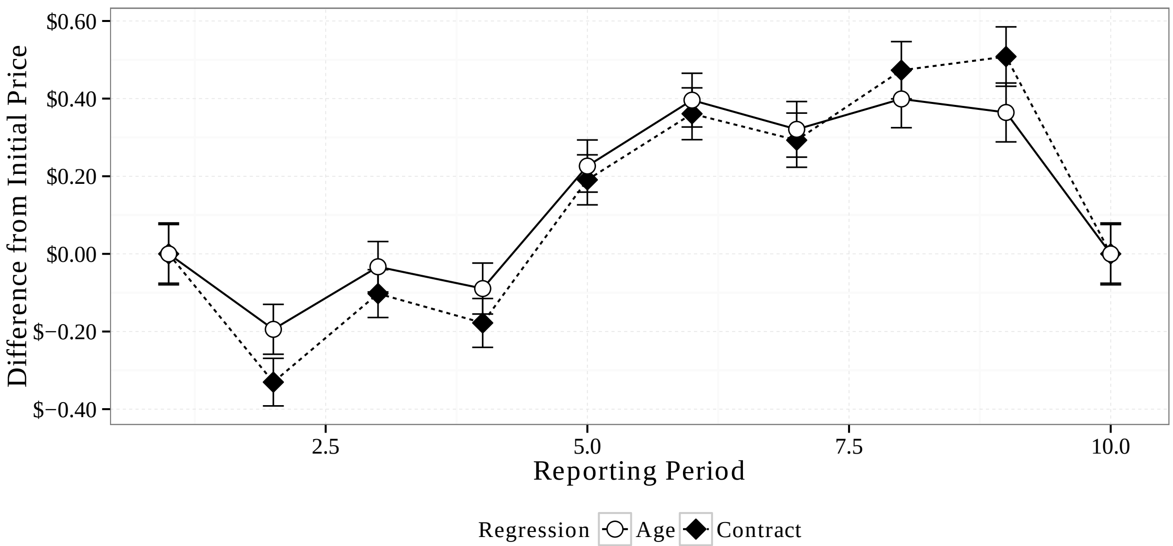
C. DYNAMIC DEVELOPMENT OF PRICES

This project focuses on the static aspects of pricing in this market. But since customers are entering a potentially longer lasting relationship with carters, they might be forward looking with regard to the prices that carters charge in the future and this might affect their initial decision to search. Sellers might, for example, engage in bait and switch pricing by which they lure customers with initial low rates and then increase their prices when the customer is "locked in". This might for example work if customers have significant switching cost. Another consideration might be that customers learn about the suppliers as time passes and decreasing search cost is reflected in a decreasing price path. To give a sense of the price developments of contracts across time I run two different regressions. Both regressions again include the same controls as the price dispersion regressions:

$$\mathbf{X} = \{\text{business type FE, recyclables FE, time FE, zip code FE, transfer station FE}, q, \ldots, q^5, \\ \text{Number of Pickup FE}\}.$$

The variable of interest is now a time dummy, which summarizes how prices move across time holding the above characteristics fixed. The first regression explores how prices develop within a customer carter relationship. The second regression looks at the pure age effect, not conditioning on a contract. The graphs below summarize the development of prices. Effects are shown relative to the excluded category, which is the initial price. The reporting periods are six month. We can see that there is an initial small but significant decline in prices and a subsequent increase.

Figure 17: Dynamic development of prices



Notes: This figure shows how the prices develop for buyers over time, relative to the initial price. A reporting period is a half-year. The data points are time-dummy estimates, where the excluded variable is the dummy on the initial time period. The dotted time series conditions on the contract, i.e. holds the carter fixed. Graphs show that the changes are modest in the initial two years.

D. DETAILS ON COST SIMULATION FROM PRICE OFFERS

In the standard asymmetric auction case without uncertainty about the number and composition of competitors the maximization problem for a firm of type k is:

$$\max_{p_k} = (p_k - c) \cdot (1 - G_k(\beta_k^{-1}(p)|\mathbf{z}))^{N_k - 1} \cdot (1 - G_{-k}(\beta_{-k}^{-1}(p)|\mathbf{z}))^{N_{-k}} \tag{12}$$

Base on this, the first order

$$\begin{aligned}
- (p_k - c) \cdot \Bigg[& \frac{1}{\beta_k'(\beta_k^{-1}(p)|\mathbf{z}))} g(\beta_k^{-1}(p)) \cdot (N_k - 1) \cdot (1 - G_k(\beta_k^{-1}(p)|\mathbf{z}))^{N_k - 2} \cdot (1 - G_{-k}(\beta_{-k}^{-1}(p)|\mathbf{z}))^{N_{-k}} \\
& + \frac{1}{\beta_{-k}'(\beta_{-k}^{-1}(p)|\mathbf{z}))} g(\beta_{-k}^{-1}(p)) \cdot N_{-k} \cdot (1 - G_{-k}(\beta_{-k}^{-1}(p)|\mathbf{z}))^{N_{-k} - 1} \cdot (1 - G_{-k}(\beta_{-k}^{-1}(p)|\mathbf{z}))^{N_{-k}} \Bigg] \\
& + (1 - G_k(\beta_k^{-1}(p)|\mathbf{z}))^{N_k - 1} \cdot (1 - G_{-k}(\beta_{-k}^{-1}(p)|\mathbf{z}))^{N_{-k}} = 0
\end{aligned} \tag{13}$$

Cancelling this simplifies to:

$$(p_k - c) \cdot \left[\frac{1}{\beta_k'(\beta_k^{-1}(p)|\mathbf{z}))} \cdot \frac{g(\beta_k^{-1}(p)) \cdot (N_k - 1)}{(1 - G_k(\beta_k^{-1}(p)|\mathbf{z}))} + \frac{1}{\beta_{-k}'(\beta_{-k}^{-1}(p)|\mathbf{z}))} \cdot \frac{g(\beta_{-k}^{-1}(p)) \cdot N_{-k}}{(1 - G_{-k}(\beta_{-k}^{-1}(p)|\mathbf{z}))} \right] = 1 \tag{14}$$

Now let $\mathcal{F}_k^0(p|\mathbf{z}_j)$ be the distribution of price offers for firm of type k. Using the insight in Guerre et al. (2000) this relates to the cost distribution $\mathcal{G}(.)$ in the following way:

$$\mathcal{F}(p|\mathbf{z}_j) = \mathcal{G}(\beta_k^{-1}(p))|\mathbf{z}_j) = \mathcal{G}(\beta_k^{-1}(p)|\mathbf{z}_j) \tag{15}$$

Taking the derivative with respect to p we can translate this into a realtionship for the densities:

$$f_k^0(p) = g(\beta_k^{-1}(p)) \cdot \frac{1}{\beta_k'(\beta_k^{-1}(p)|\mathbf{z}))} \tag{16}$$

Substituting Equation 15 and Equation 16 into Equation 14 and rearranging, the cost can be expressed in terms of observable (in this case estimated) disrtibutions which can be used for the simulation of the cost distribution:

$$c = p_k - \left[\cdot \frac{f_k^0(p) \cdot (N_k - 1)}{(1 - F_k(p|\mathbf{z}))} + \frac{f_{-k}^0(p) \cdot N_{-k}}{(1 - F_{-k}(p|\mathbf{z}))} \right]^{-1} \tag{17}$$

E. ADDITIONAL TABLES

E.1 BREAKDOWN OF BUSINESS-TYPES

Table 9: Types of Businesses signed up with carters

Business type	fraction of total	total number
Retail non - food	.410	46514
Retail - food	.138	15660
Wholesale non - food	.015	1648
Wholesale - food	.010	1048
Restaurand/bar	.107	12181
Hotel - small	.003	302
Hotel-big	.002	175
Medical offices	.028	3138
Automobile repair	.027	3100
Office building - small	.032	3620
Office building - medium	.026	2997
Office building - large	.015	1675
Light manufacturing	.020	2242
Heavy manufacturing	.003	298
Institution	.020	2277
Professional office	.043	4838
None of the above	.104	11883

Note: Fractions and totals are averaged across the nine reporting periods. Missings are declared as "none of the above".

E.2 Robustness Check: Using only low commissions

Table 10: Comparing final prices for brokered contracts with prices in non-brokered market

	(1)	(2)	(4)	(5)	(6)
	OLS	OLS	Quantile (0.25)	Quantile (0.50)	Quantile (0.75)
	r_{ijt} (rate + comission)	r_{ijt} (rate + comission)	r_{ijt} (rate + comission)	r_{ijt} (rate + comission)	r_{ijt} (rate + comission)
Broker	0.445**	0.634**	0.658**	0.407**	1.012**
	(0.133)	(0.121)	(0.0250)	(0.0360)	(0.0303)
Quantity	-0.0201**	-0.0195**	-0.0124**	-0.0140**	-0.0181**
	(0.000757)	(0.000740)	(0.000147)	(0.000222)	(0.000199)
Recyclables	-0.715**	-0.712**	-3.809**	-4.909**	-3.899**
	(0.179)	(0.175)	(0.0239)	(0.0343)	(0.0289)
Deals with broker		-2.408**			
		(0.171)			
Observations	92244	92244	100941	100941	100941
Deals with Broker	No	Yes	No	No	No
Transfer FE	Yes	Yes	No	No	No
R^2	0.266	0.273			

Note: $+\ p < 0.10$, $*\ p < 0.05$, $**\ p < 0.01$. All specifications include the following set of controls: quantity of waste, transfer-station fixed effects, zip-code fixed effects, business-type fixed effects, length-of-contract fixed effects, recyclable materials fixed effects, reporting-date fixed effects, number of weekly pickups, and the HHI index. The aggregate regressions at the zip-code level include the average quantity at the zip-code level, the average number of pickups and the average number of customers that use recyclables. Standard errors are clustered at the zip-code level.

E.3 The Fee structure for customers with different volumes

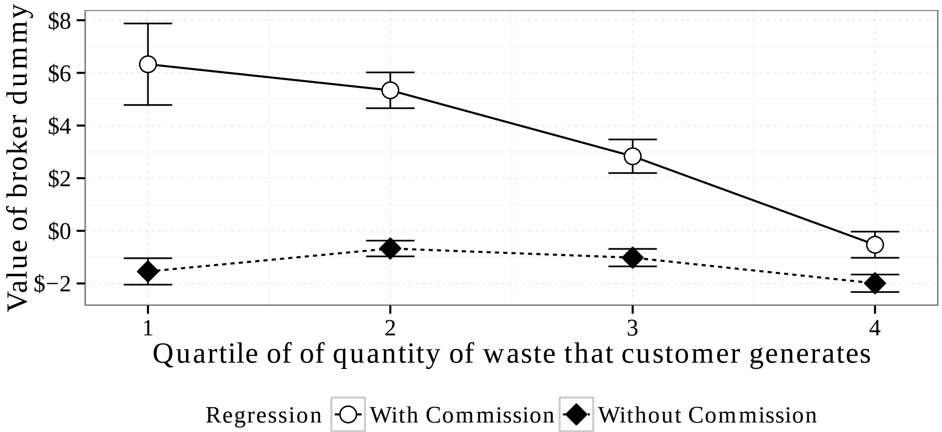
Table 2 documents the fact that amongst larger waste generators there is a higher fraction of brokered contracts. While higher quantity not necessarily means that these are also larger businesses one would still expect a strong positive correlation between the two. To the extend that this is true, it is somewhat puzzling that larger institutions outsource the service while smaller businesses search “in-house”. One would expect the opposite. The following regressions show that this might be explained by the fee structure of brokers. Table 11 repeats the main OLS specifications of Table 3 and Table 4, which explore price differences between brokered and non-brokered contracts, but splitting the effect in terms of the quartiles of the quantity of waste. The main coefficients of interest are displayed in Figure 18. While the difference between the negotiated price and the search market price is relatively flat, the discount for the negotiated price plus the commission increases as one moves to higher quantities. The prices in the broker market make it therefore increasingly attractive for large buyers to delegate their search to brokers.

Table 11: Comparing final prices for brokered contracts with prices in non-brokered market

Variable	(1 OLS) rate	(2 OLS) rate + commission
$Q_{0,25}$	2.955**	1.946**
	(0.373)	(0.458)
$Q_{25,50}$	1.097**	0.0847
	(0.270)	(0.354)
$Q_{50,75}$	-0.440	-1.469**
	(0.268)	(0.359)
$Q_{75,100}$	-0.471	-1.495**
	(0.321)	(0.398)
$\text{Brokered}_{0,25}$	-1.543**	6.328**
	(0.256)	(0.790)
$\text{Brokered}_{25,50}$	-0.671**	5.338**
	(0.152)	(0.347)
$\text{Brokered}_{50,75}$	-1.019**	2.833**
	(0.169)	(0.326)
$\text{Brokered}_{75,100}$	-1.992**	-0.528*
	(0.169)	(0.253)
Recyclables	0.0860	0.0411
	(0.200)	(0.209)
Observations	98687	95052
R^2	0.287	0.305

Note: $+\ p < 0.10$, $*\ p < 0.05$, $**\ p < 0.01$. All specifications include the following set of controls: transfer-station fixed effects, zip-code fixed effects, business-type fixed effects, length-of-contract fixed effects, recyclable materials fixed effects, reporting-date fixed effects, number of weekly pickups, and the HHI index. The aggregate regressions at the zip-code level include the average quantity at the zip-code level, the average number of pickups and the average number of customers that use recyclables. Standard errors are clustered at the zip-code level.

Figure 18: The bidding functions for customers of different type



E.4 PRICE EFFECT OF DIFFERENT SEARCH TECHNOLOGIES: FULL TABLE

Table 12: Price relative to frictionless price with when buyer hold auction versus search-market.

Buyer	Decentralized Market		Buyer holds Auction	
Type	65% carter are L	80% carter are L	65% carter are L	80% carter are L
$Q_{0,25}$				
No Recyclables	166.70	154.81	154.86	142.78
Low Recyclables	187.00	172.97	170.86	157.22
High Recyclables	191.20	175.71	173.17	160.19
$Q_{25,50}$				
No Recyclables	215.93	198.80	197.14	182.29
Low Recyclables	257.36	243.15	232.04	215.19
High Recyclables	278.96	258.96	244.49	230.07
$Q_{50,75}$				
No Recyclables	251.07	246.20	233.64	226.42
Low Recyclables	393.41	383.54	349.63	344.41
High Recyclables	438.85	404.22	392.30	364.82
$Q_{75,100}$				
No Recyclables	305.23	289.72	274.25	262.38
Low Recyclables	688.63	652.18	587.25	572.50
High Recyclables	745.51	695.35	652.86	609.98

Note: The table compares the prices customers get under their optimal search strategy (Decentralized Market) with what they would get if instead they were holding an auction where the number of competitors is given by the optimal number of search inquiries. All numbers are expressed in percentages of the frictionless price, where all carters are asked for price quotes. In each case this is done computed for the carter composition in the search market (65% are of type L) as well as the carter composition in the broker auction (80% are of type L).

E.5 Absolute Cost and Welfare Changes

Table 13: Overview of Changes when Brokers are not available

Subset	Total Customer Expenses (%) Formerly Broker	Total Customer Expenses (%) Formerly Search	Breakdown Externality (%) Contract Expenses	Breakdown Externality (%) Search Expenses	Market Search Cost UB	Market Search Cost LB	Total Service (%) Cost (%)	Total Market Cost UB ($1000)	Total Market Cost LB ($1000)
$Q_{0,25}$									
No Recyclables	171.9	112.0	99.3	12.7	129.4	-335.5	26.0	1832.2	615.6
Low Recyclables	241.0	118.3	113.6	4.7	125.1	-278.1	18.8	14.6	2.6
High Recyclables	106.2	15.8	8.2	7.6	129.1	-268.8	-5.9	98.4	1.0
$Q_{25,50}$									
No Recyclables	288.0	205.5	159.0	46.5	373.6	-357.3	89.8	6019.6	2769.1
Low Recyclables	324.1	182.4	154.5	27.9	347.3	-251.4	97.3	102.0	48.7
High Recyclables	335.8	109.8	88.1	21.8	346.6	-241.0	63.5	1075.8	375.0
$Q_{50,75}$									
No Recyclables	325.7	344.9	285.9	59.0	847.6	-480.8	208.8	10849.4	4866.7
Low Recyclables	549.2	185.2	137.2	48.0	790.5	-278.7	157.6	616.6	251.8
High Recyclables	601.9	414.2	357.9	56.3	803.8	-237.2	265.4	4305.8	2404.5
$Q_{75,100}$									
No Recyclables	3433.3	1263.5	969.1	294.4	2874.9	-1651.7	679.6	37423.5	13831.9
Low Recyclables	3530.3	1033.2	805.6	227.6	2802.3	-756.0	620.3	8620.3	3621.1
High Recyclables	3839.3	1240.0	900.6	339.4	2693.0	-486.8	656.3	15579.7	6033.2

Note: This table shows the increase in search cost and contract-cost separately for buyers that were formerly using brokers (OB) and those that did not (ONB). The changes in expenses for customers that were formerly using brokers are not broken down since the initial division in search cost and contract cost is not meaningful in this case.

Table 14: Changes to the search market when possibility of brokerage is removed

	Customer			Carter			
	Number	Number	Average	Realized Profits			Market Share
Subset	Searchers	Searches	Price	*All*	*L*	*H*	*L*
$\mathbf{Q_{0,25}}$							
No Recyclables	-6.191	-3.955	26.472	38.55	36.906	47.892	1.184
Low Recyclables	-6.191	-3.763	33.687	43.747	42.589	50.679	-0.021
High Recyclables	-6.103	-0.292	2.733	3.816	3.722	4.497	-0.451
$\mathbf{Q_{25,50}}$							
No Recyclables	-11.739	-8.376	20.228	16.218	16.291	15.745	0.15
Low Recyclables	-13.043	-10.879	27.332	17.181	17.225	17.585	-1.539
High Recyclables	-13.043	-6.09	19.381	13.269	13.196	13.922	-0.57
$\mathbf{Q_{50,75}}$							
No Recyclables	-13.043	-10.19	16.29	11.387	11.156	13.4	-0.999
Low Recyclables	-14.821	-6.172	12.563	7.503	6.902	11.316	-0.522
High Recyclables	-15.11	-14.084	25.331	12.265	12.616	11.209	-1.494
$\mathbf{Q_{75,100}}$							
No Recyclables	-16.037	-6.779	22.733	18.457	17.988	20.887	0.329
Low Recyclables	-17.56	-6.001	25.572	16.723	16.38	18.326	0.541
High Recyclables	-22.118	-4.29	27.636	16.714	16.249	20.009	-0.911

Note: All numbers are percentage changes from the original scenario observed in the data to a conterfactual scenario where broker intermediation is not allowed. There are two rows with zero changes. For these two scenarios the model predicted no brokered contracts in the original scenario.

Table 15: Counterfactual Results (absolute)

	Customer			Carter			
	Number	Number	Average	Realized Profits			Market Share
Subset	Searchers	Searches	Price	*All*	*L*	*H*	*L*
$Q_{0,25}$							
No Recyclables	100.0	13.551	11.538	355.376	354.504	360.323	0.85022
Low Recyclables	100.0	13.326	10.572	373.845	374.255	371.542	0.84886
High Recyclables	100.0	13.662	8.081	270.665	272.721	259.601	0.84327
$Q_{25,50}$							
No Recyclables	100.0	11.716	8.963	800.832	809.92	753.132	0.83995
Low Recyclables	100.0	11.567	7.603	796.605	804.911	755.026	0.83351
High Recyclables	100.0	11.704	7.33	784.34	792.2	743.253	0.83941
$Q_{50,75}$							
No Recyclables	100.0	9.422	8.274	1721.628	1749.884	1581.521	0.83218
Low Recyclables	100.0	9.015	7.132	1729.336	1757.189	1592.212	0.83117
High Recyclables	100.0	9.26	6.902	1733.357	1763.912	1584.857	0.82935
$Q_{75,100}$							
No Recyclables	100.0	9.846	8.352	6776.342	6867.924	6310.893	0.83559
Low Recyclables	100.0	9.915	6.806	6633.906	6716.399	6213.501	0.83596
High Recyclables	100.0	10.419	6.392	6539.64	6631.443	6080.567	0.83335

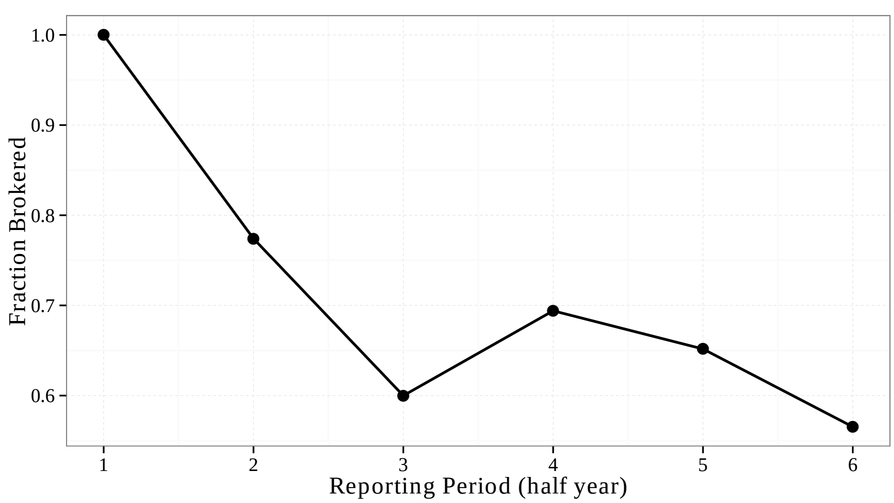
Note: The table shows the absolute values of counter-factual search cost and firm-profits if customers no longer have the opportunity to use a broker.

F. ROBUSTNESS CHECK

The full non-parametric identification of $\mathcal{H}(.|\mathbf{x})$ relies on the marginal types $\kappa_m(\mathbf{x},\mathbf{z})$, $m \in \{1,\ldots,M\}$ and κ_b fully tracing out the distribution but in the application I only use a finite number of values for z. In particular one might be worried, that without a z for which $\kappa_b \to \infty$ the counter-factual results depend too much on very high search cost expenses for types that have higher search cost than κ_b and where there is not enough discipline on the shape of the density. One can, however, bound the maximal welfare change, and the changes in search expenses, by assuming that all probability mass of those types lies at $\kappa_b + \epsilon$ for a very small ϵ. This is the minimal value that the search cost can take on given the partial knowledge about the search cost distribution. Using this bounded distribution, I find that the upper bound on the welfare change is decreased from $43.6 Million to $39.1 Million (from 12% of market volume to 10.7% of market volume) and the lower bound from $18.4 Million to $14.0 Million (from 5.1% of market volume to 3.8% of market volume). Total expenses for those buyers that were using brokers rise from 11.48% (as opposed to 17% without bounding). This shows that imposing this very extreme cut-off for the distribution does not alter the results too much, both in terms welfare and the expenses for buyers and sellers.

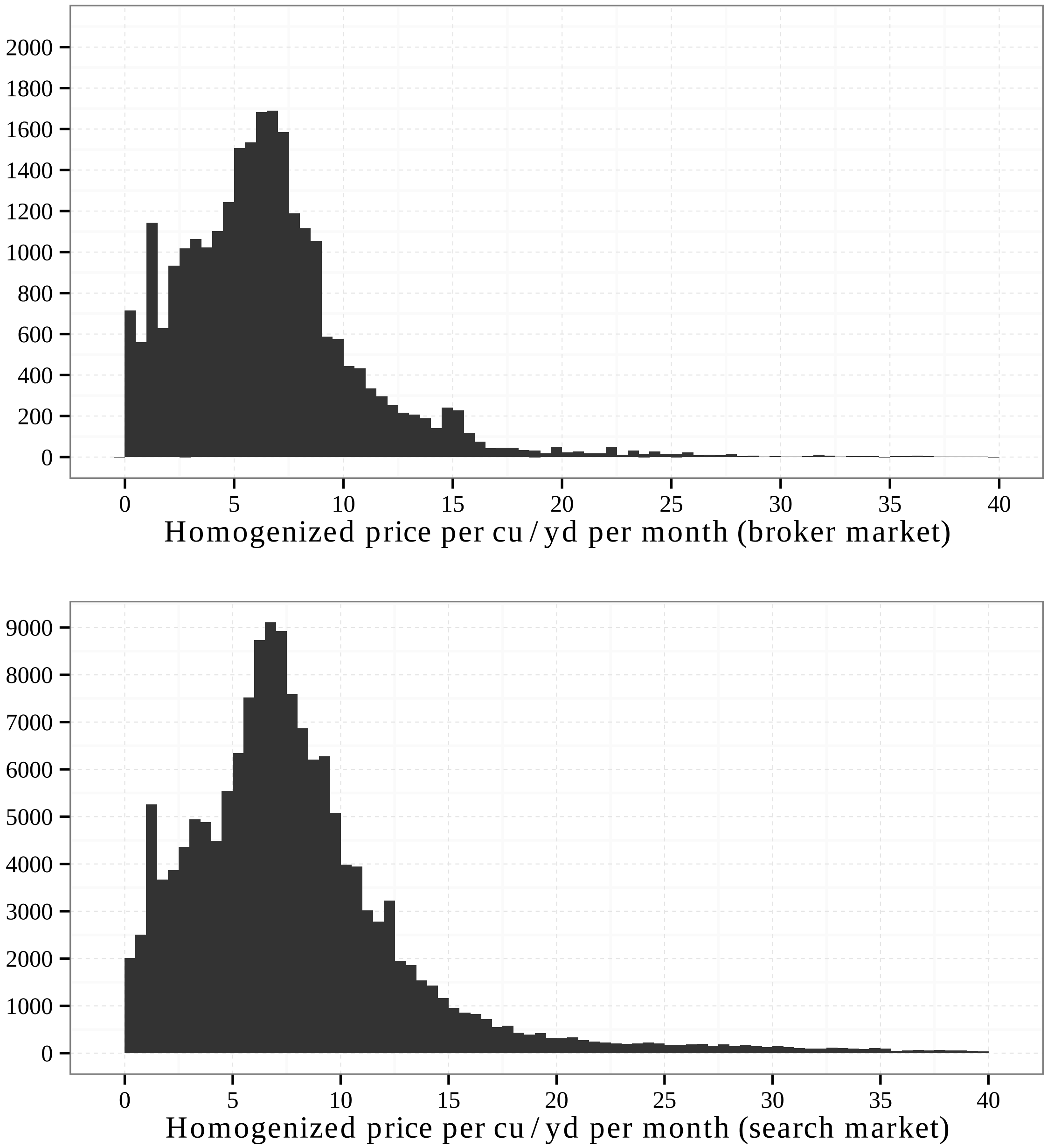
G. ADDITIONAL FIGURES

Figure 19: Customers who cut out the middlemen



Notes: This graph conditions on all customer-carter relationship that were initially brokered and tracks the percentage of ongoing contracts that are still brokered over time. We can see that about forty percent of customers drop the broker after 18 month. This is consistent with the interpretation that brokers are important for the initial match making or search and that customers want to stop paying the broker fee after they have an established relationship with a carter.

Figure 20: Homogenized Price Histograms



Notes: These histogram show prices that are homogenized along several observable dimensions of a contract. In analogy to a price deflator I first run a regression of the log-price on a set of dummy variables: the type of business, the reporting period, the transfer station as well as the number of pick-ups per week. Based on the results of this regression I compute predicted values $\hat{y}_i$ and re-base the price, dividing through the exponential of the predicted value and multiplying by the exponential of the estimated intercept: $p_i \cdot exp(\hat{b}_0)/exp(\hat{y}_i)$. The observation is now interpreted as belonging to the joint set of observations defined by the excluded variables in each set of dummy variables. The choice on these excluded variables therefore determines to what category prices are re-normalized. I choose the following: the most recent year (2014), the most common business type category (retail non-food), the average number of picksups (five) and the largest transfer station.

Figure 21: The bidding functions for customers of different type

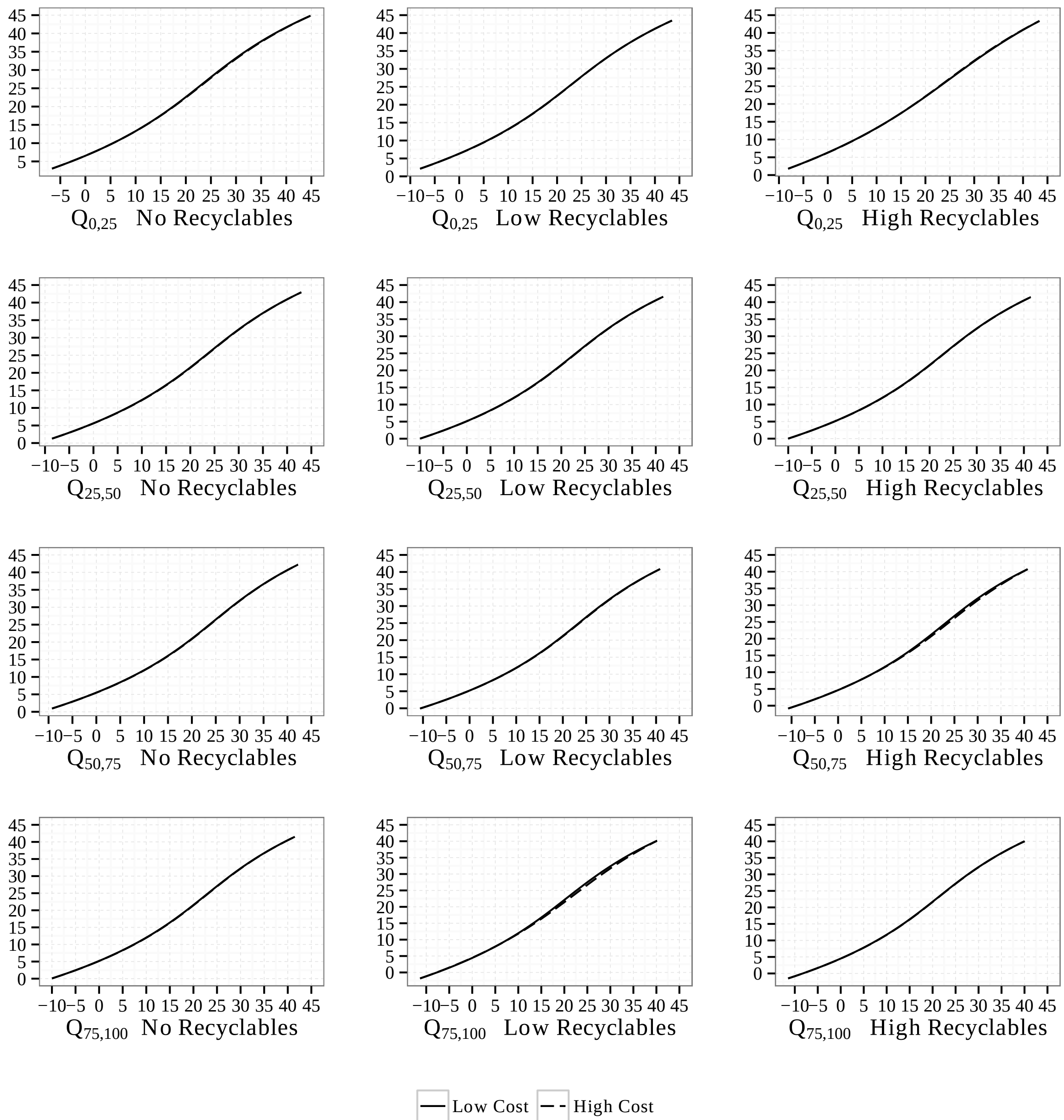
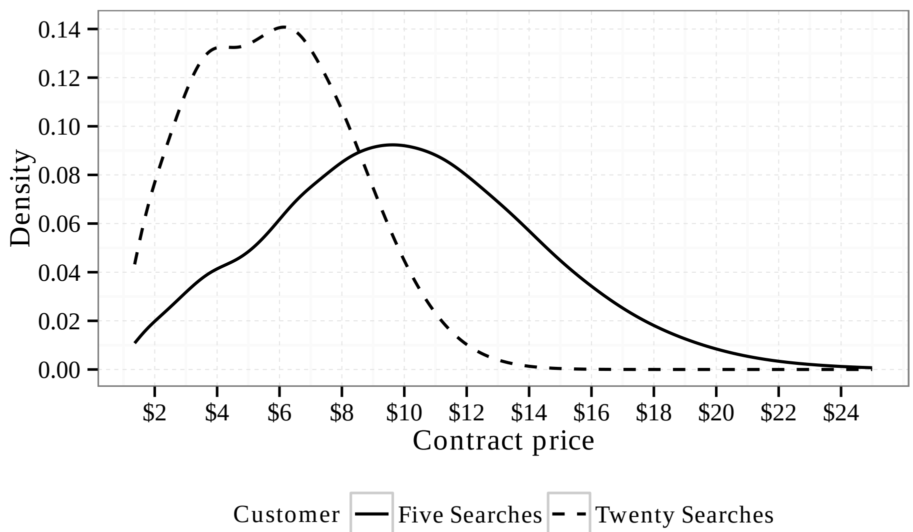


Figure 22: Price Offer Distribution for two different Customers



Notes: This figure shows the price distribution for customers in the category $Q_{25,50}$ without recyclables, for a customer who makes five searches compared to someone who makes twenty searches.