KU LEUVEN
CENTER FOR ECONOMIC STUDIES

DISCUSSION PAPER SERIES
DPS15.32

DECEMBER 2015

Intermediation in markets with buyer and seller selection: theory and an application to real estate brokerage

Bert WILLEKENS, Roel HELGERS, Maarten GOOS and Erik BUYST

Department of Economics

Faculty of Economics
And Business

Intermediation in Markets with Buyer and Seller Selection: Theory and an Application to Real Estate Brokerage*

Bert Willekens Roel Helgers Maarten Goos Erik Buyst

December 16, 2015

Abstract

Many markets are one-to-one matching markets in which match-making intermediaries enable pairs of buyers and sellers to negotiate a transaction price for a good or service. Examples are real estate markets in which realtors search for matches between potential home buyers and sellers, or labor markets in which employment services help the unemployed find work by filling firms' vacancies. This paper investigates the socially optimal size and market structure of such markets. Firstly, it shows that it is socially optimal for intermediaries to have some market power. That is, it is socially optimal that intermediaries charge a service fee that is above the per-match cost as this excludes some low valuation buyers, which are disliked by sellers, as well as some high reservation price sellers, which are disliked by buyers. Secondly, it shows that matching markets are generally characterized by an excessive number of intermediaries that operate in the market compared to what is socially optimal. When calibrating the model using data from the Belgian real estate brokerage industry, the welfare counterfactuals suggest that the observed average commission rate of 4.3% is below the socially optimal commission rate, which is estimated to be in the range between 5.1% and 24%. A welfare gain of 1% to 11% could be established when regulating broker service fees, given the number of brokers that currently operate in the market. When also regulating broker entry, a further welfare gain of 7% to 69% could be realized. Various other policy relevant welfare counterfactuals are constructed and discussed.

* Buyst, Helgers and Willekens: University of Leuven. Goos: Utrecht University. Contact: bert.willekens@kuleuven.be. We thank Jon Levin, Andras Niedermayer, Jo Van Biesebroeck, Patrick Van Cayseele, Glen Weyl and seminar and conference participants at the University of Leuven, Bocconi University (EARIE 2014), University of Alicante (ENHR 2014) and University of Reading (AREUEA 2014) for useful comments and suggestions. Willekens acknowledges financial support from the Research Foundation Flanders.

1 Introduction

Many one-to-one matching markets are characterized by a dual structure. On the one hand, there is an industry of intermediaries who guide the matching process between market participants and they usually charge a significant fee in return for this service. On the other hand, there is an outside market in which market participants independently search for a trading partner and incur their own search costs. For example, in real estate markets there is typically a real estate brokerage industry and a "for-sale-by-owner" market. In labor markets, firms can utilize the services of a recruitment agency or can internally organize the process of hiring a new employee. In e-commerce, sellers of a second-hand good can choose to post a listing on a free classified ads website, like Craigslist, or can use a centralized trading platform, like eBay or Amazon, in return for a service fee. In all of these examples, participants on both sides of the market self-select into one of the two market segments, not only based on their own preferences, but also based on their expectations about which type of trading partner they will eventually meet. So, the participation decision of each market participant entails an externality for the participants on the other side of the market and it therefore matters from a social point of view how different types of participants are selected into the market.

This paper investigates what the socially optimal size and market structure is of an industry of private intermediaries, taking into account that the service fee charged by the intermediaries influences which types of participants are selected into the intermediary market. It is important to address this question, given that the incentives of private intermediaries that maximize profits are not necessarily aligned with those of a social planner. Not only because the intermediaries might possess market power, which can certainly be the case due to the informational advantages they inherently have over the market participants, but also because they might not properly internalize the externalities present in the market when pricing their services. To address the issue, a general model of imperfect competition among intermediaries that operate in a one-to-one matching market is presented. Subsequently, the model is empirically applied to the case of real estate brokerage. Real estate is a particularly interesting application because different types of buyers and sellers typically decide to hire the services of a real estate broker compared to those who trade in the for-sale-by-owner market.[1]

In the model, interactions among market participants – referred to as buyers and sellers – and the intermediaries that operate in the market – referred to as brokers – occur in four stages.[2] In the first stage, brokers can freely enter the market as long as they expect it is profitable to do so. In the second

[1] Hendel, Nevo and Ortalo-Magné (2009), for example, provide evidence that less patient sellers and more patient buyers (to avoid patient sellers) tend to trade in the brokerage market compared to those who trade in the for-sale-by-owner market.

[2] Equivalently, for a labor market one can think of the sellers as workers, of the buyers as firms and of the brokers as recruiters. The price of the traded "good" is then the wage of the worker.

stage, the brokers that entered the market imperfectly compete to attract buyers and sellers by announcing their service fee charged to either buyers or sellers. The service fee possibly consists of a flat fee and a fee proportional to the price of the traded good and is only paid conditional on a successful transaction. In the third stage, buyers and sellers enter the intermediary market when their expected utility of participating is greater than when participating through the outside market. In the final stage, buyers and sellers that participate in the intermediary market are randomly assigned to one another and the sales price of the traded good is determined by a Nash bargain between the buyer and the seller. The assumptions in the last stage are imposed to capture the intuition that market participants care about the characteristics of their trading partner – the price a seller receives and a buyer pays depends on the reservation value of their trading partner – and that due to information imperfections there is uncertainty about which trading partner they will meet when deciding on market participation.

A first important result derived from this setting is that a social planner always charges an intermediary service fee above the per-match cost of serving buyers and sellers. The planner internalizes the externality that buyers dislike high reservation price sellers and sellers dislike low valuation buyers, which are excluded from the intermediary market by charging a relatively high service fee. It follows that some market power attributed to intermediaries is justified when they compete in a private market. The monopoly (or collusive) service fee, however, always exceeds the socially optimal fee. So, there exists an inverse u-shaped relationship between private broker market power and social value created by the intermediary industry. A second finding is that the private market outcome is generally characterized by an excessive number of intermediaries that operate in the market compared to what is socially optimal. That is, when a novel entrant steals business from incumbent brokers, valuable resources are wasted by brokers inefficiently competing to realize the same number of transactions that could also be established with fewer brokers operating in the market. Furthermore, this entry distortion is more severe when brokers possess more market power in pricing their services. So, combined with the result that some broker market power is justified to properly internalize the participation externalities of buyers and sellers, the model has nonstandard policy implications.[3]

In comparison to the private free entry equilibrium, the welfare effects are derived when a social planner optimally regulates broker service fees, broker entry or both. When regulating both, all market distortions can be eliminated by setting the service fee such that the participation externalities of

[3] Note that it is important to also account for the entry distortion on top of the distortion in the service fee, as there are typically little barriers for novel intermediaries to enter and operate in matching markets. Hsieh and Moretti (2003) and Barwick and Pathak (2015), for example, provide evidence that the entry distortion is quantitatively important in the US real estate brokerage industry.

buyers and sellers are properly internalized and by minimizing the number of brokers that operate in the market. However, the welfare gains and redistributive effects of either regulating broker service fees or broker entry are ambiguous. They depend on the underlying parameters of seller supply and buyer demand and the structure of broker costs. The model outcomes are therefore further illustrated for realistic calibrated parameter values using data from the Belgian real estate brokerage industry.[4]

The empirical results suggest that the observed average commission rate of 4.3% charged by brokers is below the socially optimal commission rate, which ranges from 5.1% to 24% for the estimated range of feasible values for the parameters of seller supply and buyer demand. This implies that the externalities present in the market are insufficiently internalized and it would be welfare improving to exclude more buyers and sellers. For the most inelastic bound on estimated supply and demand elasticities, the welfare counterfactuals suggest that a welfare gain of 19% could be established when regulating both service fees and market entry of brokers. The outcome of a social planner that regulates broker entry and allows brokers to privately compete in pricing their services is calculated to generate a welfare gain of 18%, while regulating service fees and allowing for free broker entry only implies a welfare gain of 5%. For the most elastic bound on the supply and demand elasticities, however, regulating both service fees and entry implies a welfare gain of 71%, only regulating entry results in a gain of 40% and only regulating service fees in a gain of 52%. Importantly, welfare gains for both the inelastic and elastic bounds are gains in consumer surplus attributed to buyers and sellers when regulating broker service fees. Regulating broker entry, however, always implies a loss in consumer surplus, which is compensated by a gain in broker profits.

In the literature, there is an extensive strand of research that investigates the role of intermediaries to facilitate market transactions – see, for example, Rubinstein and Wolinsky (1987), Biglaiser (1993) and Yavas (1994) for seminal contributions. Spulber (1999) provides a unified perspective on the different views on firm intermediation in the early literature, in which intermediaries usually play the role of market clearing entities and the externalities induced by the participation decision of different buyer and seller types emphasized here play no role. More recently, Niedermayer and Shneyerov (2014) and Loertscher and Niedermayer (2015) explore how an optimal market clearing mechanism can be implemented by intermediaries that charge a service fee instead of directly setting bid-ask spreads.

[4] The institutional setting of the Belgian real estate market is particularly interesting to apply the model because there are no significant barriers for new brokers to enter the market and there are no institutional restrictions for brokers to compete in pricing their services. It is therefore sensible to construct welfare counterfactuals in which both the pricing and entry behavior of brokers is affected by policy interventions. This contrasts with the US, for example, where commission rates charged by real estate brokers typically show little variation, which suggests a lack of competition among brokers in pricing their services (e.g. Hsieh and Moretti 2003).

Evaluating the optimal pricing behavior of platform businesses in the presence externalities across different groups of market participants has been the topic of interest in the so-called two-sided markets literature – e.g. Rochet and Tirole (2003), Armstrong (2006) and Weyl (2010). In this literature it is typically assumed, however, that only the size and not the composition of one group of market participants affects utility of another group. In our setting, it is precisely the changed composition of user types when more or less buyers and sellers participate in the intermediary market that drives the results. Damiano and Li (2007, 2008) analyze a similar composition effect. Although their setup is quite different from ours – e.g. they allow for complementarities in the match value function and analyze duopoly competition among endogenously differentiated platforms – Damiano and Li (2008) establish a similar result that the market outcome under duopoly can be less efficient than the monopoly outcome. The basic intuition for this result is the same as in our setting. Since market participants care about the quality of their trading partner, it can be socially efficient to exclude some participants from the market. Gomes and Pavan (2015) build on Damiano and Li (2007) to investigate optimal matching mechanisms in many-to-many matching settings.

Our work also relates to recent research on competition among service providers in markets where consumer selection plays an important role, like insurance and credit markets – e.g. Einav and Finkelstein (2011), Veiga and Weyl (2015) and Mahoney and Weyl (2015). In particular, Mahoney and Weyl (2015) demonstrate that an inverse u-shaped relationship exists between competition and welfare in markets characterized by advantageous selection. The matching markets we study can also be interpreted as being characterized by advantageous selection in the sense that buyers and sellers with a high willingness to pay for the brokerage service are assumed to be the ones that bring most value to the market through the Nash bargain. The crucial difference with insurance or credit markets, however, is that selection occurs through an externality across the market: buyers care which sellers are selected into the intermediary market and vice versa. The graphical approach to present the results is also inspired by Einav and Finkelstein (2011) and Mahoney and Weyl (2015).

Finally, there is vast body of research that investigates the inefficiencies in the US real estate brokerage industry that can be attributed to a lack of price competition among brokers. Seminal theoretical contributions that point out conditions under which fixed commission rates can be socially harmful are Yinger (1981), Crockett (1982) and Miceli (1992). Hsieh and Moretti (2003) provide supporting empirical evidence of significant social waste in the US brokerage industry due to excessive broker entry. Other recent contributions that structurally aim to quantify the entry distortions are Han and Hong (2011) and Barwick and Pathak (2015).[5] In the present paper, the entry distortion is

[5] In addition, there are several papers that investigate the question whether fixed commission rates could be the result of a competitive market outcome or are more likely to arise from (tacit) collusion among brokers, usually

evaluated when brokers do compete in pricing their services. It is particularly relevant to address this question today, given that the adoption of new information technologies seems to have intensified price competition among intermediaries, not only in real estate brokerage (e.g. USDOJ and FTC report 2007), but also for many other intermediate service providers, like travel agencies and stock brokers, as pointed out by Levitt and Syverson (2008a).

The remainder of this paper is organized as follows. Section 2 present the theoretical model and results. Section 3 proposes the methodology to empirically implement the model. Section 4 describes the data and the institutional setting of the Belgian real estate brokerage industry. Section 5 presents the results for the model calibration and welfare counterfactuals. The final section concludes.

2 Model

Consider a four-stage static model of symmetric imperfect competition among brokers who offer a service of matching buyers and sellers in a market for a homogeneous good. The implications of allowing for heterogeneous product characteristics are discussed in the next section when the methodology to implement the model empirically is introduced. The timing of the model can be summarized as follows:

Stage 1: N brokers (out of an unrestricted amount) enter the market.

Stage 2: Participating brokers simultaneously announce the brokerage fees charged to sellers and buyers in return for their service.

Stage 3: N^S sellers and N^B buyers (out of a potential mass S) enter the market through one of the brokers.

Stage 4: M transactions occur through the brokerage industry and the broker service fees are paid.

Assume that brokers, sellers and buyers are risk-neutral and that sellers and buyers have unit supply and demand, respectively. Sellers are heterogeneous in their reservation price of providing the good to the market through one of the brokers, denoted by s and assumed smoothly distributed by $F^S(.)$ with density $f^S(.)$ on $[s^L, s^H]$ with $s^H > s^L$. Similarly, buyers are heterogeneous in their valuation of purchasing the good in the brokerage market, denoted by b and assumed smoothly distributed by $F^B(.)$ with density $f^B(.)$ on $[b^L, b^H]$ with $b^H > b^L$. The outside option of not participating in the market through one of the brokers is normalized to zero for both sellers and buyers. This normalization implies that one can think of s as the common (opportunity) cost of sellers of providing

from a principle-agent perspective. Examples are Carroll (1989), Anglin and Arnott (1999), Yavas (2001), Miceli, Pancak and Sirmans (2007), Levitt and Syverson (2008a) and Fisher and Yavas (2010).

the good to the market, either through the brokerage market or the outside market, subtracted by the gain of each seller in search costs when hiring a broker. So, sellers with a *low* reservation price s are assumed to be the ones who gain relatively most from the brokerage service. Similarly, one can think of b as the common valuation for the good of buyers added by the individual gain in search cost when purchasing the good through one of the brokers compared to searching for the good through the outside market. Buyers with a *high* value of b are thus the ones that gain relatively most from the brokerage service.

Assume that the distributions of seller reservation prices and buyer valuations for the good are public information. Individual seller and buyers types, however, are *ex ante* private information, when sellers and buyers decide upon market participation (stage 3), and they become revealed *ex post* once a buyer is matched to a seller (stage 4). The remainder of this section recursively specifies the occurrence of events and reports the resulting outcomes for each stage of the model.

2.1 Individual transaction valuations (stage 4)

When sellers participate in the market by hiring a broker they are charged a fee that only has to be paid conditional on the good being sold by the hired broker. The fee possibly consists of a flat component T and a percentage fee t charged proportional to the sales price of the good. The individual transaction value of a seller type s can hence be written as:

$$(1-t)p - s - T \tag{1}$$

where p denotes the transaction price. Buyers are not directly charged for the broker service and the individual transaction value of a buyer type b can therefore be written as:

$$b - p \tag{2}$$

The fee charged to the seller, however, can (partially) be passed through in the bargain over the sales price between the buyer and the seller. More specifically, assume the transaction price is chosen to maximize an asymmetric Nash bargain:

$$\max_p (b-p)^{1-\beta}\big((1-t)p - s - T\big)^{\beta} \tag{3}$$

where $\beta \in [0,1]$ denotes the bargaining weight of sellers and $1-\beta$ is the bargaining weight of buyers.[6] This yields the following expression for the transaction price:

[6] Note that in real estate markets, the broker, rather than the seller, usually bargains over the transaction price with potential buyers (or buyer-brokers). However, a seller-broker (buyer-broker) contract typically also explicitly specifies that the broker should represent the best interest of the seller (buyer) in this process, which is assumed to be case here. More generally, this paper ignores any potential principle-agent problems concerning

$$p(b, s) = \beta b + (1 - \beta) \frac{T + s}{1 - t} \tag{4}$$

Nash bargaining implies that the transaction price at which the good is sold is match-specific and depends on the valuation of the buyer and the reservation price of the seller that are being matched. The homogeneous good is therefore allowed to be sold at dispersed prices, rather than being determined by a competitive market clearing mechanism, which would imply a single market price. This is consistent with the arguments of Stigler (1961) that price dispersion is inherent to markets with imperfect information and costly search, of which matching markets are a primary example. Baye, Morgan, and Scholten (2007) provide a further discussion on the determinants of price dispersion in markets with imperfect information. [7]

2.2 Buyer and seller participation (stage 3)

Assume that the service offered by brokers is perceived as differentiated across buyers and across sellers, for example, by the different locations of the brokers. Service differentiation is restricted, however, by the assumption that in equilibrium a symmetric and representative set of buyers and sellers is attracted by each broker. More specifically, market supply of sellers is equal to $N^S = \sum_{i=1}^{N} n_i^S$ where n_i^S is the number of sellers attracted by broker i, which is assumed to be the same across brokers: $n_1^S = \ldots = n_N^S = N^S/N \equiv n^S$. Similarly, market demand for buyers is equal to $N^B = \sum_{i=1}^{N} n_i^B$ where n_i^B is the number of buyers attracted by broker i, again assuming symmetry across brokers: $n_1^B = \ldots = n_N^B = N^B/N \equiv n^B$.

In addition, assume that the matching technology offered by the brokers is efficient and random. That is, the number of matches established by every broker is equal to $\min[n^B, n^S]$, the match probability of sellers is $\min[n^B, n^S]/n^S \equiv m^S$ and the match probability of buyers is $\min[n^B, n^S]/n^B \equiv m^B$. By broker symmetry, it follows that the equilibrium number of matches that occur through the brokerage market is equal to $M = \min[N^B, N^S]$. In what follows, broker subscripts i are omitted to minimize the notational burden.

Expected seller and buyer utility of participating through the brokerage market can be written as:

$$u^S = \Big((1 - t)p(\bar{b}, s) - s - T\Big)m^S = \beta\Big((1 - t)\bar{b} - s - T\Big)m^S \tag{5}$$

the seller-broker or buyer-broker relationship, as investigated, for example, by Rutherford, Springer and Yavas (2005) and Levitt and Syverson (2008b).

[7] Finally, it should also be noted that the specific assumption that only sellers are charged for the brokerage service and that buyers are not directly charged is imposed because this fee structure is observed in the Belgian real estate brokerage market analyzed in the empirical part of this paper. All the derived results, however, also carry through when only buyers are directly charged, or when both sellers and buyers are charged part of the fee, as formalized in Appendix A1.

$$u^B = (b - p(b,\bar{s}))m^B = (1-\beta)\left(b - \frac{T+\bar{s}}{1-t}\right)m^B \tag{6}$$

where $\bar{b}$ denotes the expected buyer valuation for the good and $\bar{s}$ the expected seller reservation price, respectively:

$$\bar{b} = \frac{S}{N^B}\int_0^{N^B/S} F^{B^{-1}}(1-x)dx \tag{7}$$

$$\bar{s} = \frac{S}{N^S}\int_0^{N^S/S} F^{S^{-1}}(x)dx \tag{8}$$

Sellers participate when $u^S \geq 0 \Leftrightarrow s \leq (1-t)\bar{b} - T \equiv \tilde{s}$, where $\tilde{s}$ denotes the reservation price of the marginal seller that participates through the brokerage market. Similarly, buyers participate when $u^B \geq 0 \Leftrightarrow b \geq \frac{T+\bar{s}}{1-t} \equiv \tilde{b}$, where $\tilde{b}$ denotes the marginal buyer valuation. Market supply of sellers and market demand for buyers can thus be summarized as:

$$N^S = SF^S(\tilde{s}) = SF^S\big((1-t)\bar{b} - T\big) \tag{9}$$

$$N^B = S\left(1 - F^B(\tilde{b})\right) = S\left(1 - F^B\left(\frac{T+\bar{s}}{1-t}\right)\right) \tag{10}$$

Expression (9) shows that market supply of sellers depends negatively on the service fees T and t charged by brokers, as one would expect. Specific to our setting, however, is that seller supply also depends positively on the expected buyer valuation $\bar{b}$. All else equal, when sellers expect that buyers with a higher valuation participate in the market, more sellers participate because they expect to receive a higher price for their property. This in turn implies that seller supply is characterized by a negative externality induced by the participation decision of buyers. As illustrated by expression (7), the expected buyer valuation depends negatively on the number of buyers that participate. This is because the marginal buyer always has a lower valuation for the good than inframarginal buyers and hence the participation of this marginal buyer drives down the average valuation of all the buyers that participate in the market. Similarly, expression (10) shows that market demand for buyers depends negatively on the service fees. In addition, it is characterized by a negative externality induced by the participation decision of sellers, through the expected seller reservation price $\bar{s}$. Low reservation price sellers enter the market first and hence increased seller participation raises the sales price buyers expect to pay, which in turn reduces buyer demand.[8]

[8] Note that there is another channel through which externalities can result from the participation decision of users on either side. As is clear from the expressions (5) and (6) for expected seller and buyer utilities, respectively, the match probabilities of users on both sides, $m^S = \min[N^B, N^S]/N^S$ and $m^B = \min[N^B, N^S]/N^B$, also depend on the participation decision of users on both sides. The assumption that the matching technology is

Using expressions (9) and (10) and the definitions of $\tilde{s}$ and $\tilde{b}$ allows us to write the market clearing flat fee T and proportional fee t as a function of marginal and average preference values of sellers and buyers:

$$T = \frac{\tilde{b}\tilde{s} - \bar{b}\bar{s}}{\bar{b} - \tilde{b}} \tag{11}$$

$$1 - t = \frac{\tilde{s} - \bar{s}}{\bar{b} - \tilde{b}} \tag{12}$$

Expressions (11) and (12) can be interpreted as a system of inverse demand equations, in which $\tilde{b} = F^{B^{-1}}(1 - N^B/S)$ and $\tilde{s} = F^{S^{-1}}(N^S/S)$, as follows from (9) and (10), and $\bar{b}$ and $\bar{s}$ are given by expressions (7) and (8). In what follows, it is assumed that any equilibrium market allocation N^B, N^S is uniquely established through the two market clearing values of the pricing instruments T and t that follow from (11) and (12).[9]

2.3 Imperfect broker competition (stage 2)

2.3.1 Broker profits and welfare

Expected broker profits can be written as:

$$\pi = (AR - MC)\min[n^B, n^S] - FC \tag{13}$$

where $MC \geq 0$ denotes a constant per-match cost incurred by each broker when matching buyers and sellers and $FC \geq 0$ denotes a fixed cost incurred by each broker to operate in the market, independent of the number of matches established. AR is defined as the expected or average per-match revenue:

$$AR \equiv T + t\bar{p} \tag{14}$$

efficient, however, will imply that profit-maximizing brokers always balance the market by attracting the same number of buyers and sellers. This in turn implies that the match probabilities of both buyers and sellers are equal to 1 in equilibrium and that these additional externalities play no role. See, Goos, Van Cayseele and Willekens (2014) for a more general treatment of the implications of matching frictions on the optimal pricing behavior of platform businesses.

[9] In other words, we assume that brokers can always resolve the coordination problem they face to attract two distinct user groups in the presence of indirect network externalities. This coordination problem is well-known from the two-sided markets literature and various solutions were proposed, for example, by Caillaud and Jullien (2003), Weyl (2010) and White and Weyl (2015). We do not explicitly deal with the issue here given that it is precisely an important part of the “job” of brokers in matching markets to resolve the coordination problem. Brokers can credibly commit to sellers to search for a buyer to their best effort, given that payments to the broker only occur when a transaction is actually established. By this logic, the coordination problem is less of an issue in markets where intermediaries are involved in the trading process between participants and charge conditional payments compared to the classic two-sided market examples where the platform has no direct control over the interactions between attracted user groups, like payment card networks or newspapers.

in which $\bar{p} \equiv p(\bar{b}, \bar{s})$ denotes the expected transaction price of transactions that occur through the brokerage market, which by the symmetry assumptions is the same across brokers. Using expressions (4), (11) and (12), the average transaction price can be written as:

$$\bar{p} = \beta\bar{b} + (1-\beta)\tilde{b} \tag{15}$$

In addition, using expressions (11), (12) and (15), average per-match revenue can be written as a function of marginal and average user types on both sides of the market:

$$AR = \beta\left(\bar{b} - \tilde{s}\right) + (1-\beta)\left(\tilde{b} - \bar{s}\right) \tag{16}$$

Given that the marginal and average buyer types ($\tilde{b}$ and $\bar{b}$, respectively) are strictly decreasing in the number of buyers attracted into the brokerage market and the marginal and average sellers types ($\tilde{s}$ and $\bar{s}$, respectively) are strictly increasing in the number of sellers, expression (16) implies that expected per-match revenue is strictly decreasing in both the number of buyers and sellers that participate in the market by hiring a broker.

To model symmetric imperfect competition among brokers in providing their service to the market, we follow the approach of Weyl and Fabinger (2013) by which the degree of imperfect competition can be captured by a single "conduct parameter", extended to our setting where competing brokers attract and efficiently match two distinct user groups. To do so, the assumption is imposed that strategic interactions among brokers are restricted such that average per-match revenue for individual brokers is strictly decreasing in the number of users attracted on both sides of the market, i.e. $dAR/dn^I < 0$ for $I = B, S$, which is the equivalent to assuming that firms face downward sloping individual demand curves.

In the current setting, this implies that individual broker profits, given by expression (13), are strictly decreasing in the number of users on one side of the market if the attracted number of users on that side exceeds the number of users attracted on the other side, i.e. $d\pi/dn^I < 0$ if $n^I > n^J$ for $I \neq J$. This in turn implies that any profit maximizing equilibrium must always be balanced, i.e. $n^S = n^B = n$ or, equivalently, $N^S = N^B = M$. If not, brokers can always raise profits by lowering the number of users on the long side of the market. This result directly follows from the assumption that the matching technology available to brokers is efficient and it conveniently allows us to convert the problem of brokers competing to attract users on two distinct sides into a problem where the brokers compete in a single quantity (n) by using one of the available pricing instruments (e.g. T). The other available pricing instrument (e.g. t) is simply adjusted to ensure the balanced market condition holds and therefore no longer needs to be considered as a strategic decision variable.

Following Weyl and Fabinger (2013), instead of explicitly modeling the interactions among competing brokers, it is assumed that in any imperfectly competitive equilibrium the elasticity-adjusted Lerner index is set equal to a conduct parameter θ, which in our model satisfies:

$$\frac{AR - MC}{AR}\left(-\frac{dM}{dAR}\frac{AR}{M}\right) = \theta \tag{17}$$

where $\theta \in [0,1]$ when the broker services are substitutes, which is assumed to be the case. As formalized by Weyl and Fabinger (2013), this framework nests a broad range of imperfect competition models, among which monopoly or cartel ($\theta = 1$); Bertrand ($\theta = 0$); Cournot ($\theta = 1/N$); Bresnahan (1989)'s constant conjectural variations model ($\theta = (1 + R)/N$ where $dM/dn = 1 + R$); and symmetrically differentiated Nash-in-prices and monopolistic competition (for which θ is not a constant). However, we do not derive explicit conditions for these models for our setup, given that none of the results hinge on the specific underlying model of imperfect competition. The only thing that matters here is that any outcome on the continuum between monopoly and Bertrand is a feasible imperfect competition equilibrium.

To evaluate market efficiency, the private market equilibrium is compared to the outcome determined by a Pigouvian planner that optimally chooses the number sellers N^S and buyers N^B attracted in the brokerage industry to maximize total social value, taking the number of brokers that operate in the market as given. Total social value generated in the market is equal to the sum of total industry profits $\Pi \equiv \pi N$ and total consumer surplus CS, defined as the sum of total buyer and seller surplus, which can be written as:

$$CS = \left(\beta(\tilde{s} - \bar{s}) + (1 - \beta)(\bar{b} - \tilde{b})\right)\min[N^B, N^S] \tag{18}$$

By combining equations (13), (16) and (18), total social value W simplifies to:

$$W = (\bar{b} - \bar{s} - MC)\min[N^B, N^S] - FC\,N \tag{19}$$

Given that $\bar{b} - \bar{s}$ is strictly decreasing in N^B and N^S, the Pigouvian planner always balances the market, i.e. $N^B = N^S = M$, because welfare is strictly decreasing in participation on the long side of the market. This again conveniently allows us to simplify the social optimization problem to a problem with a single decision variable, in this case M.

2.3.2 Private market outcome

Proposition 1 summarizes the private market equilibrium when an exogenous number N of symmetric brokers operate the market. The result follows from equating N^S and N^B to M in expression (16) for average per-match revenue, differentiating with respect to M and substituting the solution in the imperfect competition equation (17).

Proposition 1 *Optimal private broker behavior implies that the equilibrium number of matches M established through the brokerage market satisfies:*

$$AR - MC = \theta(MS + ET) \tag{20}$$

where MS denotes marginal consumer surplus, defined as dCS/dM, which can be written as:

$$MS = \beta \frac{F^S(\tilde{s})}{f^S(\tilde{s})} + (1-\beta)\frac{1 - F^B(\tilde{b})}{f^B(\tilde{b})} \tag{21}$$

and ET refers to an "externality tax", raised to internalize the cross-side participation externalities in buyer demand and seller supply, which can be written as:

$$ET = \beta(\bar{b} - \tilde{b}) + (1-\beta)(\tilde{s} - \bar{s}) \tag{22}$$

Expression (20) shows that the markup of average per-match revenue over per-match cost is increasing in the conduct parameter θ, ranging from zero under Bertrand competition ($\theta = 0$) to $MS + ET$, which is the monopoly markup ($\theta = 1$). The first term, MS, denotes marginal consumer surplus, which in a standard monopoly model is equal to the inverse hazard rate (or semi-elasticity) of demand and coincides with the classic Cournot distortion. In the present setting, MS is equal to the weighted sum of inverse hazard rates of seller supply and buyer demand, where the weights are equal to the bargaining weight of users on these respective sides. This is intuitive: if one side possesses no bargaining power in determining sales prices, users on that side capture no surplus from transactions and hence no surplus can be extracted by brokers from that side, independent of the elasticity of demand or supply. The second term, ET, refers to an externality tax raised by brokers to internalize the negative cross-side externalities present in the market. That is, brokers want to avoid attracting too many buyers because more buyers imply a lower average buyer valuation for the good, which in turn is disliked by the sellers because they expect to receive a lower price for their properties. Similarly, too many sellers imply a high average reservation price of sellers, which is disliked by buyers because they expect to pay a higher price for the good. To account for this, brokers charge a higher markup than they would without externalities.

From expression (22) it is clear that the magnitude of ET depends on the spread between average and marginal user types on both sides of the market or, in other words, on the degree of heterogeneity in user types. For example, when buyers are homogeneous in their valuation, sellers are indifferent to which buyer they will be matched and the participation decision of the marginal buyer causes no externalities. In this case, $\bar{b} = \tilde{b}$ and the first term in ET disappears because there is no externality for brokers to internalize on the buyer side. In contrast, when dispersion in buyer valuations is large, the spread between the marginal and average buyer valuation will be large and that marginal buyer entails a large externality. The tax raised to internalize this externality is precisely the spread between the average and marginal buyer valuation, weighted by the bargaining strength of sellers. Similarly, the tax to internalize the externality on the seller side is equal to spread between the marginal and average seller (where the former has a higher reservation price than the latter which is disliked by buyers), weighted by the measure of buyer bargaining power.

To sum up, proposition 1 demonstrates that under Bertrand competition ($\theta = 0$) the markup of average per-match revenue over per-match cost is equal to zero, whereas under monopoly pricing ($\theta = 1$) it is equated to a weighted version of the classic Cournot distortion plus a tax imposed to internalize the negative cross-side externalities present in the market. Depending on the degree of competition among brokers in pricing their services, any markup in between these two bounds is a feasible private market outcome. To evaluate the distortions that might arise from private broker behavior, we now turn to the socially optimal market outcome.

2.3.3 Socially optimal outcome

Proposition 2 summarizes the social optimum chosen by a Pigouvian planner. The result follows from equating N^S and N^B to M in expression (19) for total social value and rewriting the first-order condition with respect to M. The socially optimal degree of broker competition is derived from equating the private and social first-order conditions.

Proposition 2 *At the first-best social optimum, the equilibrium number of matches M^* established through the brokerage market satisfies:*

$$AR - MC = ET \tag{23}$$

This implies that the socially optimal degree of competition among brokers in a private market satisfies:

$$\theta^* = \frac{ET}{MS + ET} \tag{24}$$

Expression (23) demonstrates that a Pigouvian planner also internalizes the selection effect by taxing the negative externalities induced by the participation decision of users on both sides. Furthermore, it does so exactly to the same extent a monopolist does in the private market. The externality tax is strictly positive in the presence of heterogeneity in buyer and/or seller types, which implies that Bertrand competition among brokers ($AR = MC$) is not socially optimal. In this case, broker fees are too low and the equilibrium number of matches too high compared to the social optimum because the participation externalities present in the market are not properly internalized. On the other hand, the monopoly outcome can never be efficient because, on top of the externality tax, broker fees are marked up by the weighted Cournot distortion, which results in upward distorted broker fees and hence insufficient participation of buyers and sellers. So, in a private market there exists an intermediate degree of imperfect competition θ^* which establishes the first-best social optimum. Expression (24) shows that θ^* depends on the magnitude of MS relative to ET. When marginal consumer surplus (the Cournot distortion) is small relative to the externality tax, the desired degree of market power is large and vice versa. Which of both measures is largest depends on the underlying distributions of user types and relative bargaining weights, as is clear from expressions (21) and (22).

To further illustrate the intuition of propositions 1 and 2, **Figure 1** graphically summarizes the results for linear buyer demand and seller supply. The number of transactions that occur through the brokerage market (M) are on the horizontal axis and broker revenues and costs are on the vertical axis. The AR curve in the figure illustrates that the expected per-match revenue of brokers decreases in the number of transactions that occur in the brokerage market. The marginal revenue curve, given by $MR = AR - MS - ET$, always lies below the average revenue curve. Bertrand equilibrium is characterized by the point where the AR curve crosses the constant marginal cost curve and monopoly (or cartel) equilibrium by the point where the marginal revenue curve crosses the marginal cost curve. As formalized in proposition 1, depending on brokers' market power measured by the conduct parameter θ, private market equilibrium lies somewhere on the continuum in between the monopoly and Bertrand outcome. The social optimum is established at the point where the average revenue curve crosses the upward sloping social cost curve. The social cost of attracting buyers and sellers is equal to marginal cost plus the tax that should be raised to internalize the participation externalities of buyers and sellers: $SC = MC + ET$. In the presence of heterogeneity in buyer and seller types, the social optimum on the average revenue curve always lies in between the Bertrand and monopoly outcomes. So, there exists an intermediate degree of broker competition θ^* for which the incentives of the social planner and the private brokers are aligned, as formalized in proposition 2.

2.4 Free broker entry (stage 1)

2.4.1 Free entry equilibrium

In the first stage of the model brokers can freely enter the market and they will do so as long as profits of the marginal entrant are weakly positive. Ignoring the integer constraint on the number of brokers, this implies that in a free entry equilibrium individual broker profits must be equal to zero:

$$\pi = (AR - MC)\frac{M}{N} - FC = 0 \tag{25}$$

The number of brokers that enter the market depends on the markup they expect to receive in the second stage, given by expression (20). For example, when brokers collude on charging the monopoly service fee, expected per-match revenue (AR) and hence the number of transactions that occur through the brokerage industry (M) are independent of the number of brokers that enter the market. In this case, N is equal to $(AR - MC)M/FC$. At the other extreme, equation (25) shows that Bertrand equilibrium ($AR = MC$) is not feasible in the presence of a positive fixed cost. More generally, when market power in the second stage is sufficiently large to cover the fixed cost of at least one entrant, the number of brokers that operate the market follows from the zero-profit condition (25), where AR and M depend on N through the private first-order condition (20). In what follows, the free entry equilibrium number of brokers that operate the market is denoted as N^{FE}.[10] In the free entry equilibrium the average per-match revenue earned by brokers is always equal to the average per-match cost, $AR = AC$, where $AC = MC + FC/(M/N)$, as follows from rewriting expression (25). Greater market power in the second stage induces more brokers to enter the market, such that the number of transactions per broker (M/N) falls and hence the average cost incurred by each broker increases.

Figure 2 graphically illustrates the free entry equilibrium, which is characterized by the crossing of the average revenue curve and the average cost curve. Two cases are drawn. The average cost curve AC_1 crosses the average revenue curve above the social optimum and hence the average service fee is too high and too few transactions occur through the brokerage market compared to what is socially optimal. This case is more likely to occur when either fixed costs are high or when broker entry is high because brokers possess market power in setting their service fees (in stage 2), or both. In the extreme case where fixed operating costs are such that only one broker can enter the market, it will set the monopoly service fee and the AC curve will cross the AR curve at the monopoly outcome. Alternatively, when brokers collude to charge the monopoly service fee, the AC curve will also cross the AR curve at the monopoly outcome, even when fixed costs are relatively small. Many brokers will

[10] Following Mankiw and Whinston (1986), the free entry equilibrium is unique when assumptions (a), (b) and (c) specified in proposition 3 below are satisfied.

enter the market, which pushes up the AC curve, and every broker will only carry out a few but highly profitable transactions. In the second case, the average cost curve AC_2 crosses the average revenue curve below the social optimum, the average service fee is too low and too many transactions occur through the brokerage market because the negative participation externalities are not properly internalized. This is more likely to occur when fixed costs are small or when brokers possess limited market power in setting their service fees, or both.

2.4.2 Socially optimal entry

To evaluate how the private entry decision of brokers potentially distort the market outcomes, we follow Mankiw and Whinston (1986) by comparing the private free entry equilibrium to that of a social planner who optimally chooses the number of brokers that operate the market, taking private broker behavior once they enter the market as given. That is, the planner maximizes W, given by expression (19) in which $N^B = N^S = nN$, by optimally choosing N, taking into account that the number of buyers and sellers attracted by individual brokers n is affected by N through the private first-order condition in the second stage of the model. The results are summarized in proposition 3.[11]

Proposition 3 *If for any N: (a) $dM/dN = n + Ndn/dN > 0$, (b) $Ndn/dN < 0$ and (c) $AR - MC > 0$, then the free entry equilibrium number of brokers N^{FE} strictly exceeds the socially optimal number of brokers, denoted by N^{SE}.*

The result that the private free entry equilibrium is always characterized by excessive entry is consistent with the findings of Mankiw and Whinston (1986), who demonstrate under the same set of assumptions (a)-(c) that in standard oligopoly models there is always excessive entry in the presence of fixed costs. The intuition is that private brokers do not account for the fact that they “steal business” from the incumbent brokers. That is, when a new broker enters, the market expands (assumption (a)) in the sense that more matches will be established through the brokerage market, but if the market expansion is smaller than the individual number of matches established by the incumbent brokers prior to the entry decision of the marginal entrant, this entrant also steals business from the incumbent brokers (assumption (b)). Absent of fixed costs, business-stealing has no social cost, i.e. generated revenues in the market are simply divided among more brokers. In the presence of fixed costs, however, business-stealing implies that investments in fixed costs are wasted from a social point of view, given that the same market outcome could also be established by less brokers and hence less investments in fixed costs. The presence of fixed costs also implies that assumption (c) required for the result in proposition 3 to hold – that brokers charge a strictly positive markup over marginal cost – is satisfied.

[11] The proof of proposition 3 can be found in Appendix A2.

2.5 Policy implications

The policy implications that follow from the results in propositions 1-3 are summarized in corollary 1. The first implication directly follows from combining the results in propositions 2 and 3. The second implication follows from the proof of proposition 3. The third implication follows from maximizing total social value in expression (19) with respect to $M = N^B = N^S$, while also allowing for the number of brokers that operate in the market to depend on M through the free entry condition (25), which is equivalent to maximizing consumer surplus, given by expression (18).

Corollary 1 *(i) The first-best social optimum can be established by setting the service fees charged by brokers such that the average per-match revenue equates the social cost to attract buyers and sellers and by minimizing the number of brokers that operate in the market:*

$$AR = MC + ET \text{ and } N \to 0 \tag{26}$$

(ii) When a social planner chooses the optimal number of brokers that operate in the market, while allowing them to privately compete in pricing their services once they have entered the market, the equilibrium number of matches M^{SE} established through the brokerage market satisfies:

$$AR = MC + ET + \frac{FC}{dM/dN} \tag{27}$$

where the market expansion effect of the marginal entrant (dM/dN) follows from differentiating the private first-order condition (20).

(iii) When a social planner sets the service fees to optimize the number of matches established in the brokerage market, while allowing brokers to freely enter the market, the equilibrium number of matches M^{SM} satisfies:

$$AR \to MC \text{ such that } N \to 0 \tag{28}$$

Corollary 1 shows the model outcomes when a social planner optimally regulates the service fees charged by brokers, broker entry or both. **Figure 3** illustrates the welfare effects. As a benchmark, the top left panel of Figure 3 plots a possible observed free entry equilibrium. In this case, the number of matches M^{FE} is determined by the point where the average cost curve (AC) crosses the average revenue curve (AR). Social value generated in the brokerage market is equal to surface A below the AR curve *minus* surface B below the ET curve, where the latter captures the social cost of the externalities present in the market. In free entry equilibrium brokers earn zero profits, so all surplus generated in the market is consumer surplus attributed to buyers and sellers. The remaining three panels in Figure 3 illustrate the implications of imposing the different policies described in corollary 1.

Firstly, when a social planner can regulate both brokerage service fees and market entry of brokers, implication *(i)* in corollary 1 applies. The planner equates average per-match revenue earned by brokers to the social cost of attracting buyers and sellers and minimizes the number of brokers to carry out the transactions. In the model, no integer constraint is imposed on the number of brokers and there are no constraints on the number of transactions a single broker can realize, so the socially optimal number of brokers approaches zero. Of course, in practice brokers have time constraints and therefore there is a limit to the number of transactions a single broker can establish in a given time period. The planner should thus approximate the number of brokers required to realize the desired number of transactions, while minimizing the amount of business brokers steal from one another when operating in the market. The top right panel in figure 3 illustrates the social first-best when a single broker can realize all desired transactions. Social surplus generated by the brokerage industry is equal to consumer surplus (surface A *minus* surface B) *plus* the profits earned by the brokerage industry (surface C). Total social value is unambiguously higher compared to the free entry equilibrium, although there might be shift in surplus from buyers and sellers to the brokers when the average service fee in the social first-best is higher than in the free entry equilibrium – as it is drawn in Figure 3.

Secondly, when a social planner can only influence the entry process of brokers, but not their pricing behavior once they have entered the market, implication *(ii)* in corollary 1 applies. Expression (26) shows that the markup earned by brokers in this case is higher compared to the social first-best. The social planner not only internalizes the externalities induced by the participation decision of buyers and sellers (ET), but also the fixed costs that brokers incur to operate in the market (FC) divided by the market expansion effect of the marginal entrant (dM/dN). The additional markup is larger when fixed entry costs are larger and when the market expansion effect ($dM/dN = n + Ndn/dN$) relative to the business-stealing effect (Ndn/dN) is smaller or, in other words, when the social cost induced by the marginal entrant is higher. Note that to implement this policy, the planner has to know how the optimal pricing behavior of brokers is affected by changes in the number of brokers that operate in the market, i.e. how θ is affected by N. The bottom left panel of Figure 3 illustrates the outcome, imposing Bresnahan (1989)'s constant conjectural variations model: $\theta = (1 + R)/N$, where R is calculated from the free entry equilibrium and is assumed to remain constant as the number of brokers changes. The figure demonstrates that consumer surplus is smaller and profits of the brokerage industry are larger compared to the social first-best.

Thirdly, implication *(iii)* in corollary 1 applies when the social planner can influence the pricing behavior of brokers, but not their entry decision. In this case, independent of the markup chosen by the planner, brokers enter the market until they all earn zero profits and hence brokers bring no surplus to the market. The planner therefore maximizes total consumer surplus, which is strictly increasing in M, as follows from expression (18). So, it is optimal to set average per-match revenue arbitrarily close to the per-match cost ($AR \rightarrow MC$), which minimizes the number of brokers that enter the market ($N \rightarrow 0$). Again, in practice the social planner should account for the time constraints of brokers and should target the service fees such that a minimal number of brokers enter the market to realize the desired transactions. The bottom right panel of Figure 3 illustrates the outcome when the service fees are set such that a single broker enters the market. The figure demonstrates there are no broker profits in this case and that more buyer and sellers participate in the market compared the social first-best. This comes at the expense, however, of a higher social cost due to the externalities present in the market (surface B) and therefore total surplus is smaller compared to the social first-best.

In general, and not just for the case drawn in Figure 3, interventions *(i)*, *(ii)* and *(iii)* are always (weakly) welfare improving compared to any observed free entry equilibrium, which can be anywhere on the continuum between the monopoly and Bertrand outcome, as discussed above. The welfare gain is always largest when imposing the social first-best (case *(i)*). However, which of the second-best cases *(ii)* or *(iii)* generates most welfare gains is ambiguous. It depends on the parameters of seller supply and buyer demand and on the cost structure of brokers and therefore essentially is an empirical question. The remainder of this paper empirically applies the model to the case of the real estate brokerage and further discusses the practical implications of the theoretical results.

3 Empirical methodology

This section presents a methodology to quantify the parameters of the theoretical model.[12] It is assumed that the following cross-sectional data are available for one or multiple local markets in which brokers compete for transactions (e.g. a city in the case of real estate brokerage) within a given time frame (e.g. one or multiple years). Firstly, at the market-level: the number of transactions carried out by the brokerage industry (M) relative to the potential number of transactions (S); the number of brokers that operate in the market (N); and some (in)direct measures of broker costs (MC and FC) – e.g. Hsieh and Moretti (2003) use the wage earned by employees in other service industries within local markets as a proxy for the opportunity cost to operate as a real estate broker, but direct cost measures are preferred. Secondly, at the transaction-level: a representative sample of brokered transactions, with details on the (average) service fees charged by the brokers; sales prices and product characteristics of the traded good; and some measures of buyer and seller characteristics. Finally, it is

[12] As a reminder, the exogenous parameters in the model are the parameters of the distributions of buyer demand and seller supply ($F^B(.)$ and $F^S(.)$, respectively), seller bargaining weight (β), market size (S), broker per-match (MC) and fixed (FC) costs and the parameter(s) of the underlying model of broker competition that determine broker market power (θ). The endogenous outcome variables are the number of transactions that occur in the brokerage market (M) and the number of brokers that operate in the market (N).

useful to observe some broker characteristics or to observe multiple transactions carried out by the same broker to control for broker heterogeneity, as they are assumed to be homogeneous in the model.

3.1 Parametric specification of seller supply and buyer demand

Assume that buyer valuations are uniformly distributed over the interval $[b^L, b^H]$ and that seller reservation prices are uniformly distributed over the interval $[s^L, s^H]$. The model then implies that buyers with a valuation in the range $[\tilde{b}, b^H]$ participate in the brokerage market, where $\tilde{b}$ is the valuation of the marginal buyer, and sellers participate when their reservation price is in the range $[s^L, \tilde{s}]$, where $\tilde{s}$ is the reservation price of the marginal seller. Given that market participants are assumed to be randomly assigned to one another, it follows that the prices at which the good is sold are distributed by a symmetric triangular distribution.[13] The lowest possible price at which the good is sold occurs when buyer type $\tilde{b}$ is matched to seller type s^L. The sales price is then equal to $p(\tilde{b}, s^L) = \beta\tilde{b} + (1-\beta)\frac{T+s^L}{1-t} \equiv p^{MIN}$, which is observed with probability zero. Similarly, the highest possible sales price is $p(b^H, \tilde{s}) = \beta b^H + (1-\beta)\frac{T+\tilde{s}}{1-t} \equiv p^{MAX}$, again observed with probability zero. The average sales price is the average of the minimum and maximum price: $\bar{p} = (p^{MIN} + p^{MAX})/2$, which is most likely to be observed.

In addition, the market clearing flat fee T and proportional fee t satisfy expressions (11) and (12), in which the marginal and average buyer and seller valuations can be written as a function of the fraction of buyers and sellers that participate in the market and the distributional parameters of buyer and seller reservation values. Combining expressions (11) and (12) with those for the minimum and maximum sales prices therefore allows to solve for the four relevant distributional parameters b^L, b^H, s^L and s^H as a function of the market clearing service fees (T and t), the fraction of buyers and sellers that participate in the market ($\frac{M}{S} = \frac{N^B}{S} = \frac{N^S}{S}$), seller bargaining weight (β) and the minimum and maximum sales price (p^{MIN} and p^{MAX}). By the assumption of linear supply and demand this system of equations has an analytical solution.

The average flat and proportional service fee and the fraction of buyers and sellers that participate in the market are assumed to be observed. So, it remains to obtain a proxy for seller and buyer bargaining weights and the minimum and maximum price of properties sold in the brokerage market to derive the parameters of supply and demand. The key problem to obtain a proxy for these measures using transaction data is that the theoretical model assumes that a homogeneous good is traded in the market,

[13] To see this, note from expression (4) that the sales price is a weighted sum of the buyer valuation and the seller reservation price and it is a familiar statistical property that any weighted sum of two independent continuous uniform random variables is distributed by a symmetric triangular distribution (e.g. Grinstead and Snell 1997).

whereas in practice traded goods are often heterogeneous in many dimensions. In other words, the model assumes that all dispersion in sales prices, measured by the difference between p^{MAX} and p^{MIN}, can be attributed to heterogeneity in buyer and seller characteristics, while in practice a large part of dispersion in sales prices can also be attributed to differences in the characteristics of the good – e.g. the size, location and age of a real estate property. A methodology therefore introduced that allows us to derive an upper and a lower bound on the dispersion of sales prices that can be attributed to heterogeneity in buyer and seller characteristics. To do so, we build on the hedonic pricing model of Rosen (1974) and the extension of Harding, Rosenthal and Sirmans (2003) that allows for bargaining among market participants. The estimated bounds on price dispersion can subsequently be used to obtain bounds on the range of feasible values for the parameters of supply and demand.[14]

3.2 Estimating residual sales price dispersion

Consider the following imperfectly competitive hedonic pricing model:

$$p_{gsb} = X_g \alpha^G + X_s \alpha^S + X_b \alpha^B + e_g + e_s + e_b \tag{29}$$

where p_{gsb} denotes the sales price of good g when being sold by seller s to buyer b. X_g denotes a vector of observable characteristics of the heterogeneous good sold in the market and α^G is the vector with corresponding coefficients that can be interpreted as the value a specific characteristic of the good on average contributes to the sales price of the good. In addition, as in Harding, Rosenthal and Sirmans (2003), and opposed to the competitive hedonic pricing model of Rosen (1974), it is assumed that not only the characteristics of the good influence the price at which it is sold, but also the characteristics of the buyer and seller involved in the transaction. The intuition is that not all values of the product characteristics (α^G) are always known to all buyers and sellers and these informational imperfections leave room for bargaining over the sales price. This is typically the case in markets that are thin because the traded good is very heterogeneous, as in real estate markets. The vector X_s contains seller characteristics and the corresponding coefficient vector α^S measures how much these characteristics contribute in determining the sales price. Similarly, the vector X_b with coefficients α^B captures how buyer characteristics contribute to the sales price. The residuals e_g, e_s and e_b capture unobserved heterogeneity in product, seller and buyer characteristics, respectively.

[14] Note that we face a nonstandard identification problem. With data on prices and quantities of goods sold by firms in standard product markets there are many techniques available in the literature to estimate consumer demand and firm market power. See, for example, Bresnahan (1989), Perloff, Karp and Golan (2007) and Einav and Levin (2010) for reviews. These techniques do not account, however, for the role of intermediaries. In the empirical literature on two-sided markets, there are some papers that estimate market power of platforms in the presence of externalities among different types of consumer groups (e.g. Rysman 2004; Lee 2013; Jeziorski 2014), but they typically do not allow for bargaining among matched trading partners. Finally, Bajari and Benkard (2005) propose a more general methodology than ours to estimate parameters of consumer demand and seller supply using the hedonic approach, but without intermediaries.

To estimate the upper bound on the dispersion of observed sales prices that can be attributed to buyer and seller characteristics, the following hedonic pricing regression can be estimated:

$$p_{gsb} = X_g \alpha^G + \varepsilon_{gsb} \tag{30}$$

where the dispersion in the error term ε_{gsb} is interpreted as the residual dispersion in sales prices that can be attributed to buyer and seller characteristics *and* to unobserved heterogeneity in product characteristics. Thus, if all relevant characteristics of the good that influence the sales price would be observed, the variance of e_g in expression (29) would be zero, and ε_{gsb} would solely capture buyer and seller heterogeneity. If not all relevant product characteristics are observed, the variance of e_g is positive, and ε_{gsb} overestimates the heterogeneity in sales prices that can be attributed to buyers and sellers. Therefore the dispersion of ε_{gsb} is interpreted as an upper bound for the dispersion in sales prices that comes from buyer and seller heterogeneity. The values of p^{MIN} and p^{MAX} that follow can be obtained by fitting the symmetric triangular distribution to the distribution of the residuals around the predicted value of the regression.

To estimate the lower bound we want to estimate how much the terms $X_s \alpha^S$ and $X_b \alpha^B$ in expression (29) contribute to the variation in sales prices. If we would estimate equation (29), however, the obtained coefficients for these terms would likely be biased in the presence of unobserved product heterogeneity. This because different buyers and sellers are expected to differently value product characteristic and hence the component e_g in the error term will be correlated with the regressors in X_s and X_b. More specifically, when $e_g = X_s \delta^S + X_b \delta^B + e_g{}'$, where δ^S and δ^B measure how much sellers and buyers value the unobserved product characteristics, e_g in expression (29) is clearly correlated with X_s and X_b when δ^S and δ^B differ from zero. To solve this, we follow Harding, Rosenthal and Sirmans (2003) by introducing two symmetry assumptions. Firstly, that the valuation of identical buyers and sellers for the unobserved product characteristics is the same, that is $\delta^S = \delta^B$. Secondly, that the way identical buyers and sellers can influence the sales price through the bargaining process is the same in magnitude but opposite. That is, $\alpha^S = -\alpha^B$, which implies that the amount by which a certain degree of education, for example, allows a seller to push up the sales price is the same as it allows a buyer with the same educational level to push it down. Accounting for this allows us to rewrite equation (29) as follows:

$$p_{gsb} = X_g \alpha^G + \alpha(X_s - X_b) + \delta(X_s + X_b) + e_g{}' + e_s + e_b \tag{31}$$

in which the term $\alpha(X_s - X_b)$ estimates how much buyer and seller attributes contribute to the variation in sales prices through the bargaining process and the term $\delta(X_s + X_b)$ estimates the valuation of buyers and sellers for unobserved product characteristics. The values of p^{MIN} and p^{MAX}

can now be obtained by fitting the symmetric triangular distribution to the predicted values of the term $\alpha(X^S - X^B)$ around the predicted value of the regression.

As a final step, consistent with the assumptions in the empirical specification, it is assumed that the Nash bargaining game in the theoretical model is symmetric. That is, the bargaining weight of both buyers and sellers is one half: $\beta = 1 - \beta = 0.5$. Using the estimated bounds for p^{MAX} and p^{MIN} then allows us to calculate bounds for the values for the distributional parameters of the model, as described in the previous subsection. In addition, if either MC or FC iss observed, the other cost measure of the two can be calculated by using the zero profit condition (25). Then, using MC and the parameters of seller supply and buyer demand, broker market power (θ) can be calculated from the private first-order condition (20), which closes the model.

4 Data

4.1 Transaction-level data

The main dataset used for the analysis is a sample of 26,986 residential real estate properties that were sold in Belgium through one of 97 real estate agencies of a large franchise system in the period 2005-2014.[15] **Table 1** provides descriptive statistics on sales prices and service fees charged by brokers (the latter are only available since 2011). An average property is sold for €215,579, ranging from €90,000 at the 5th percentile to €392,500 at the 95th percentile. The average flat fee charged by brokers is €2,786, ranging from €0 to €6050 and the average proportional fee is 3%, ranging from 0 to 4.2%. This implies that brokers charge on average a total service fee of €9,182 or a commission rate of 4.3% for an average priced property.[16]

[15] The sample is restricted to houses, excluding apartments, for which another 10,666 transactions are observed. The same qualitative results are obtained when only using apartments in the analysis below, or when using both houses and apartments. Using both complicates the regression analysis because some observable property characteristics might affect the price of houses and apartments differently. Therefore, we prefer to exclude apartments from the sample.

[16] Note that, especially compared to the US where brokers usually charge a fixed commission rate of 5 or 6% (e.g. Hsieh and Moretti 2003), the service fees in our sample show strong variation. Furthermore, the service fees are on average lower and brokers use various pricing strategies – for 8% of the transactions only a flat fee was charged, for 17% only a proportional fee and for 75% a combination of both. These observations suggest that price competition among brokers is stronger in the Belgian brokerage industry than in the US industry. We believe that the crucial institutional difference that makes the Belgian market more competitive than the US market is that buyers (almost) never hire a broker in their search for a real estate property and only sellers hire brokers to sell their properties. This allows brokers to supply their services more independently than in an MLS system where real estate agents rely heavily on their colleagues to sell properties and can be penalized when deviating from the conventional commission rate (e.g. Levitt and Syverson 2008a). The reason for the absence of buyer representation by brokers in the Belgian market is likely due to the fact that every real estate transaction has to be approved and concluded by a notary, who supervises that all legal and administrative requirements are satisfied. Notaries thus essentially take up the role of guiding buyers through the process of buying a property which is executed by private brokers in the US. Note that notaries also receive a fee for this service, however,

In addition, the dataset contains the initial listing price when properties were first brought on the market, time-to-sell and a broad range of observable property characteristics, such as size, age and number of bedrooms. Importantly, the exact location of properties is also observed, which allows us to construct measures such as the distance to the closest city center, distance to the capital city (Brussels) or distance to the nearest train station. For about half of the transactions, the data also contains the previous address of the buyer of a property. Observing the previous location of residence of buyers and sellers allows us to construct indirect measures of buyer and seller characteristics using publicly available administrative data for local living areas, for example, on median income, age and educational level of the population.[17]

4.2 Market-level data

The transaction data are complemented with aggregate data on the Belgian real estate brokerage industry. More specifically, data were collected on total market size, market share of the brokerage industry, broker entry and broker costs. **Table 2** summarizes the data using 2013 as the reference year.

Firstly, as a measure of market size, the total number of real estate transactions that occurred in Belgium in 2013 is used, calculated from publicly available administrative data from Statistics Belgium. In total there were 123,652 registered real estate transactions, of which 80,491 were houses and 43,161 apartments. So, the measure of market size consist of all the properties that were sold through the brokerage industry *plus* all the properties that were sold in the outside market (for-sale-by-owner).[18]

Secondly, a proxy for the fraction of transactions that occurred through the brokerage industry is obtained from a survey conducted by the Policy Research Center for Housing. The survey questioned 10,000 households that were randomly selected from the civil register about their current housing status. In the period 2009-2013, 710 of these households purchased a real estate property and 397 of them, or approximately 56%, claim that they bought the property from a seller that was assisted by a real estate broker.

Thirdly, two measures for the number of brokers that operate in the Belgian real estate market were collected. The first measure comes from data provided by the professional association of real estate

they are not private entities. Both the number of notaries and the fee they can charge for their service is highly regulated.

[17] Table B1 in the appendix reports the descriptive statistics for the observable property characteristics and Table B2 for the proxies of buyer and seller characteristics.

[18] The measure of market side should be interpreted as a lower bound on the actual potential market size, given that the measure does not include properties that were put up for sale, but remained unsold. In addition, it is possible that there were some buyers and sellers that would have entered the real estate market under different conditions (e.g. should broker service fees have been lower), but eventually decided not to enter.

brokers in Belgium (BIV), which contains the address of all registered members on the 1st of January 2011, 2012 and 2013. The data show that the number of registered brokers remains stable over these three years and in 2013 there were 8963 registered brokers. The advantage of this measure is that registration with the professional association is mandatory in Belgium, which implies that all persons who are licensed to broker real estate transactions are included. The problem, however, is that not all brokers that are included in the list are necessarily active (full-time), so this number should be interpreted as an upper bound. As a second measure for the number of brokers active in the market we collected data from the largest online real estate listing platform in Belgium (www.immoweb.be). On the 12th of December 2013, 3303 real estate agencies had at least one active real estate listing on the website. Assuming that the number of real estate brokers per agency is similar to that in the BIV-data (approximately 1.36), there were 4494 real estate brokers active on the listing platform. Of course, it is unlikely that every broker in Belgium had an active listing on that day, so this measure is interpreted as a lower bound on the number of brokers that operate in the market. The average of both measures is 6,728.

Finally, data were collected from various sources on the advertisement and administrative costs to sell a real estate property in Belgium, resulting in a proxy of €983 for the monetary per-match cost.[19] In addition, when assuming that brokers can freely enter the market and earn zero profits, the costs to operate as a broker should also include the income a broker could earn when practicing a different profession. As a proxy for this opportunity cost, the average yearly wage of employees working in other service sectors than real estate brokerage is used, which was equal to €48,525 in 2013, as reported by the National Bank of Belgium.

Using a measure for the per-match cost MC, the implied fixed cost FC can be calculated from the zero profit condition (25). The key question, however, is whether the opportunity cost to operate as a broker should be included in the measure for marginal costs or for fixed costs. On the one hand, it can be argued that the opportunity cost reflects the value of time that brokers invest in selling real estate properties. In this case, the opportunity cost (divided by the average number of yearly transactions per broker) should be included in the measure for MC, which results a proxy of €5,696 for MC and of €35,882 for FC. The amount of €35,882 then serves as a proxy for the monetary operating costs that brokers incur on a yearly basis independent of the number of properties they sell. These might be costs linked to office space, office supplies, purchasing or leasing a car, obtaining the broker license, the franchise fee, etc. On the other hand, as argued by Hsieh and Moretti (2003), for example, when brokers can freely enter the market, they are likely to waste valuable time and other recourses while inefficiently competing for transactions, especially when broker commission rates are fixed, as is

[19] See Table B3 in the appendix for details.

typically observed in the US. By this logic, in the extreme case when all broker time is unproductive, the opportunity cost should be fully included in the measure of fixed costs. When the estimated monetary per-match cost of €983 is used as a proxy for MC, the implied fixed cost FC is €84,389. This measure then includes both the opportunity cost and the other fixed monetary operating costs. It seems reasonable to assume, however, that at least part of the time spent by brokers is productive, especially in our setting where broker commission rates are not fixed. For the baseline calibration of the model, the average proxy for MC of €3,339 and for FC of €60,141 is used, which each include half of the opportunity cost.

5 Model calibration and welfare counterfactuals

Using the data described in the previous section, this section first calibrates the outcomes of the theoretical model by applying the empirical methodology proposed in section 3. Subsequently, different welfare counterfactuals are constructed and discussed. Finally, sensitivity analysis is provided using alternative measures for broker costs and for the number of brokers that operate in the market.

5.1 Model calibration

The first step is to obtain values for the parameters of buyer demand (b^L, b^H) and seller supply (s^L, s^H). To do so, remember that an estimate is required for dispersion in sales prices that can be attributed to buyer and seller characteristics and that the proposed methodology allows to estimate bounds on this dispersion. For the upper bound, after estimating the hedonic pricing equation (30), the top panel of **Figure 4** plots a Kernel density of the residuals around the predicted value of the regression.[20] In addition, the figure plots the fitted symmetric triangular distribution that minimizes the distance between the kernel density and the fitted distribution. The implied minimum sales price (p^{MIN}) is €141,064 and the maximum sales price (p^{MAX}) is €289,494 around the average of €215,279. So, the spread in sales prices due to heterogeneity in buyer and seller reservation values is therefore estimated to be €148,430. For the lower bound, after estimating equation (31), the bottom panel of Figure 4 plots a kernel density of the predicted values of the term $\alpha(X^S - X^B)$ around the predicted value of the regression and the corresponding fitted triangular distribution.[21] The estimate for the minimum price is €206,807 and for the maximum price €223,751, implying a spread of €16,944.

The implied values for the parameters of supply and demand are reported in **Table 3**. For the upper bound on sales price dispersion, the valuation of buyers ranges from €61,547 to €326,601 and the

[20] The results of estimating regression equation (30) using OLS are reported in the first column of Table B4 in the appendix.
[21] The regression results of estimating equation (31) using OLS are reported in the second column of Table B4 and Table B5 in the appendix.

reservation price of sellers from €98,081 to €355,259. For the lower bound on sales price dispersion, the spread ranges from €197,729 to €227,986 for buyer valuations and from €193,766 to €223,124 for seller reservation prices. Note that by construction dispersion in buyer valuations and seller reservation prices for the upper bound is larger than for the lower bound, which implies that for the upper bound buyer demand and seller supply are relatively inelastic with respect to changes in the broker service fee compared to the lower bound. For the upper bound the average revenue curve, plotted in Figures 1-3 above, is therefore relatively inelastic and thus relatively steep. For the lower bound the average revenue curve is flatter.[22]

Also note that the difference between the upper and the lower bound in the dispersion of sales prices attributable to buyer and seller heterogeneity is large, €148,430 versus €16,944, respectively. This suggests that in the estimation for the upper bound there are still many property characteristics that are unobserved. Similarly, for the lower bound, there are likely many other unobserved buyer and seller characteristics that influence the sales price of properties. If all relevant property, buyer and seller characteristics would be observed, both estimates would yield the same spread. So, the question arises which of the two measures comes closest to reality. For the upper bound the spread implies, for example, that if the average seller would be lucky and be matched with the highest valuation buyer, the property would sell for €252,386. If unlucky and being matched to the lowest valuation buyer that participates in the market, the property would only sell for €178,171. Similarly, for the lower bound the property of the average seller would sell for a price ranging between €219,515 and €211,043. Intuitively, the spread of about €8,500 that can be attributed to "luck" in meeting the best trading partner in the lower bound perhaps comes closer to reality than the spread of about €74,000 implied by the upper bound. In what follows, the results are always reported for both the upper and the lower bound.

As a second step, various outcome variables of the model can be calculated by combining the obtained values for the parameters of supply and demand with the market-level data reported in Table 2. **Table 4** shows the calibrated values for the outcome variables that determine the optimal private service fee, as described in proposition 1. The table shows that the observed total service fee ($AR = €9{,}182$) is significantly above marginal cost ($MC = €3{,}339$). As shown by expression (20) in proposition 1, in a private market this markup is comprised of the sum of marginal surplus MS and the tax raised to internalize the participation externalities of buyers and sellers ET, weighted by the measure of broker market power θ. For the estimated upper bound of sales price dispersion, which corresponds to

[22] More specifically, the estimates for the parameters of supply and demand imply that the elasticity of the average revenue curve at the observed outcome is 0.04 for the upper bound on price dispersion and 0.37 for the lower bound. This implies that an increase in the average commission rate from the current 4.3% to 5.3%, for example, decreases the number of transactions in the brokerage market by 1% for the inelastic and by 9% for the elastic AR curve, respectively.

inelastic seller supply and buyer demand, $MS = €146{,}224$ and $ET = €73{,}112$ are relatively large compared to when supply and demand are elastic, $MS = €16{,}692$ and $ET = €8{,}346$. Given that expression (20) is assumed to hold as an identity, corresponding broker market power is relatively small for inelastic compared to elastic supply and demand ($\theta = 0.026$ and $\theta = 0.233$, respectively). Finally, total social value generated by the Belgian real estate brokerage industry in 2013 is estimated to be about 5 billion for the inelastic and 578 million for the elastic bound on supply and demand.

5.2 Welfare counterfactuals

In this subsection, the observed private market outcomes described in Table 4 are compared to those determined by a social planner. The three scenarios described in corollary 1 are considered and the results are summarized in **Table 5**. The top panel of Table 5 corresponds with case *(i)* in corollary 1 and reports the model outcomes when the social planner chooses the optimal number of transactions in the brokerage market, while minimizing the number of brokers that operate in the market. The results show that for both measures of inelastic and elastic supply and demand the planner attracts less buyers and sellers (a fraction of 0.43 and 0.52, respectively) compared to the observed private market outcome (where a fraction of 0.56 of the transactions occur in the brokerage market). This implies that the observed average service fee is below the socially desired level and the participation externalities of buyers and sellers are insufficiently internalized. The current commission rate is on average 4.3% and the optimal counterfactual commission rates are 5.1% for elastic and 23.7% for inelastic supply and demand, respectively. Assuming that all the transactions can be realized by a single broker, imposing the social first-best would imply a welfare gain ranging from 19% for inelastic to 71% for elastic supply and demand. In practice, more than one broker is required of course to realize the desired number transactions. So, when appointing a realistic number of brokers the welfare gain would be smaller, as these brokers have to incur fixed operating costs. In addition, note that imposing the first-best implies a loss in consumer surplus allocated to buyers and sellers and the net gain comes from increased broker profits.

The middle panel of Table 5 corresponds with case *(ii)* in corollary 1 and reports the model outcomes when the social planner determines the number of brokers that operate in the market, while allowing them freely compete in pricing their services once they have entered the market. To do this, an assumption has to be made on how broker market power is affected when the number of brokers in the market changes. More specifically, Bresnahan (1989)'s constant conjectural variations model is imposed for which $\theta = (1 + R)/N$ where $dM/dn = 1 + R$. The conjectural variations parameter R can be calculated from the estimates of θ for the observed market outcome, as reported in Table 4. In this case, the social planner reduces broker entry from the current 6728 to 2,261 for elastic and to 517 for inelastic supply and demand. The corresponding commission rates increase from the current 4.3%

to 7.3% and 25.3%, respectively. This regulation entails an estimated welfare gain between 18% for inelastic and 40% for elastic supply and demand compared to the observed market outcome. Note that this policy implies an even larger loss in consumer surplus and a comparable gain in broker profits compared to the social first-best.

The bottom panel of Table 5 corresponds with case *(iii)* in corollary 1. The counterfactual is constructed should the service fee be set such that exactly one broker enters the market. That is, the service fee is equated to the average cost of a single broker, which implies a commission rate of 1.6%. In this case, a fraction of 0.58 of the buyers and sellers participate in the brokerage market for inelastic supply and demand and a fraction of 0.69 for elastic supply and demand. Given that the counterfactual service fee is now below the observed service fee, there is a gain in consumer surplus of 5% to 52%. Broker profits remain zero, as the free entry condition continues to apply under this scenario.

Overall, Table 5 suggests that the effectiveness of regulating broker entry or broker service fees crucially depends on how sensitive participation of buyers and seller is to changes in the service fee. For the inelastic bound, regulating broker entry is more effective than regulating the service fees, while the reverse holds for the elastic bound. In addition, when regulating broker entry, there can be important redistributive effects that shift surplus from buyers and sellers to brokers, which a social planner might want to take under consideration. If a regulator can only regulate entry, but nevertheless is only concerned with consumer surplus and not with broker profits, one possible solution is to sell licenses to brokers – i.e. impose a lump sum tax to operate as a broker. At a right price, this can induce the optimal number of brokers to enter the market under the second scenario in Table 5, while broker profits would remain zero. The revenues of this taxation could then be redistributed to buyers and sellers through other real estate market policies.

5.3 Sensitivity analysis

As discussed in section 4, some of the parameter values used for the baseline calibration of the model might suffer from measurement errors. This section discusses the sensitivity of the results to deviations of the model parameters from their baseline values.

To start, remember that the proxy for marginal costs includes half of the opportunity cost to operate as a broker, measured by the wage brokers could potentially earn when working in a different service sector. This implicitly assumes that half of the effort of brokers goes to productively selling real estate properties and half is unproductive effort spent on marketing their services and competing with other brokers for transactions. The middle two columns of **Table 6** present the model outcomes should all effort be unproductive. In this case the opportunity cost is fully included in the measure for fixed costs (€84,389) and only the monetary costs of marketing and selling a real estate property are included in the measure for marginal costs (€983). The final two columns present the opposite case where all effort is assumed to be productive ($FC = €35{,}882$ and $MC = €5{,}696$). The table shows that the results are robust to the alternative specifications of broker costs. The observed service fee (€9,182) always remains too low compared to the socially optimal fee, although it comes very close to the social optimum for elastic supply and demand and the lower bound on marginal costs. Intuitively, the estimated welfare gains from all policy interventions are larger for the lower bound on marginal costs compared to the baseline case and they are smaller for the upper bound. The only qualitative difference compared to Table 5 is that for the upper bound on marginal costs regulating broker entry is now more effective than regulating the service fee for both inelastic and elastic supply and demand.

In addition, remember from Table 2 that an upper and a lower bound on the number of brokers active in the market is observed and the average of both was used for the baseline calibration. **Table 7** reports the model outcomes for the upper and the lower bound. Again, none of these alternative specifications qualitatively alter the conclusions of the baseline case. Finally, the robustness of the results was also tested for possible measurement errors in the parameter values of market size (S), brokerage industry market share (M/S) and buyer and seller bargaining weight (β). For reasonable deviations from their baseline values, none of these qualitatively alter the conclusions of the baseline specification and the results are therefore omitted.

6 Conclusion

This paper aimed to make two contributions. Firstly, to present a general model of imperfect competition among intermediaries that operate in a one-to-one matching markets, in which the intermediaries are allowed to freely enter the market and flexibly compete in pricing their services. The model showed that some private broker market power is justified from a social perspective, such that the broker service fee properly internalizes the participation externalities of buyers and seller. In addition, it showed that generally an excessive number of intermediaries operate in a private market compared to what is socially desirable. The second contribution is to derive policy implications from this setting and to quantify the effects of various counterfactual regulatory interventions using data from the Belgian real estate brokerage industry. The counterfactuals suggest that regulating broker entry is more effective when seller supply and buyer demand are relatively insensitive to changes in the service fee charged by brokers. In contrast, targeting broker service fees is more effective when supply and demand is elastic. A regulator should be cautious, however, about redistributive effects that shift surplus from buyers and sellers to brokers when regulating broker entry, whereas regulating service fees always increases consumer surplus.

In the theoretical model some simplifying assumptions were made that abstract from important realistic features of matching markets. It is important to explore the implications of relaxing these assumptions in future work. To start, the fact was ignored that matching markets not only clear on prices, but also on the time dimension. For example, it is well-known that in real estate markets there exists a tradeoff for sellers between selling quickly and selling at a high price (e.g. Han and Strange 2015). Exploring the implications of broker competition in a dynamic setting that explicitly models the search process of buyers and sellers is an important direction for future research. In addition, the model abstracted from the use of a list price as a strategic instrument for sellers (or brokers) to market the good sold in the market (e.g. Albrecht, Gautier and Vroman 2015). Exploring broker competition in a model of directed search with posted prices could render interesting new insights. We also abstracted from issues concerning the principle-agent relationship between sellers and brokers. For example, the incentives of a seller and a broker in the marketing process of a real estate property are not necessarily aligned (e.g. Rutherford, Springer and Yavas 2005, Levitt and Syverson 2008b). It would be interesting to further explore the incentive effects of broker competition, as in Fisher and Yavas (2010). Finally, we ignored possible heterogeneities in the quality of the services offered by different brokers and other possible institutional differences across local markets. With additional data on broker and local market characteristics the analysis can be further refined. In this light, it is also of particular interest to allow for heterogeneous outcomes in the outside market, as in Hendel, Nevo and Ortalo-Magné (2009).

For the empirical analysis, it would be interesting to apply our framework using data from other real estate markets and compare those to our results. Of particular interest, is to apply our methodology to case of the US, which allows to test whether the conventional commission rate of 5-6% charged by real estate brokers is in the range of the socially optimal commission rate. In addition, our model can also be applied to other matching markets, such as labor markets or second hand goods markets, for which different policy implications might apply than those derived for real estate markets in the present paper.

Appendices

Appendix A1: Equivalence different fee structures

In the main text only sellers and not buyers are directly charged for the brokerage service. By the assumption of Nash bargaining, however, the service fee can partially be passed through to buyers. Because Nash bargaining is efficient, the model outcomes are independent to whether the service fee is charged to sellers or buyers. To see this, consider the opposite case than the one analyzed in the main text where only buyers and not sellers are directly charged. In this case, the individual transaction valuation of a buyer type b is equal to $b-(1+t)p-T$ and of a seller type s is equal to $p-s$. Nash bargaining implies that the transaction price when a buyer type b and a seller type s are matched is $p(b,s)=\beta(b-T)/(1+t)+(1-\beta)s$. The inverse demand equations can then be written as $T=(\tilde{b}\tilde{s}-\bar{b}\bar{s})/(\tilde{s}-\bar{s})$ and $1+t=(\bar{b}-\tilde{b})/(\tilde{s}-\bar{s})$ and the average sales as $\bar{p}=\beta\tilde{s}+(1-\beta)\bar{s}$. Combining these expressions yields the following expression for average revenue: $AR\equiv t\bar{p}+T=\beta(\bar{b}-\tilde{s})+(1-\beta)(\tilde{b}-\bar{s})$, which is identical to expression (16) in the main text. So, expression (13) for broker profits and expression (19) for welfare are also identical and all optimal pricing results carry through independent to which side of the market the service fee is charged.

Appendix A2: Proof proposition 3

Differentiating expression (19) for total social value, in which $N^S=N^B=nN$, with respect to N yields:

$$\frac{dW}{dN}=\left(\tilde{b}-\tilde{s}-MC\right)\left(n+N\frac{dn}{dN}\right)-FC \qquad (A1)$$

Equating expression (A1) to zero, using the expressions for AR (16) and ET (22), that $n+Ndn/dN=dM/dN$ and rewriting yields expression (27). Note that dM/dN can be written as a function of M by solving the private first-order condition (20) for N as a function of M (the solution is unique by assumption (b)) before differentiating. So, expression (27) can be written solely as a function of M (independent of N) and hence can be solved for the equilibrium number of matches M^{SE} at the social optimum.

The excessive entry result follows from adding to and subtracting from expression (A1) expression (13) for individual broker profits, in which $n^S=n^B=n$ and $N^S=N^B=nN$, which after rewriting yields:

$$\frac{dW}{dN}=\pi+(AR-MC)N\frac{dn}{dN}-ET\left(n+N\frac{dn}{dN}\right) \qquad (A2)$$

Expression (A2) illustrates the distortions that result from free entry in the private market relative to the social optimum. Under free entry in the private market individual broker profits equate zero ($\pi = 0$), while entry is socially optimal when the impact of the marginal entrant on social welfare is zero ($dW/dN = 0$). So, expression (A2) implies that private entry and socially optimal entry coincide when the sum of the second and the third term equals zero. When the sum of these terms is negative, there is excessive entry. This because $d\pi/dN < 0$, so $dW/dN = 0$ only holds when the number of brokers is smaller than under private entry. By assumptions (a)-(c) and the fact that $ET > 0$, the last two terms in (A2) are strictly negative and hence the private free entry equilibrium is unambiguously characterized by excessive entry. QED

References

Albrecht, J., Gautier, P. and Vroman, S. 2015. "Directed search in the housing market" forthcoming in *Review of Economic Dynamics*.

Anglin, P. and Arnott. R. 1999. "Are brokers' commission rates on home sales too high? A conceptual analysis" *Real Estate Economics*, *27*(4):719-749.

Armstrong, M. 2006. "Competition in Two-Sided Markets" *RAND Journal of Economics*, 37(3):668–691.

Bajari, P. and Benkard, L. 2003. "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach" *Journal of Political Economy*, 113(6):1239-1276.

Barwick, P.J. and Pathak, P.A. 2015. "The Costs of Free Entry: An Empirical Study of Real Estate Agents in Greater Boston" *RAND Journal of Economics*, 46(1):103–145.

Baye, M., Morgan, J. and Scholten, P. 2007. "Information, Search and Price Dispersion" *Handbook on Economics and Information Systems*, T. Hendershott (ed.), Elsevier.

Biglaiser, G. 1993. "Middlemen as Experts" *RAND Journal of Economics*, 24(2):212-223.

Bresnahan, T.F. 1989. "Empirical Studies of Industries with Market Power" in *Handbook of Industrial Organization*, Vol.2:1011-57, R. Schmalensee and R. Willig (eds.), Amsterdam: North-Holland.

Caillaud, B. and Jullien, B. 2003. "Chicken and Egg: Competition among Intermediation Service Providers" *RAND Journal of Economics*, 34(2):309–328.

Carroll, W. 1989. "Fixed-percentage commissions and moral hazard in residential real estate brokerage" *The Journal of Real Estate Finance and Economics*, *2*(4):349-365.

Crockett, J.H. 1982. "Competition and Efficiency in Transacting: The Case of Residential Real Estate Brokerage" *Journal of the American Real Estate and Urban Economics Association*, 10(2):209-227.

Damiano, E. and Li, H. 2007. "Price Discrimination and Efficient Matching" *Economic Theory*, 30:243–263.

Damiano, E. and Li, H. 2008. "Competing Matchmaking" *Journal of the European Economic Association*, 6:789–818.

Einav, L. and Finkelstein, A. 2011. "Selection in Insurance Markets: Theory and Empirics in Pictures" *Journal of Economic Perspectives*, 25(1):115–138.

Einav, L. and Levin, J. 2010. "Empirical Industrial Organization: A Progress Report" *Journal of Economics Perspectives*, 24(2):145-162.

Fisher, L.M. and Yavas, A. 2010. "A case for percentage commission contracts: The impact of a "Race" among agents" *The Journal of Real Estate Finance and Economics*, *40*(1):1-13.

Genesove, D. and Mayer, C. 2001. "Loss aversion and seller behavior: evidence from housing markets" *Quarterly Journal of Economics*, 116(4):1233-1260.

Gomes, R. and Pavan, A. 2015. "Many-to-Many Matching and Price Discrimination" Forthcoming in *Theoretical Economics*.

Goos, M., Van Cayseele, P.J.G. and Willekens, B. 2014. “Platform Pricing in Matching Markets” *Review of Network Economics*, 12(4):437-457.

Grinstead, C.M. and Snell, J.L. 1997. “Introduction to probability” American Mathematical Society.

Han, L. and Hong, S.-H. 2011. "Testing cost inefficiency under free entry in the real estate brokerage industry" *Journal of Business & Economic Statistics*, 29(4):564-578.

Han, L. and Strange, W.C. 2015. “The Microstructure of Housing Markets: Search, Bargaining, and Brokerage” *Handbook of Regional and Urban Economics*, Vol. 5, G. Duranton, J.V. Henderson and W.C. Strange (eds.), Elsevier.

Harding, J.P., Rosenthal S.S. and Sirmans, C.F. 2003. "Estimating Bargaining Power in the Market for Existing Homes" *Review of Economics and Statistics*, 85(1):178-88.

Hendel, I., Nevo, A. and Ortalo-Magné, F. 2009. “The relative performance of real estate marketing platforms: MLS versus FSBO Madison.com” *American Economic Review,* 99(5):1878-98.

Hsieh, C.T. and Moretti, E. 2003. “Can Free Entry be Efficient? Fixed Commissions and Social Waste in the Real Estate Industry” *Journal of Political Economy*, 111(5):1076-1122.

Jeziorski, P. 2014. “Effects of Mergers in Two-Sided Markets: The U.S. Radio Industry” *American Economic Journal: Microeconomics*, 6(4): 35-73.

Lee, R.S. 2013. “Vertical Integration and Exclusivity in Platform and Two-Sided Markets” *American Economic Review*, 103(7):2960–3000.

Levitt, S. and Syverson C. 2008a. “Antitrust implications of outcomes when home sellers use flat-fee real estate agents” *Brookings-Wharton Papers on Urban Affairs, 2008.*

Levitt, S. and Syverson, C. 2008b. “Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions” *Review of Economics and Statistics*, 90(4):599–611.

Loertscher, S. and Niedermayer, A. 2015. "Percentage Fees in Thin Markets: An Optimal Pricing Perspective" Working Paper.

Mahoney, N. and Weyl, G. 2015. “Imperfect Competition in Selection Markets” Working paper.

Mankiw, G. and Whinston, M. 1986. “Free Entry and Social Inefficiency” *RAND Journal of Economics*, 17(1):48-58.

Miceli, T.J. 1992. "The Welfare Effects of Non-Price Competition among Real Estate Brokers" *Journal of the American Real Estate and Urban Economics Association*, 20(4):519-532.

Miceli, T.J., Pancak, K.A. and Sirmans, C.F. 2007. "Is the Compensation Model for Real Estate Brokers Obsolete?" *Journal of Real Estate Finance and Economics*, 35(1):7-22.

Niedermayer, A. and Shneyerov, A. 2014. “For-Profit Search Platforms” *International Economic Review*, 55(3), 765–789.

Perloff, J.M., Karp, L.S. and Golan, A. 2007. “Estimating Market Power and Strategies” Cambridge University Press.

Rochet, J-C. and Tirole, J. 2003. "Platform Competition in Two-Sided Markets" *Journal of the European Economic Association*, 1:990-1029.

Rosen, S. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition" *Journal of Political Economy*, 82(1):34-55.

Rubinstein, A. and Wolinsky, A. 1987. “Middlemen” *Quarterly Journal of Economics*, 102:581-594.

Rutherford, R.C., Springer T.M. and Yavaş, A. 2005. “Conflicts between Principals and Agents: Evidence from Residential Brokerage” *Journal of Financial Economics*, 76(3):627–65.

Rysman, M. 2004. “Competition between Networks: A Study of the Market for Yellow Pages” *Review of Economic Studies*, 71(2):483–512

Spulber, D.F. 1999. “Market Microstructure: Intermediaries and the Theory of the Firm” Cambridge: Cambridge University Press.

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

US Department of Justice and Federal Trade Commission. 2007. "Competition in the Real Estate Brokerage Industry".

Veiga, A. and Weyl, G. 2015. "Product Design in Selection Markets" Working Paper.

Weyl, G. 2010. "A Price Theory of Multi-Sided Platforms" *American Economic Review*, 100(September 2010):1642-1672.

Weyl, G. and Fabinger, M. 2013. "Pass-Through as an Economic Tool: Principle of Incidence under Imperfect Competition" *Journal of Political Economy*, 121(3):528-583.

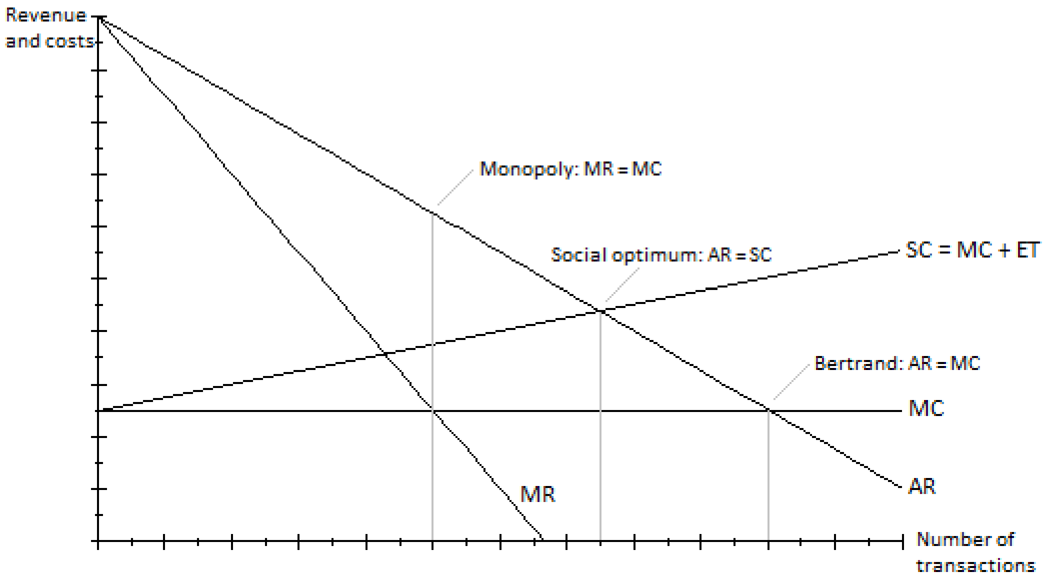
White, A. and Weyl, G. 2015. “Insulated Platform Competition” Working Paper.

Yavas, A. 2001. "Impossibility of a Competitive Equilibrium in the Real Estate Brokerage Industry" *Journal of Real Estate Research*, 21(3):187-200.

Yavas, A. 1994. "Middlemen in Bilateral Search Markets" *Journal of Labor Economics*, 12:406-429.

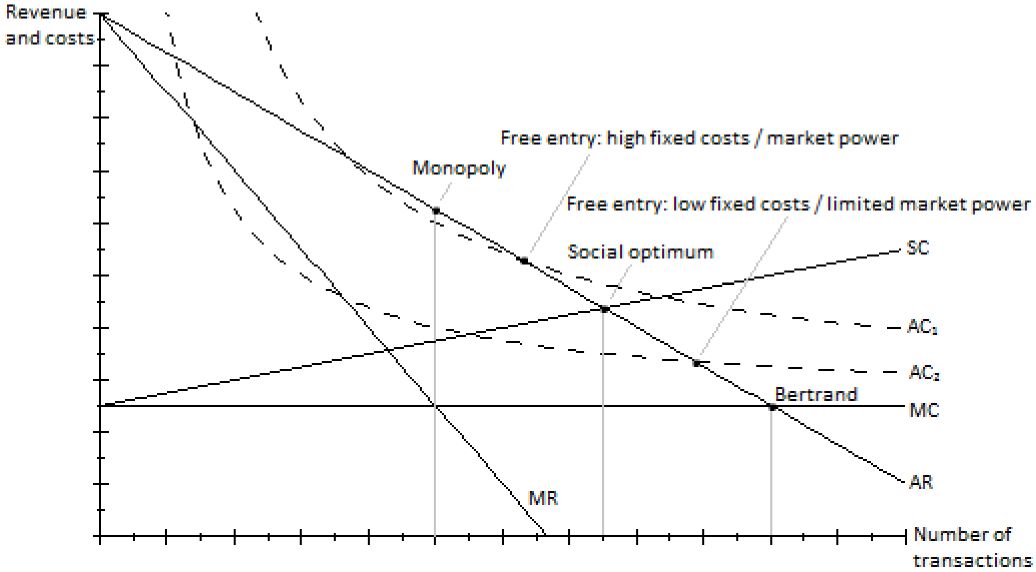
Yinger, J. 1981. "A Search Model of Real Estate Broker Behavior" *American Economic Review*, 71(4):591-604.

Figure 1: Monopoly, Bertrand and social optimum



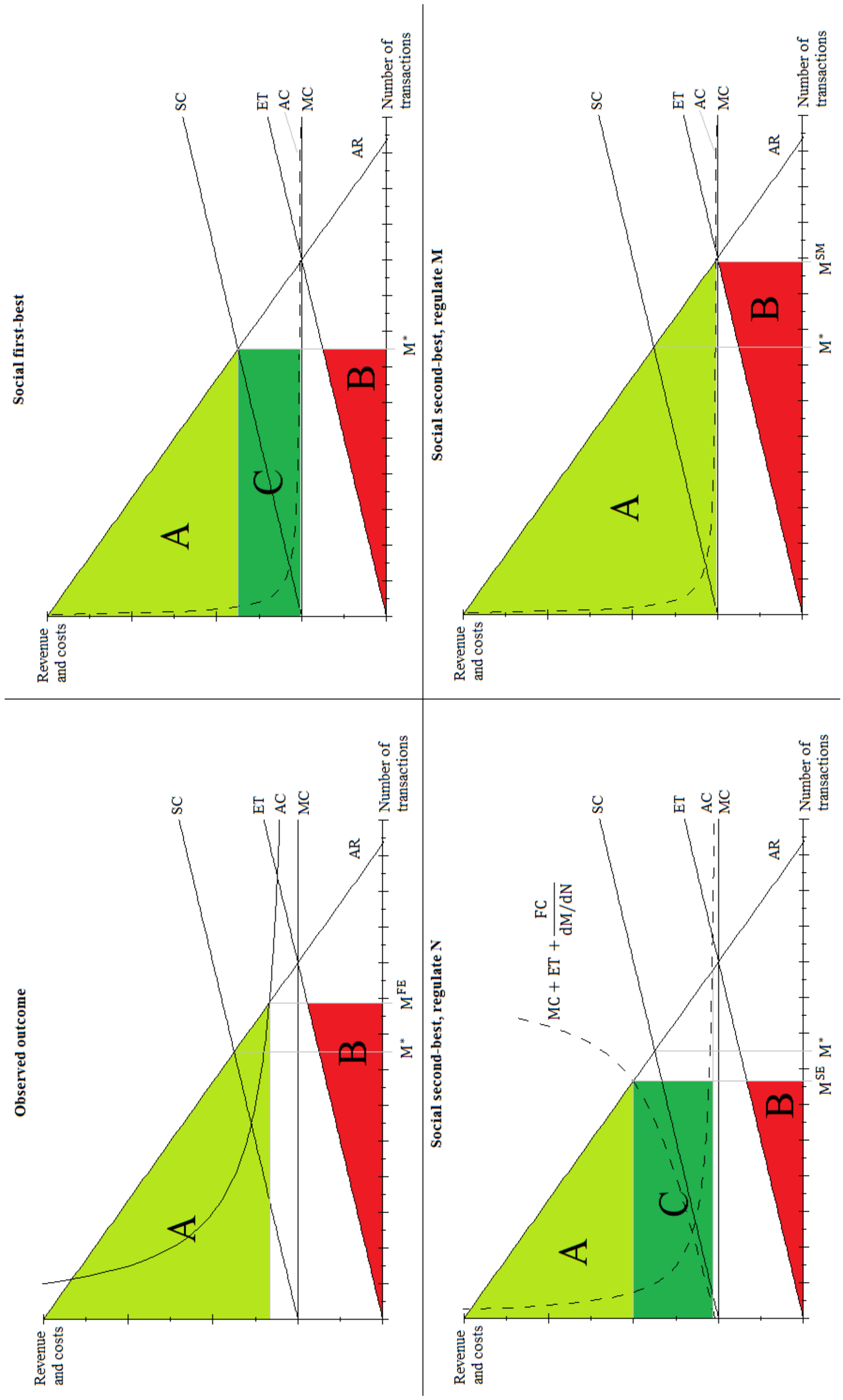
Note: figure 1 assumes that market size is equal to one ($S = 1$), bargaining weights are symmetric ($\beta = 0.5$), seller reservation prices and buyer valuations are uniformly distributed on a unit interval ($s \sim U[0,1]$, $b \sim U[1,2]$) and $MC = 0.5$.

Figure 2: Free entry equilibria



Note: in addition to the assumptions in figure 1, figure 2 assumes $FC = 0.01$ and that $N = 35$ for AC_1 and $N = 15$ for AC_2.

Figure 3: Policy implications



Note: in addition to the assumptions in figure 1, figure 3 assumes for the observed outcome that $FC = 0.01$ and $N = 15$, which implies that broker market power is $\theta = 0.13$. In the social first-best: $AR = SC$ and $N = 1$. In the social second-best when the social planner regulates entry: $AR = MC + ET + FC/(dM/dN)$ where dM/dN follows from the private FOC, imposing Bresnahan (1989)'s constant conjectural variations model $\theta = (1 + R)/N$. In the social second-best when the social planner regulates the number of transactions in the brokerage market: $AR = AC$ and $N = 1$.

Figure 4: Observed and fitted residual sales price distributions

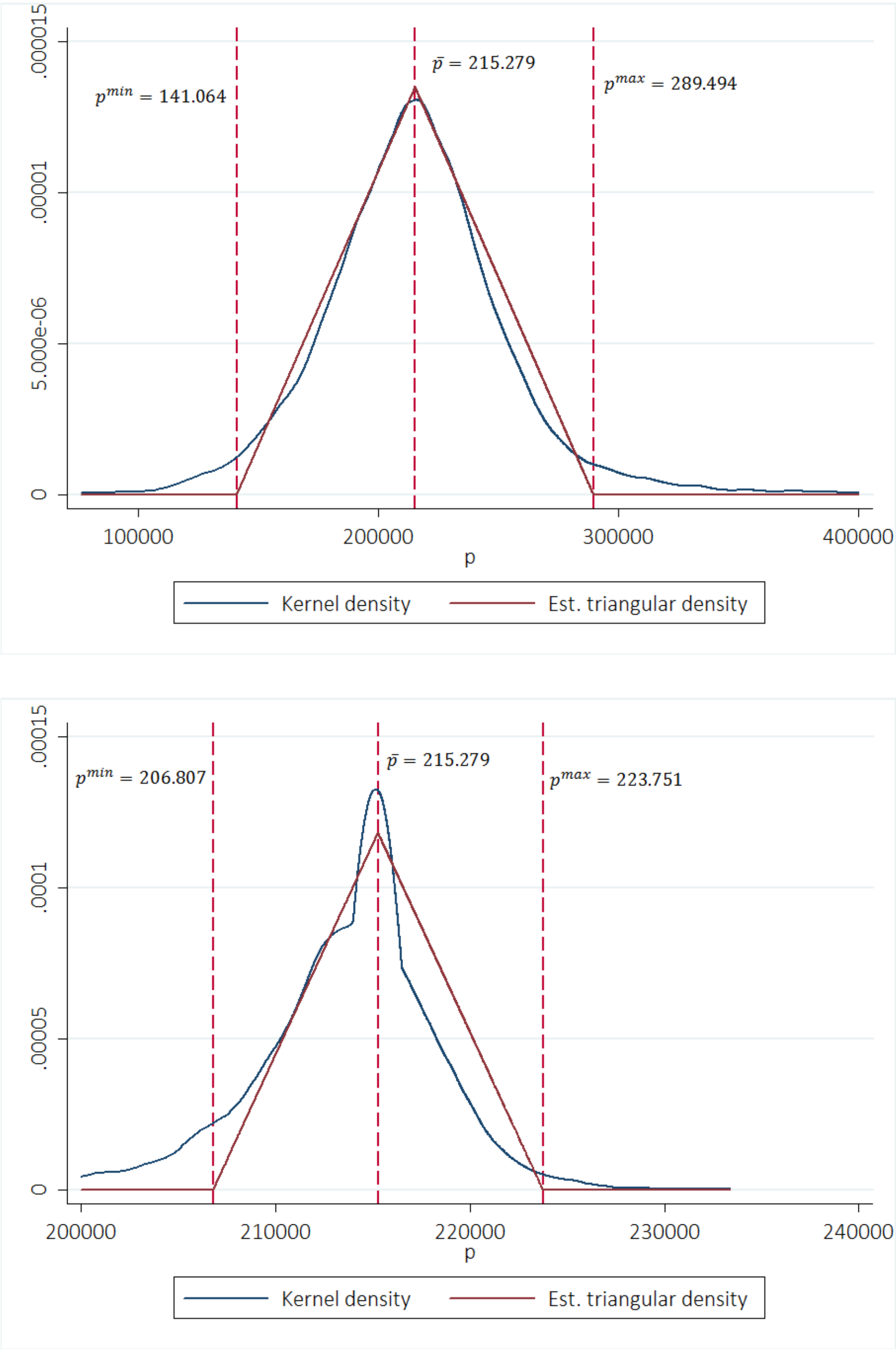


Table 1: Descriptive statistics sales prices and service fees

Description	In model	Obs.	Mean	St. Dev.	P5	P95
Sales price (in €)	p	26,986	215,279	92,763	90,000	392,500
Flat service fee (in €)	T	9,367	2,786	1,844	0	6,050
Proportional service fee	t	9,503	0.030	0.010	0	0.042

Table 2: Market-level data

Description	In model	Mean	Range
Market size	S	123,652	
Market share brokers	M/S	0.56	
# Brokers	N	6,728	4,494 (Immoweb) - 8,963 (BIV)
Marginal costs (in €)	MC	3,339	983 (excl. opp. cost) - 5,696 (incl. opp. cost)
Fixed costs (in €)	FC	60,141	35,882 (excl. opp. cost) – 84,389 (incl. opp. cost)

Table 3: Estimated supply and demand parameters

Parameter	Upper bound	Lower bound
b^L	61,547	197,729
b^H	326,601	227,986
s^L	98,081	193,766
s^H	355,259	223,124

Table 4: Baseline calibration

Variable	Inelastic S & D	Elastic S & D
AR (in €)	9,182	9,182
MC (in €)	3,339	3,339
MS (in €)	146,224	16,692
ET (in €)	73,112	8,346
θ	0.026	0.233
W (in €)	5,062,680,199	577,929,349

Table 5: Welfare counterfactuals

Scenario	Variable	Inelastic S & D	Elastic S & D
Regulate service fee, N=1	M/S	0.431	0.518
	AR	59,634	11,059
	$\bar{p}$	240,885	216,231
	$AR/\bar{p}$	0.237	0.051
	N	1	1
	W^*	6,002,970,151	988,995,085
	Π	3,001,455,005	494,467,472
	CS	3,001,515,146	494,527,613
	W^*/W	1.185	1.711
Regulate entry, flexible service fee	M/S	0.427	0.408
	AR	61,310	15,996
	$\bar{p}$	241,736	218,737
	$AR/\bar{p}$	0.248	0.073
	N	517	2.261
	W^{SM}	5,971,346,023	808,151,940
	Π	3,029,111,538	501,952,042
	CS	2,942,234,485	306,199,898
	W^{SM}/W	1.179	1.398
Regulate service fee, flexible entry	M/S	0.575	0.691
	AR	3,340	3,340
	$\bar{p}$	212,314	212,313
	$AR/\bar{p}$	0.016	0.016
	N	1	1
	W^{SE}	5,335,986,832	879,120,106
	Π	0	0
	CS	5,335,986,832	879,120,106
	W^{SE}/W	1.053	1.521

Table 6: Welfare counterfactuals - sensitivity with respect to MC

Scenario	Variable	$MC = 983$ Inelastic s & d	$MC = 983$ Elastic s & d	$MC = 5{,}696$ Inelastic s & d	$MC = 5{,}696$ Elastic s & d
Regulate service fee, N=1	M/S	0.435	0.557	0.426	0.478
	AR	57,867	9,292	61,402	12,827
	N	1	1	1	1
	W^*	6,129,218,905	1,145,638,920	5,877,982,929	843,805,245
	W^*/W	1.210	1.982	1.161	1.460
Regulate entry, flexible service fee	M/S	0.427	0.399	0.425	0.418
	AR	61,064	16,370	62,038	15,521
	N	701	2,556	316	1,783
	W^{SE}	6,067,995,065	837,666,228	5,866,594,606	766,480,880
	W^{SE}/W	1.198	1.449	1.158	1.326
Regulate service fee, flexible entry	M/S	0.580	0.743	0.568	0.637
	AR	984	984	5,697	5,697
	N	1	1	1	1
	W^{SM}	5,448,21,335	1,018,364,459	5,224,881,689	750,057,080
	W^{SM}/W	1.076	1.762	1.032	1.297

Table 7: Welfare counterfactuals - sensitivity with respect to N

Scenario	Variable	$N = 4{,}494$ Inelastic s & d	$N = 4{,}494$ Elastic s & d	$N = 8{,}963$ Inelastic s & d	$N = 8{,}963$ Elastic s & d
Regulate service fee, N=1	M/S	0.431	0.518	0.431	0.518
	AR	59,634	11,059	59,634	11,059
	N	1	1	1	1
	W^*	6,002,940,254	988,965,188	6,002,985,148	989,010,082
	W^*/W	1.185	1.711	1.185	1.711
Regulate entry, flexible service fee	M/S	0.426	0.407	0.426	0.407
	AR	61,352	15,997	61,357	15,996
	N	345	1,510	688	3,012
	W^{SE}	5,971,346,336	808,151,939	5,971,346,321	808,151,941
	W^{SE}/W	1.179	1.398	1.179	1.398
Regulate service fee, flexible entry	M/S	0.574	0.690	0.574	0.690
	AR	3340	3340	3340	3340
	N	1	1	1	1
	W^{SM}	5,355,966,901	879,100,175	5,335,996,830	879,130,104
	W^{SM}/W	1.053	1.521	1.053	1.521

Table B1: Descriptive statistics property characteristics

Variable	Obs.	Mean	St. Dev.
List price (in €)	26,899	234,002.10	101,431.79
Sales price (in €)	26,986	215,278.89	92,762.61
Days-on-market	26,839	118.69	114.36
Living surface (in m2)	24,758	178.47	63.054
Lot size (in m2)	26,361	726.79	892.6
Year of construction	26,588	1956.09	32.089
# Bedrooms	26,986	3.118	0.945
# Garages	26,909	0.819	0.726
Terraced	26,986	0.332	0.471
Semi-detached	26,986	0.255	0.436
Detached	26,986	0.411	0.492
Terrace	24,260	0.703	0.456
Elevator	13,969	0.003	0.058
Central heating	26,900	0.714	0.451
Heating material: gas	26,901	0.555	0.496
Heating material: electricity	26,901	0.072	0.259
Condensing boiler	26,790	0.045	0.207
Underfloor heating	26,879	0.028	0.165
Glazing: single	26,923	0.362	0.48
Glazing: double	26,923	0.775	0.417
Glazing: triple	26,923	0.005	0.073
Kitchen: luxuriously finished	26,915	0.063	0.244
Kitchen: dishwasher	26,915	0.292	0.454
State: luxuriously finished	26,898	0.049	0.216
State: ready to move in	26,898	0.585	0.492
State: minor refreshments necessary	26,898	0.196	0.397
Various: fireplace in living	23,636	0.123	0.329
Various: alarm	23,636	0.065	0.246
Environment: residential	24,417	0.155	0.362
Environment: villa district	24,417	0.049	0.216
Dist. center village	26,986	0.953	0.867
Dist. Brussels	26,986	61.262	33.465
Travel time to Brussels (in minutes)	26,986	57.159	21.288
Dist. nearest city	26,986	14.024	9.89
Dist. highway	26,986	5.484	5.337
Dist. train station	26,986	3.833	3.667
Year of sale	26,983	2010.21	2.673

Note: The sample of brokered real estate transactions contains besides information on prices (sales and listing price) and liquidity (time-on-market) also a very detailed description of the features of every property. The characteristics reported do not only describe the size of every dwelling (terraced vs. (semi-)detached, interior space, lot size, # bedrooms, # garages, # bathrooms), but also provide detailed information concerning the heating system (type (central heating), material (gas, electricity,…), elements (underfloor heating, accumulators,…)), isolation (single vs. double vs. triple glazing), state of the dwelling (ready to move in, luxuriously finished,…), and its environment (residential, villa districts,…) and location (distance to different amenities, major cities). For several rooms, such as the kitchen (well-maintained, dishwasher, ceramic stove,…), bathroom (bath, shower,…) and basement (wine cellar,…), the realtor registered the features present.

Table B2: Descriptive statistics buyer & seller characteristics

	Buyers			Sellers		
Variable	Obs.	Mean	St. Dev.	Obs.	Mean	St. Dev.
Avg. age pop.	11,454	40.69	3.81	26,918	40.65	3.49
% married	11,444	0.52	0.12	26,897	0.54	0.1
Avg. size household	11,444	2.34	0.31	26,897	2.4	0.26
Med. tax. Income (in €)	11,424	21,837	3819.94	26,822	22,577	3861.98
% higher education	11,480	0.28	0.08	26,980	0.27	0.08
% Owners	11,480	0.68	0.18	26,980	0.72	0.15

Note: To construct measures of buyer and seller characteristics administrative data are used at the level of statistical sectors in Belgium. The 19,781 statistical sectors in Belgium are the lowest territorial level at which Statistics Belgium gathers information and, on average, have a surface of 1.5km^2, and are home to approximately 550 inhabitants and/or 240 households. Given that we know the exact location of every dwelling and the previous address for a subsample of buyers, we can spatially join the respective x- and y-coordinates with the appropriate statistical sectors using the spatial join module in Quantum GIS. From Statistics Belgium we retrieved yearly data on different demographic variables and taxable incomes for every statistical sector. We either observed or managed to calculate the average age of the population, the percentage of reference persons of households who are married, the average size of a household and the median taxable income for every local living area. From the Census 2011 we furthermore retrieved the percentage of the population that finished higher education (where higher education is defined as a university degree or higher) and the percentage of owner-occupied houses in the total housing stock.

Table B3: Marginal costs

What?	Details	Costs (range)	Costs (mean)	Source
Listing	*www.immoweb.be*	€100-€150	€125	*www.immoweb.be*
Other promotional activities	"For sale" sign, advertisement in local newspapers	100	€100	Own estimate
Special information duty	Building permits,...	€20 - €100	€60	*www. okra.be*
Certificates	Energy Performance	€150-€450 (€200-€600, according to www.immoweb.be)	€350	
	Electricity	€120	€120	
	Soil	€50	€50	
	Oil fuel tank	€65-€225	€145	
Title of land		€25	€25	
Information from property/charge registers	Mortgage	€16.50	€16.50	*www.kadaster.be*
	Cadaster	€16.50	€16.50	
Total:			€983	

Note: Table B3 provides an overview of the monetary costs incurred by real estate agents when selling a property. Information from www.immoweb.be, the largest online listing service in Belgium, suggests that a listing costs between €100 and €150. Other online listing platforms in Belgium are usually free of charge. We assume that other promotional activities, such as a "for sale" sign and advertisement in local newspapers and so on, cost another €100. Since real estate agents also help sellers gather the necessary documents these costs are also initially incurred by the real estate agent. From www.okra.be and www.immoweb.be we learned that these costs are in total between €430 and €1120. Information from property registers finally contribute another €33. Given all these costs we calculate a monetary per-match cost of €983 for a representative transaction.

Table B4: Regression results

Variable	Inelastic S & D	Elastic S & D
Semi-detached	7,479***	4,903***
	(790.6)	(1,183)
Detached	19,948***	17,423***
	(1,040)	(1,593)
Living surface	422.5***	363.2***
	(30.55)	(48.63)
Living surface sq.	-0.255***	-0.164
	(0.0731)	(0.119)
Lot size	45.18***	48.57***
	(1.368)	(2.078)
Lot size sq.	-0.00427***	-0.00472***
	(0.00026)	(0.00036)
Terrace	8,480***	6,797***
	(757.7)	(1,144)
Central heating	8,557***	7,635***
	(1,271)	(1,874)
Condensing boiler	9,407***	10,747***
	(1,535)	(2,611)
Underfloor heating	20,322***	20,463***
	(2,243)	(3,561)
Glazing: single	-9,799***	-10,387***
	(754)	(1,126)
Glazing: double	707.3	1,493
	(832.8)	(1,236)
Glazing: triple	10,687***	15,627***
	(3,801)	(5,471)
Kitchen: luxuriously finished	10,450***	8,550***
	(1,564)	(2,299)
Kitchen: dishwasher	11,318***	12,020***
	(775.1)	(1,207)
State: luxuriously finished	27,548***	25,673***
	(1,982)	(2,952)
State: ready to move in	14,842***	14,946***
	(771.9)	(1,152)
State: minor refreshments necessary	-2,267***	-3,917***
	(736.9)	(1,146)
Various: fireplace in living	6,846***	5,992***
	(1,043)	(1,635)
Various: alarm	20,341***	18,301***
	(1,583)	(2,632)
Environment: residential	15,673***	14,679***

	(1,068)	(1,703)
Environment: villa district	18,145***	13,731***
	(1,905)	(2,941)
Dist. highway	-195.1	-655.1*
	(234.4)	(347.7)
Dist. train station	-305.3	-208.5
	(250.4)	(385.6)
Dist. Brussels	-1,472***	-655
	(390.7)	(639.9)
Dist. Brussels sq.	6.689**	5.699
	(2.993)	(4.819)
Observations	18,812	8,083
R-squared	0.828	0.844
# Bedrooms	YES	YES
# Garages	YES	YES
Building period	YES	YES
Other quality controls	YES	YES
Other location controls	YES	YES
Municipality FE	YES	YES
Broker FE	YES	YES
Year-District FE	YES	YES
Buyer-Seller characteristics	NO	YES

$* = 10\%$, $** = 5\%$ and $*** = 1\%$ significance level

Note: Table B4 presents the regression results. Whereas most hedonic house price analyses use a log-transformed dependent variable, house prices are not log-transformed here since the purpose is for the residuals to capture the price spread (in euros). For some independent variables, such as interior space and lot size, the regression is therefore augmented with a quadratic term to capture possible nonlinearities. The first column of table B4 presents the estimated coefficients for the hedonic pricing regression (30) without buyer and seller characteristics. Almost all coefficients show the expected signs and are statistically significant. For example, the sales price of a dwelling is positively related to its interior space and lot size, but an additional square meter is less valuable for a large dwelling than for a smaller one. Also observe that (semi-)detached houses are more expensive than terraces ones. Furthermore, note that all the characteristics that relate to the quality of the property show their expected signs.[23]

Following Harding, Rosenthal and Sirmans (2003), in the second column of table B4 buyer and seller characteristics are included in the regression analysis. Unlike Harding, Rosenthal and Sirmans (2003), however, these variables are not observed at the individual level. Instead, their local living area counterparts are used as a proxy variable, as described in table B2. In addition, Table B5 reports the regression coefficients for the "bargaining effect" (α) and the "demand effect" (δ) in expression (31). There are significant positive demand effects from the percentage of people with a college education or higher, the percentage of people that is married, and the average age of the population. There is a significantly negative demand effects from the percentage of owners. The only significant bargaining effect comes from the percentage of people that enjoyed higher education, which suggests that the price of housing is increasing whenever the level of education of the seller is relatively high compared to that of the buyer.

[23] Recall from table B1 that our dataset also contains information on listing prices and time-on-market, which are not included as regressors. Although the (initial) listing price and the time-on-market might seem suitable to control for the effect of unobservables, they are also likely to be correlated with seller characteristics, as shown by Genesove & Mayer (2001), for example, which makes them unsuitable to estimate a proper upper bound for the spread in sales prices that can be attributed to buyer and seller characteristics.

Table B5: Bargaining and demand effects

Variable	Bargaining/ Demand	
Distance buyer-property	–	-53.79*
		(29.56)
% Higher education	α	47,868***
		(7,084)
	δ	63,127***
		(6,876)
% Owners	α	-6,002
		(4,688)
	δ	-14,544***
		(4,631)
% Married	α	5,664
		(11,155)
	δ	23,878**
		(10,410)
Avg. size household	α	1,121
		(3,946)
	δ	-5,762
		(3,700)
Avg. age population	α	252.3
		(162)
	δ	410.0***
		(155)
Ln(med. tax. inc.)	α	-5,422
		(4,456)
	δ	424.6
		(4,682)
Male	α	-17,144
		(11,658)
	δ	-12,111
		(10,412)

$* = 10\%$, $** = 5\%$ and $*** = 1\%$ significance level