GOVERNANCE AND THE EFFICIENCY
OF ECONOMIC SYSTEMS
G E S Y

Discussion Paper No. 435

Assessing the Performance of Simple Contracts Empirically: The Case of Percentage Fees

Simon Loertscher*
Andras Niedermayer**

* University of Melbourne
** University of Mannheim

September 20, 2012

Assessing the Performance of Simple Contracts Empirically: The Case of Percentage Fees[*]

Simon Loertscher[†]   Andras Niedermayer[‡]

September 20, 2012

Abstract

This paper estimates the cost of using simple percentage fees rather than the broker optimal Bayesian mechanism, using data for real estate transactions in Boston in the mid-1990s. This counterfactual analysis shows that intermediaries using the best percentage fee mechanisms with fees ranging from 5.4% to 7.4% achieve 85% or more of the maximum profit. With the empirically observed 6% fees intermediaries achieve at least 83% of the maximum profit and with an optimally structured linear fee, they achieve 98% or more of the maximum profit.

Keywords: brokers, simple mechanisms, percentage fees, real estate brokerage.

JEL-Classification: C72, C78, L13

1   Introduction

Real world economic agents often employ simple mechanisms. Examples include uniform pricing by iTunes, cost-reimbursement contracts in procurement, and percentage fees employed by credit card companies and real estate brokers.[1]   On its

[†]Department of Economics, Economics & Commerce Building, University of Melbourne, Victoria 3010, Australia. Email: simonl@unimelb.edu.au.

[‡]Economics Department, University of Mannheim, L7, 3-5, D-68131 Mannheim, Germany. Email: aniederm@rumms.uni-mannheim.de.

[1]See, respectively, Shiller and Waldfogel (2011), Rogerson (2003), Shy and Wang (2011) and Hsieh and Moretti (2003).

surface, at least, this simplicity contrasts with the prescriptions and predictions of Bayesian mechanism design, which suggests that optimal mechanisms will in general be rather sophisticated. For example, percentage fees are optimal mechanisms for brokers if and only if the supply function brokers face is isoelastic (Loertscher and Niedermayer, 2012), begging the question whether the choice of mechanisms is, in reality, primarily driven by concerns of simplicity and practicality rather than guided by insights uncovered by economic theory.

In this paper, we investigate how much profit is sacrificed by the use of simple mechanisms rather than optimal ones in the case of real estate brokerage. Using a structural model and the data set of Genesove and Mayer (2001), which was generously provided to us by the authors, our counterfactual analysis shows that intermediaries who employ the best proportional fee mechanisms with fees ranging from 5.4% to 7.4% achieve 85% or more of the maximum profit. With the empirically observed 6% fees intermediaries achieve at least 83% of the maximum profit and with an optimally structured linear fee, they achieve 98% or more of the maximum profit. Seemingly very little profit is lost using very simple mechanisms, which suggests that concerns of practicality and economic principles may be well aligned.

We also make two methodological contributions. First, we show how the combination of different numerical techniques can reduce computational time by several orders of magnitude and hence make the estimation computable in a practically useful amount of time. Second, our results suggest that for future research a simpler approach can be taken: a family of functions – Generalized Pareto distribution functions – turn out to be a good approximation of the seller's supply function. This family of functions allows for a closed-form solution for the problem at hand.[2]

The fact that simple mechanisms can achieve a large percentage of the optimal surplus or profit has been shown in a variety of contexts. McAfee (2002), Rogerson (2003) and Chu and Sappington (2007) provide theoretical analysis of simple mechanisms

---

[2]Our theoretical results (Loertscher and Niedermayer, 2012) show that Generalized Pareto distributions are a necessary and sufficient condition for linear fees to be exactly optimal. Therefore, we could not use this simplifying functional form for the current analysis whose purpose is to find out how well linear fees perform relative to optimal fees.

for assortative matching, incentives in procurement, and cost-sharing, respectively.[3] Recent empirical work includes Shiller and Waldfogel (2011), who demonstrate that uniform pricing by iTunes is close to optimal, and Chu, Leslie, and Sorensen (2011), who study a simple form of bundling, called bundled-size pricing, with an application to the pricing of theater tickers. To the best of our knowledge, the present paper is the first to quantify the performance of simple fee setting mechanisms by brokers, with an application to the real estate brokerage industry. By providing empirically based measures that quantify how well the simple mechanisms employed by real estate brokers fare, our paper also contributes to the literature on real estate and real estate brokerage, which has witnessed an upsurge of interest over the past decade; see, for example, Genesove and Mayer (1997, 2001), Hsieh and Moretti (2003), Hendel, Nevo, and Ortalo-Magné (2009), and Genesove and Han (2012).

While our question and setup are quite different, our paper has similarities to the empirical work on auctions, such as Donald and Paarsch (1993), Bajari (1997), Bajari and Hortaçsu (2003), and Shneyerov (2006) because we model the bargaining procedure as an auction. The similarity to Bajari and Hortaçsu (2003) goes even further: the auctions for which they estimate the optimal reserve prices are run by a profit maximizing intermediary – eBay – that also charges a transaction fee. Some of our methods should be applicable in a modified way to analyze the fee setting behavior of intermediaries in setups beyond real estate brokerage.

The remainder of this paper is structured as follows. Section 2 introduces the theory. The empirical analysis and results are described in Section 3. Sections 4 and 5 discuss results and methodological contributions. Section 6 concludes. The appendix contains additional results and a description of the numerical methods.

2   Model

The theoretical model is a slight generalization of the dynamic random matching model analyzed in Loertscher and Niedermayer (2012, LN hereafter).[4] LN have the

---

[3]Holmstrom and Milgrom (1987) analyze the optimality of simple, linear incentive schemes in an intertemporal principal-agent model.

[4]The model underlying our empirical analysis is more general in that we allow for many buyers to be matched to any broker-seller pair. We also allow for a more general bargaining procedure than

following stage game for buyers viewing a real estate property that a seller offers for sale through a broker. Every buyer's reservation price – the maximum amount he is willing to pay – v is his private information and drawn independently from the distribution F, which is common knowledge. The seller's opportunity cost of selling – the minimum amount he is willing to accept – c is also private information and drawn from the commonly known distribution G. The densities of the distributions are f and g with the respective supports $[\underline{v}, \overline{v}]$ and $[\underline{c}, \overline{c}]$. LN model these distributions as endogenous outcomes of a larger market interaction: both traders take into account their option value of trading with other potential trading partners. Further, high cost sellers and low valuation buyers must spend more time searching for a trading opportunity than others and are hence overrepresented in the market relative to the distribution of entrants. Since our estimation of the parameters of the model is agnostic about the market equilibrium concept that generates these endogenous distributions F and G, we will first describe the predictions of the model based directly on the endogenous distributions. We will need to make some (limited) assumptions about the equilibrium concept once we perform the counterfactual analysis.

2.1 Deriving the Data Generating Process

As Hsieh and Moretti (2003) note, the fees charged by brokers are almost always very close to 6%. As an approximation of the bargaining process, we model the behavior of market participants in the following way. The seller sets an optimal reserve price p for an auction with a 6% fee. The buyers participate in a second price auction. If trade occurs, the broker receives 6% of the transaction price. The exposition becomes simpler by defining some additional functions.

Define the virtual valuation function Φ and the virtual cost function Γ as

$$\Phi(v) := v - \frac{1 - F(v)}{f(v)} \quad \text{and} \quad \Gamma(c) := c + \frac{G(c)}{g(c)}.$$

Further, define the (price) elasticity of demand η_d and the (price) elasticity of supply

the take-it-or-leave-it offers underlying most of the theoretical analysis in LN.

4

η_s as[5]

$$\eta_d(v) := \frac{vf(v)}{1-F(v)} \quad \text{and} \quad \eta_s(c) := \frac{cg(c)}{G(c)}.$$

Observe that $\Phi(v) = v(1 + 1/\eta_d(v))$ and $\Gamma(c) = c(1 + 1/\eta_s(c))$. Let $\omega(\hat{p})$ be the fee charged by the broker, leaving $\hat{p} - \omega(\hat{p})$ to the seller if a transaction occurs at price $\hat{p}$. For the empirically observed fees of 6%, $\omega(\hat{p}) = 0.06\hat{p}$. Assume that $k \geq 1$ buyers with valuations that are independently drawn from F participate in a second price auction where the reserve price p is set by the seller. The seller's problem is then to choose p to maximize

$$k(p-\omega(p))(1-F(p))F(p)^{k-1}+k\int_p^{\overline{v}}(y-\omega(y))(1-F(y))(k-1)F(y)^{k-2}f(y)dy+cF(p)^k.$$
(1)

The above equation is the same as in standard auction theory, except that the seller has to pay the fee $\omega(\cdot)$. The first term in the sum is the seller's profit if only one buyer bids above the reserve price p, in which case he receives $p - \omega(p)$. The second term is the expected net price $y - \omega(y)$ given by the second highest bid if two or more buyers bid above p. The third term represents the case in which no one bids above p. For $\omega(\hat{p}) = 0.06\hat{p}$, we have $\hat{p} - \omega(\hat{p}) = 0.94\hat{p}$, so that the problem simplifies to maximizing

$$0.94\left\{kp(1-F(p))F(p)^{k-1} + k\int_p^{\overline{v}}y(1-F(y))(k-1)F(y)^{k-2}f(y)dy + \frac{c}{0.94}F(p)^k\right\}$$

over p. Solving the first-order condition yields

$$p = \Phi^{-1}(c/0.94),$$
(2)

which is independent of k, reflecting a result that is well-known from auction theory. Therefore, it does not matter whether the seller knows the number of buyers at the time he sets the reserve price.

If a buyer and a seller trade, they leave the market. If there is no trade, the buyer and seller stay in the market until the next rematching, which happens after time τ

---

[5]To see that these are the price elasticities, interpret the probability that a buyer is willing to buy and a seller is willing to sell at price p as quantity demanded and quantity supplied, which are denoted $q^d(p) := 1 - F(p)$ and $q^s(p) := G(p)$, respectively. The price elasticities being defined as $-q^{d'}(p)p/q^d(p)$ and $q^{s'}(p)p/q^s(p)$, $\eta_d(p) = pf(p)/(1 - F(p))$ and $\eta_s(p) = pg(p)/(G(p))$ follows immediately after differentiation and substitution.

has passed. With exogenous probability $1 - \epsilon$ a seller or buyer drop out of the market without trading. The dropout probability is a simple device to take into account a variety of sources of impatience, for example, traders may have deadlines regarding when they want to move or buyers may have to rent temporarily at a high price if they do not buy.

The number of buyers who visit a seller in a given period is a random variable that follows a Poisson process with arrival rate $\lambda := \lambda_0 N_B / N_S$, where $\lambda_0 \in (0, 1]$ is the probability that a randomly picked buyer has a preference for this seller's house and N_B and N_S are the masses of buyers and sellers in the market, respectively.[6] The Poisson distribution of the number of buyers is founded in random matching theory. The probability $\pi_k(\lambda)$ of being matched to k buyers with $k = 0, 1, 2, ..$ is then $\pi_k(\lambda) = e^{-\lambda} \lambda^k / (k!)$. Accordingly, the probability that a seller who sets the reserve price p does not sell in a given period is

$$\hat{F}(p) := \sum_{k=0}^{\infty} \pi_k(\lambda) F(p)^k = e^{-\lambda(1 - F(p))}. \tag{3}$$

Note that (3) implies that $F(p) = 1 + \ln(\hat{F}(p))/\lambda$. This allows us to express the virtual valuation $\Phi(v)$ without λ and F and in terms of $\hat{F}$ only:

$$\Phi(v) = v + \frac{\hat{F}(v)}{\hat{f}(v)} \ln(\hat{F}(v)),$$

where $\hat{f}$ is the density of $\hat{F}$.

The time on market t of a property has a geometric distribution with the distribution function $1 - (\epsilon \hat{F}(p))^{t/\tau}$ and mean

$$T(p) = \frac{\tau}{1 - \epsilon \hat{F}(p)}. \tag{4}$$

The probability that a property sells in period t is $(1 - \hat{F}(p))(\epsilon \hat{F}(p))^{t/\tau}$. Taking the sum of this geometric series over t from 0 to infinity gives us the probability $1 - \hat{F}_\infty(p)$ that a house that is offered at price p is ever sold:

$$1 - \hat{F}_\infty(p) = \frac{1 - \hat{F}(p)}{1 - \epsilon \hat{F}(p)}. \tag{5}$$

---

[6]We model heterogeneity of houses and preferences in a simple, binary way. With probability λ_0, a buyer has a preference for a given house and values it according to v drawn from $F(v)$, and with probability $1 - \lambda_0$ she does not like the house in the sense that her valuation of it is 0.

When taking the model in LN to the data, we make two modifications of the baseline model in addition to those on matching and bargaining. First, the price p is a quality-adjusted reserve price, which is observed with an error. Second, time on market is observed with an error. The details of how we model these errors are spelled out in Section 3.2.

2.2 Model with Optimal Fee Setting

We can now consider the model for the counterfactual analysis. We need one additional assumption for this: when the broker proposes the exchange mechanism for the buyers and the seller, he faces the competitive threat that the buyer and seller may go to a different broker in the future if they do not like the mechanism he offers. We model this competitive threat in the spirit of Burdett and Mortensen (1998) in the following way, with similarities to Wolinsky (1986)'s and Anderson and Renault (1999)'s models of imperfect competition. The broker can only extract rents from the current trade, but not from future trades.[7] Hence, the broker's mechanism depends directly on the equilibrium distributions F and G for the current trade, which he takes as exogenously given. This assumption means that the broker does not try to change the distributions F and G by, for example, reducing the option value of future trade or by letting certain types of traders cumulate more in the market.

The broker designs a mechanism that maximizes his own profits given the incentive compatibility and individual rationality constraints of participants. Arguments similar to those in LN show that no mechanism exists that generates higher profits than a fee setting mechanism, which is defined as follows. The broker announces a fee function $\omega(p)$, then the seller sets a reserve price p, and an arbitrary number of buyers, which is drawn from a Poisson distribution with arrival rate λ, participate in a second price auction.[8] If there is trade, the winning buyer pays the transaction

---

[7]Alternatively, one could allow the level of competition between brokers to vary. Appendix C in LN provides a model that has the required flexibility. In the spirit of Burdett and Mortensen (1998), it assumes that every seller who is matched to a broker in a given period does not stay with the same broker in the next period with probability $\nu \in [0, 1]$. The present model corresponds to the case $\nu = 1$.

[8]English auctions, which are strategically equivalent to second price auctions in our environment with private values, appear to be a good description of the bargaining between sellers and buyers in real estate markets. Whenever a seller (or his broker) receives a new offer, he will advise the bidder

price $\hat{p}$, which is given by the maximum of the second highest bid and the seller's reserve price p. Therefore, the seller nets $\hat{p} - \omega(\hat{p})$ and the broker nets $\omega(\hat{p})$ if trade occurs. If there is no trade, the payoff is zero for all players in the current period. The optimal fee function is

$$\omega(p) = p - E_v[\Gamma^{-1}(\Phi(v))|v \geq p] = p - \int_p^{\overline{v}} \Gamma^{-1}(\Phi(v)) \frac{[\ln \hat{F}(v)]'}{\ln \hat{F}(p)} dv, \tag{6}$$

which induces a seller with cost c to set the price

$$P(c) := \Phi^{-1}(\Gamma(c)), \tag{7}$$

as shown in Proposition 1 in LN. It can also be shown that an overall increase in the elasticity of supply $\eta_s(c)$ leads to an overall decrease in fees $\omega(p)$ and, through this channel, to an overall decrease in the gross prices $P(c)$. An overall increase in the elasticity of demand $\eta_d(v)$ leads to a decrease in prices $P(c)$, which has an ambiguous effect on fees.

Constant elasticities of supply are helpful to obtain further insights about fees. If the seller's cost has the distribution $G(c) = c^\alpha$, the elasticity of supply is constant with $\eta_s(c) = \alpha$ and the optimal fee is proportional to the price: $\omega(p) = p/(1 + \alpha)$. A similar, more general observation is that for mirrored Generalized Pareto distributions[9] $G(c) = [(c - \underline{c})/(\overline{c} - \underline{c})]^\alpha$ the optimal fee function is linear, $\omega(p) = p/(1 + \alpha) - \underline{c}/(1 + \alpha)$ (see Proposition 4 in LN). Note that for linear virtual cost distributions, the fees are determined solely by the seller's elasticity of supply and are independent of the buyer's distribution.

Factors driving optimal fees towards linearity Beyond being analytically convenient, one may wonder whether there are reasons to expect mirrored Generalized

---

with the current highest bid. With the empirically observed proportional fees, the same results are obtained assuming a first price auction even if the buyers are not aware of the number of competitors when placing their bids as long as all buyers are ex ante equally well (or equally poorly) informed of this number; see Krishna (2002, p. 34) for details. However, for the counterfactual analysis that uses the broker optimal fee, which will typically not be proportional, additional structure would have to be imposed for first price auctions.

[9]These are equivalent to distributions with linear virtual cost functions, which have linear inverse hazard rates $G(c)/g(c)$. Here we focus on mirrored Generalized Pareto distributions with a finite lower bound. A mirrored Generalized Pareto distribution with an infinite lower bound is a (mirrored) exponential distribution $G(c) = \exp(-\alpha(\overline{c} - c))$.

Pareto distributions and the linear optimal fees they imply to be plausible descriptions of reality. Three forces drive optimal fees towards linearity. LN identify two forces that lower the elasticity of supply and make it "more constant". The first one arises in a static setup (i.e. when $\epsilon = 0$) and the second one results from dynamics. To understand the static effect, assume that the seller's opportunity cost of selling stems from price offers for his property that are generated outside the broker's platform. If the seller receives one outside price offer c drawn from $G(c)$, the implied elasticity of supply is $\eta_s(c) = cg(c)/G(c)$. If the seller gets $\alpha > 1$ i.i.d. outside price offers drawn from G, then the distribution of the best offer is $G(c)^\alpha$, which implies elasticity $\alpha\eta_s(c) > \eta_s(c)$. This clearly lowers fees and makes them "more linear" in the following sense. Let $\omega_\alpha(p)$ be the optimal fee function associated with the distribution $G(c)^\alpha$. This function $\omega_\alpha(p)$ lies within a concave upper bound $\overline{\omega}_\alpha(p)$ and a convex lower bound $\underline{\omega}_\alpha(p)$, and the difference $\overline{\omega}_\alpha(p) - \underline{\omega}_\alpha(p)$ decreases in α and goes to 0 for all p as α goes to infinity (Proposition 2 in LN).

The dynamic effect arises because high cost sellers set higher prices and have a lower probability of sale. Hence higher cost sellers are overrepresented in the market compared to the entrant population. This transformation from inflow to stock distributions increases the density of sellers with higher costs. The dynamic cumulation effect increases the elasticity of demand and lowers fees. It also makes the inverse elasticity $1/\eta_s(c)$ "more constant" and fees "more linear" (see Propositions 6 and 7 and the numerical results in Section 3.3 in LN), with a variety of measures being used for "closeness of the fees to linearity": the absolute curvature of the virtual cost function $|\Gamma''(c)|$, the distance between the convex hull and the concave hull of Γ, and the ratio of profits generated by the best linear fee and by the optimal (non-linear) fee.

A further effect is based on Extreme Value Theory in statistics (see, for example, Coles (2001)). If only the most motivated sellers enter the market (only those whose costs are below an upper threshold c^*), the distribution of sellers in the market is $G(c)/G(c^*)$. The second theorem of extreme value theory (Pickands-Balkema-de Haan theorem) states that as c^* goes to the lower bound of the support $\underline{c}$, the trun-

cated distribution $G(c)/G(c^*)$ converges to a Pareto distribution (after an appropriate rescaling). This asymptotic result holds for arbitrary G satisfying some weak regularity conditions.[10]

Optimal, linear and proportional fees We now bring the theoretical model to the data and turn to the empirical question of how brokers fare when they employ linear or proportional fees rather than optimal fees. The optimal profit of an intermediary given k buyers is

$$E_{v_i,c}[\max_{i=1..k}\{\Phi(v_i) - \Gamma(c), 0\}] = E_{\tilde{v}\sim F(\tilde{v})^k,c}[\max\{\Phi(\tilde{v}) - \Gamma(c), 0\}]$$

Taking expectations over k we obtain

$$
\begin{aligned}
\Pi_{\text{optimal}} &= \sum_{k=0}^{\infty} \int \int \pi_k(\lambda) \max\{\Phi(\tilde{v}) - \Gamma(c), 0\} dG(c)[F(\tilde{v})^k]' d\tilde{v} \\
&= \int \int \max\{\Phi(\tilde{v}) - \Gamma(c), 0\} dG(c)[\sum_{k=0}^{\infty} \pi_k(\lambda) F(\tilde{v})^k]' d\tilde{v} \\
&= \int \int \max\{\Phi(\tilde{v}) - \Gamma(c), 0\} dG(c)[\hat{F}(\tilde{v})]' d\tilde{v}, \quad\quad (8)
\end{aligned}
$$

where the second equality follows by moving the sum into the brackets and the third makes use of the definition of $\hat{F}$ in (3). This means that we can compute Π_{optimal} without knowledge of λ and F, provided we have $\hat{F}$ and G. This result is important for our analysis because we can identify $\hat{F}$ but not λ and F separately from the data, as discussed below.

The expected profit of a seller with cost c and reserve price p facing $k \geq 1$ buyers is given in (1) for arbitrary $\omega(\cdot)$ (and not only the optimal fee function). Taking expectations over k we obtain

$$
\begin{aligned}
&(p - \omega(p)))(1 - F(p))[\sum_{k=1}^{\infty} \pi_k k F(p)^{k-1}] \\
&+ \int_p^{\overline{v}} (y - \omega(y))(1 - F(y))[\sum_{k=1}^{\infty} \pi_k k F(y)^{k-1}]' dy + \sum_{k=0}^{\infty} c F(p)^k \\
&= (p - \omega(p)))(-\ln \hat{F}(p))\hat{F}(p) + \int_p^{\overline{v}} (y - \omega(y))(-\ln \hat{F}(y))\hat{f}(y) dy + c\hat{F}(p). \quad (9)
\end{aligned}
$$

---

[10]Formally, the distribution has to be in the domain of attraction of an extreme value distribution. As noted in the literature on extreme value theory, all continuous "textbook" distributions satisfy these conditions. Examples are the normal, exponential, Cauchy, Beta, and uniform distributions. See Coles (2001) for more details.

Note that for $k \geq 1$, $\pi_k k = \lambda \pi_{k-1}$, and therefore $\sum_{k=1}^{\infty} \pi_k k F(p)^{k-1} = \lambda \sum_{k=0}^{\infty} \pi_k k F(p)^k = \lambda \hat{F}(p)$. Replacing $\sum_{k=1}^{\infty} \pi_k k F(p)^{k-1}$ by $\lambda \hat{F}(p)$ and $1 - F$ by $-\ln(\hat{F})/\lambda$, we obtain the equality. Therefore, the seller's expected profit can be written in terms of $\hat{F}$ only, irrespective of λ and F. The seller's optimal reserve satisfies the first-order condition

$$0 = [(p - \omega(p) - c)(-\ln \hat{F}(p))]'.$$

In addition, the intermediary's expected profit as a function of p given an arbitrary fee function ω will also be independent of λ and F as it is simply

$$\Pi(p, \omega) = \omega(p)(-\ln \hat{F}(p))\hat{F}(p) + \int_p^{\overline{v}} \omega(y)(-\ln \hat{F}(y))\hat{f}(y)dy, \tag{10}$$

which can be derived in the same way as (9). Our counterfactual analysis is concerned with the intermediary's profit and how the seller's maximization problem changes as fees change. Since these profits can be expressed in terms of $\hat{F}$ alone, we only need to identify $\hat{F}$ rather than λ and F separately.

For a linear fee $\omega_{\xi,\zeta}(p) = \xi p + \zeta$ the seller's optimal reserve price is $P_{\xi,\zeta}^{-1}(c) = \Phi^{-1}((c + \zeta)/(1 - \xi))$, which can be derived from the first-order condition. The intermediary's expected profit (taking expectations also over prices) is

$$\Pi_{\xi,\zeta} = E_c[\Pi(P_{\xi,\zeta}(c), \omega_{\xi,\zeta})]. \tag{11}$$

Maximizing with respect to the slope ξ and the intercept ζ of the fee yields the best linear profit $\Pi_{\text{linear}} = \max_{\xi,\zeta} \Pi_{\xi,\zeta}$. Note that under a linear fee, the seller's optimal reserve is $P_{\xi,\zeta}(c) = \Phi^{-1}((c + \zeta)/(1 - \xi))$. In addition to the performance of linear fees, we will estimate the profits generated by the best proportional fee function $\Pi_{\text{proportional}} = \max_\xi \Pi_{\xi,0}$ and by the 6% fee which is a reasonable approximation of the fee typically chosen by brokers, $\Pi_{6\%} = \Pi_{0.06,0}$.

3 Data, Identification, and Estimation

3.1 Data

The data set we use is the one constructed and used by Genesove and Mayer (2001). These data track individual properties in the condominium market in downtown

Boston and contain the date of the entry and exit of a property, listing price, and, if applicable, sale price, and property characteristics. Importantly, where available the data contain the sale price of previous transactions, which Genesove and Mayer (2001) also used to account for unobserved property heterogeneity in constructing a quality-adjusted price as discussed below. The data set includes property listings from January 6, 1990 to December 28, 1997 and property delistings (due to sale or withdrawal) from May 10, 1990 to March 16, 1998. We have 5792 observations in total.

As we are adding complexity to our model by solving for the optimal mechanism, we prefer to reduce complexity in other dimensions. In particular, we use a model that assumes a stationary environment. For stationarity to be a plausible assumption, we only use data from April 1, 1993 to April 1, 1996 because the changes in the real estate price index were relatively modest during this period. As an additional measure to reduce the effect of price changes over time, we do separate estimations for individual years, that is the year starting from April 1, 1993; from April 1, 1994; and from April 1, 1995. There is a trade-off when choosing the time length of the interval for which one does a separate estimation: a shorter interval reduces the effect of price changes over time and increases the weight of the effect of cross-sectional price variation. However, it also leads to fewer observations per estimation.[11] A calendar year also appears to be the appropriate choice for the time interval on the ground that the standard deviation of the price index within a year between 1993 and 1995 is much smaller than the standard deviation of the quality adjusted price. A further indicator is a measure stemming from a widely documented stylized fact in real estate markets: the correlation between price and time on market is weakly positive in cross-sectional data and negative in longitudinal data.[12] An intuitive explanation for the former

---

[11]Another effect is that some property is listed at the end of one period and sold at the beginning of another. As the results of the counterfactual analysis performed for different years are similar, we expect this effect to be small.

[12]Kang and Gardner (1989) provide empirical evidence that time on market increases with price in cross-sectional data – both based on their own dataset and on a review of other empirical work. Similar findings are reported in Glower, Haurin, and Hendershott (1998) and Genesove and Mayer (1997, 2001). The empirical literature typically finds a negative correlation between prices and vacancies – the latter can be seen as a proxy for time-on-market. Quigley (1999) investigates the effect of economic cycles on the housing market using international data on housing. He finds a

observation is that expensive houses need more time to sell. An explanation for the latter is that in times of booms, houses sell faster and at higher prices. Time on market increases in quality adjusted price for the years we include in our estimation. Further, the relative change of the real estate price index is generally small for the years considered. Table 1 shows the standard deviation of the price index within the time interval, the standard deviation of quality adjusted prices, the change of the price index, and the slope of time on market in price. Fig. 1 displays the movement of the real estate price index. The table and the figure suggest that for the time period 1993 to 1995, a time interval of one year appears to be sufficiently short to be considered cross-sectional. Excluding data for the years before 1993 and after 1995 has the additional advantage of avoiding truncation issues, which would occur for the first two and last two years in our dataset.

Measures of Stationarity			
Year	Index Std/Mean	Price Std/Mean	Relative ΔIndex	Slope Time on Market – Price Coeff. (Std. Err.)
1990	0.110	0.199	-0.241	-87.6 (28.2)
1991	0.021	0.204	-0.030	39.0 (20.9)
1992	0.051	0.214	0.131	67.0 (21.5)
1993	0.016	0.203	0.006	21.3 (22.0)
1994	0.020	0.185	-0.003	19.2 (22.0)
1995	0.026	0.190	0.067	40.1 (20.2)
1996	0.041	0.204	0.097	-7.6 (15.6)
1997	0.020	0.201	0.044	7.2 (9.8)

Table 1: Standard deviation of the real estate index divided by its mean in a given time interval, standard deviation of the quality adjusted price divided by its mean in a given time interval, change of the index divided by the index at the beginning of the interval, and slope time on market – quality adjusted price. The slope is the coefficient β_1 in the regression $T = \beta_0 + \beta_1 P + \epsilon$, where T is the time on market and P the quality adjusted price.

There are 2455 observations between April 1, 1993 and April 1, 1996. We exclude data with a quality-adjusted price larger than two and less than half as well as

---

negative correlation between vacancies and prices. See also the overview about empirical findings on the relation of vacancies and prices provided in Wheaton (1990).
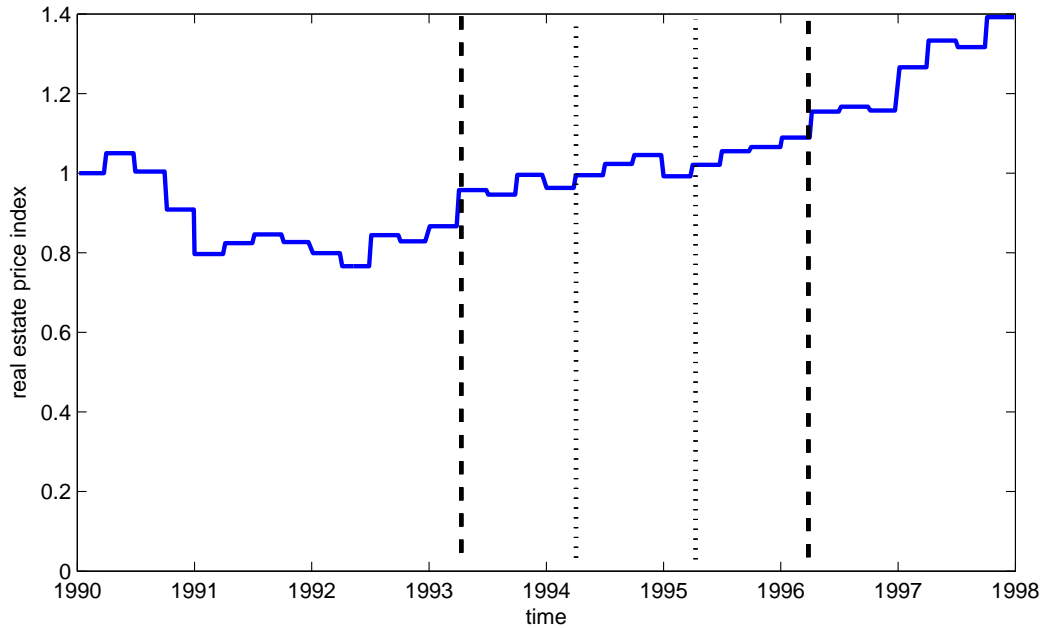
Figure 1: Development of the real estate price index in Boston. Data between the two dashed lines were used for estimations.

properties that were on the market for more than two years. This applies to 5.0% of the observations and results in 2333 remaining observations. The reason for the exclusion is that a property that is offered at less than half of or more than two times the previous transaction price (adjusted by the movement of the real estate price index) is likely to have undergone significant changes in quality or to constitute an error within the dataset. Similarly, a property on the market for more than two years was probably not seriously marketed. This exclusion does not change the estimation results qualitatively as shown in the robustness checks reported in the appendix.

Table 2 contains descriptive statistics of the included data. The average ratio of transaction price over list price is remarkably similar to the one found by Merlo and Ortalo-Magné (2004, Table 1), who use data from two regions in the United Kingdom (UK). Unsold houses stayed longer on the market than houses that did sell. If a house is delisted and relisted within four weeks, it is considered to be the same transaction. For further details about the data, see Genesove and Mayer (2001).

Descriptive Statistics			
Variables	All Houses	Sold Houses	Unsold Houses
Observations	2333	1522	811
Listing Price	$223,077 ($177,736)	$231973 ($172,861)	$206,383 ($185,501)
Quality Adjusted Listing Price	1.139 (0.219)	1.125 (0.220)	1.165 (0.215)
$100\frac{\text{Transaction Price}}{\text{Listing Price}}$		92%	
Time on Market	148 (134) days	130 (123) days	182 (147) days

Table 2: Sample Means (Standard Deviations) of Descriptive Statistics

3.2 Estimating the Dynamic Random Matching Model with Measurement Error

We use the quality index constructed by Genesove and Mayer (2001) which is based on previous transaction prices and a quarterly real estate price index P_q^{index}. Formally, in quarter q the measured "objective" quality index $\hat{\vartheta}_i$ of house i which was previously traded in quarter $q' < q$ for the price $\hat{P}_{iq'}$ is

$$\hat{\vartheta}_i = \frac{P_q^{\text{index}}}{P_{q'}^{\text{index}}}\hat{P}_{iq'}.$$

We assume that the true quality index ϑ_i is measured with a multiplicative error ϵ_i^Q, which may stem from an imperfect measurement of the quality index or changes in the quality of a property, that is, the measured quality is $\hat{\vartheta}_i = \vartheta_i\epsilon_i^Q$.[13] While the price index P_q^{index} is adjusted every quarter, our analysis is based on annual data. In what follows we therefore drop the time-index for individual observations.

Although it would be interesting to have a complex empirical model that takes into account many effects besides those of our model, it is also interesting and informative to analyze how well our baseline model can explain observations if we add two minimal modifications to this model.

First, the listing price is an imperfect proxy of the seller's reserve price as a sizable number of properties sell below the listing price. We account for this by assuming that the true quality-adjusted reserve price of seller i, denoted by p_i, involves

---

[13]This measure of quality adjustment is appropriate if all house prices move in proportion from one year to another. It neglects the impact of the seller specific type at time q' and of structural changes in demand from q' to time q.

a multiplicative error term ϵ_i^D, which may be thought of as a "discount" on the listing price. If P_i is the observed listing price of object i, the observed quality- and discount-adjusted reserve price p_i satisfies $P_i = p_i \epsilon_i^D \hat{\vartheta}_i$, or equivalently,

$$p_i = \frac{P_i}{\epsilon_i^D \hat{\vartheta}_i} = \frac{P_i}{\epsilon_i^P \vartheta_i}, \tag{12}$$

where the error of the proxy is $\epsilon_i^P := \epsilon_i^Q \epsilon_i^D$. We denote the density of ϵ_i^P as $h_p(\epsilon_i^P)$ and assume that it is log-normal.

Second, the true time on market t_i is also observed with an error, denoted ϵ_i^T, so that

$$T_i = t_i + \epsilon_i^T,$$

where T_i is the observed time on the market. The error ϵ_i^T may arise because a broker starts to show the property some time after it has been listed or because a property is delayed in being delisted after the buyer and the seller agreed on a deal. In the data set we use (described above), most properties are listed and delisted on a Sunday, Thus, we essentially have weekly data and delay happens at least until the end of the week.

Denote the joint density of prices and times on market as predicted by our baseline model as $h_{tps}(p, t, s)$ and the density of the error term for the time on market ϵ_i^T as $h_t(\epsilon_i^T)$. Let s_i be 1 if the house was sold and 0 otherwise.

First, we describe the prediction of the baseline model without error terms. The empirical inverse price function is denoted by $P_I(p) = 0.94\Phi(p)$. This function gives us the cost c of a seller who will optimally set the price $p = \Phi^{-1}(c/0.94)$. As sellers' costs have density $g(c)$, the *steady state* density of prices is proportional to $g(P_I(p))P_I'(p) =: g_p(p)$. Sellers that spend a long time on the market are over represented in steady state compared to the entrant population. Hence, to obtain the *entrant* distribution of prices, we have to divide by the average time on market, denoted $T(p)$. The *entrant* density of prices is therefore given as $\sigma g(P_I(p))P_I'(p)/T(p) =: g_{p0}(p)$, where σ is a constant that ensures that the density adds up to 1.[14]

Since rematching occurs every τ periods, and a house drops out of the market with probability $1 - \epsilon$ and is sold with probability $1 - \hat{F}(p)$, the probability that a

---

[14]In the matching model, σ is the mass or stocks of sellers in the market.

house is still on the market after t periods is $(\epsilon \hat{F}(p))^{t/\tau}$. Hence the joint distribution of p, t, and s has the density

$$h_{tps}(t,p,s) = \begin{cases} (1 - \hat{F}(p))(\epsilon \hat{F}(p))^t g_{p0}(p) & \text{if } s = 1, \\ (1 - \epsilon) \hat{F}(p)(\epsilon \hat{F}(p))^t g_{p0}(p) & \text{if } s = 0. \end{cases} \tag{13}$$

As the quality-adjusted reserve price p and time on market t are observed with noise ϵ^P and ϵ^T, the likelihood of an observation $X_i = (T_i, P_i, s_i)$ given the parameter vector θ as specified below is

$$l(X_i|\theta) = \sum_{k=1}^{T_i/\tau} \int_{-\infty}^{\infty} h_{tps}(T_i - k\tau, P_i/\epsilon^P, s_i) h_t(k\tau) h_p(\epsilon^P) d\epsilon^P \tag{14}$$

where $k\tau$ is the summation variable representing the error term ϵ^T.

3.3 Identification

Identification without measurement errors If the quality-adjusted price and time on market were observable without the errors ϵ^P and ϵ^T, it would be easy to see that our model is non-parametrically identifiable given observations of quality-adjusted price, time on market and whether a house was sold: Rearranging (4) and (5), the expressions for time on market as a function of the quality-adjusted price $T(p)$ and for the probability of ever selling $1 - F_\infty(p)$, we obtain

$$1 - \hat{F}(p) = \frac{1 - \hat{F}_\infty(p)}{T(p)/\tau},$$

$$\epsilon = \frac{T(p_2) - T(p_1)}{T(p_2)(1 - \hat{F}_\infty(p_1)) - T(p_1)(1 - \hat{F}_\infty(p_1))},$$

$$\tau = \frac{T(p_1)\hat{F}_\infty(p_2) - T(p_2)\hat{F}_\infty(p_1)}{\hat{F}_\infty(p_2) - \hat{F}_\infty(p_1)},$$

where p_1 and p_2 are two arbitrary prices (or – with some modification of the equations – price segment).[15] This makes $\hat{F}$, ϵ, and τ non-parametrically identifiable. Knowing $\hat{F}$, the distribution of sellers' costs G is non-parametrically identifiable by the distribution of prices G_p via the relationship $G(c) = G_p(\Phi^{-1}(c/0.94))$ because a seller with cost c sets the optimal reserve price $\Phi^{-1}(c/0.94)$. Given $\hat{F}$ and G, our theory provides the unique best-response fee-setting mechanism of the broker.

---

[15]The simplest example would be a price that never leads to trade and a price that leads to instantaneous trade, $p_2 = \overline{v}$ and $p_1 = \underline{v}$. This simplifies the expressions to $\tau = T(\underline{v})$ and $\epsilon = (T(\overline{v}) - T(\underline{v}))/T(\overline{v})$. In practice, one would want to take two different price segments and expectations over them, rather then two prices.

Identification with measurement errors With errors in the measurement of the quality adjusted discounted reserve price ϵ^P and of time on market ϵ^T, the argument is more involved, but identification is still possible. Redefine the observed probability of ever selling $1 - \hat{F}_\infty(P_i/\vartheta_i)$ and the time on market of sold and unsold houses $T^s(P_i/\vartheta_i)$ and $T^u(P_i/\vartheta_i)$ as functions of the observed quality adjusted listing price P_i/ϑ_i rather than the true (quality- and discount-adjusted) reserve price p_i. Further, let $\hat{T}^s(P_i/\vartheta_i) := T^s(P_i/\vartheta_i) + E[\epsilon^T]$ and $\hat{T}^u(P_i/\vartheta_i) := T^u(P_i/\vartheta_i) + E[\epsilon^T]$ be the observed average times on market. The probability of ever selling given P_i/ϑ_i and ϵ_i^P is $\text{Prob}(s = 1|P_i/\vartheta_i, \epsilon_i^P) = (1 - \hat{F}(P_i/(\vartheta_i\epsilon_i^P)))/(1 - \epsilon\hat{F}(P_i/(\vartheta_i\epsilon_i^P)))$ and the probability of never selling $\text{Prob}(s = 0|P_i/\vartheta_i, \epsilon_i^P) = (1 - \epsilon)\hat{F}(P_i/(\vartheta_i\epsilon_i^P))/(1 - \epsilon\hat{F}(P_i/(\vartheta_i\epsilon_i^P)))$. Given the unconditional density $h_p(\epsilon^P)$, the conditional densities are $h_p(\epsilon^P|P_i/\vartheta_i, s = 1) \propto h_p(\epsilon^P)/\text{Prob}(s = 1|P_i/\vartheta_i, \epsilon^P)$ and $h_p(\epsilon^P|P_i/\vartheta_i, s = 0) \propto h_p(\epsilon^P)/\text{Prob}(s = 0|P_i/\vartheta_i, \epsilon^P)$ by Bayes' Law. This gives us

$$1 - \hat{F}_\infty(P_i/\vartheta_i) = E_{\epsilon_i^P \sim H_p}\left[\frac{1 - \hat{F}(P_i/(\vartheta_i\epsilon_i^P))}{1 - \epsilon\hat{F}(P_i/(\vartheta_i\epsilon_i^P))}\right], \tag{15}$$

$$\hat{T}^s(P_i/\vartheta_i) = E_{\epsilon_i^P \sim H_p(\cdot|P_i/\vartheta_i, s=1)}\left[\frac{\tau}{1 - \epsilon\hat{F}(P_i/(\vartheta_i\epsilon_i^P))}\right] + E[\epsilon^T],$$

$$\hat{T}^u(P_i/\vartheta_i) = E_{\epsilon_i^P \sim H_p(\cdot|P_i/\vartheta_i, s=0)}\left[\frac{\tau}{1 - \epsilon\hat{F}(P_i/(\vartheta_i\epsilon_i^P))}\right] + E[\epsilon^T]$$

Note that (15) and $\hat{T}^u(P_i/\vartheta_i) - \hat{T}^s(P_i/\vartheta_i)$ do not require any knowledge about the distributions of ϵ^T and c and identify $\hat{F}$ and H_p for given ϵ and τ. The density of time on market t conditional on a particular price P_i/ϑ_i, $E_{\epsilon^P \sim H_p, \epsilon^T \sim H_t}[(\epsilon\hat{F}(P_i/(\vartheta_i\epsilon^P)))^{t/\tau - \epsilon^T}|\epsilon^T \leq t/\tau]$ identifies H_t. This uses only one price P_i/ϑ_i. The different densities of t for two additional prices P_j/ϑ_j and P_l/ϑ_l pin down ϵ and τ. Only the seller's distribution G remains to be identified. It is non-parametrically identifiable in the same way as without measurement errors ϵ^P and ϵ^T because $G(c) = G_p(\Phi^{-1}(c/0.94))$.

3.4 Estimation Procedure

While our model is non-parametrically identifiable in principle, our estimation is based on a parametric specification of the model and on Bayesian estimation methods. The main reason is that our main hypothesis is – loosely speaking – that if one

was to pick distributions F, or equivalently $\hat{F}$, and G randomly, then "most of the time" (or "on average") linear fees would perform close to optimally. Drawing $\hat{F}$ and G from the Bayesian posteriors distribution is a natural choice and also provides a clearer meaning to "on average". Further, the standard maximum likelihood estimator is biased because of the nuisance parameters and possible discontinuities in the likelihood function. Using a Bayesian estimator avoids the problems arising for the maximum likelihood estimator. Section 5 discusses methodology in more detail.

For our estimation we make the following functional form assumptions and take the following parametrization. We assume that $\hat{F}(v)$ and $G(c)$ are Beta distributions in the sense that $(v - \underline{v})/(\overline{v} - \underline{v})$ and $(c - \underline{c})/(\overline{c} - \underline{c})$ are Beta-distributed with respective parameters (α_F, β_F) and (α_G, β_G), where the density of the Beta distribution for x_i is proportional to $x_i^{\alpha_i - 1}(1 - x_i)^{\beta_i - 1}$ with $i = F, G$ and $x_F = (v - \underline{v})/(\overline{v} - \underline{v})$ and $x_G = (c - \underline{c})/(\overline{c} - \underline{c})$. The error in time on market ϵ^T follows a geometric distribution with parameter β_T, whose probability mass function is proportional to $e^{-\epsilon^T/\beta_T}$. Finally, the error in the quality-adjusted price ϵ^P is assumed to be normally distributed with mean 0 and variance σ_p^2. The advantage of using Beta distributions is that they are flexible in shape and specialize to linear virtual cost and valuation functions for $\beta_G = 1$ and $\alpha_F = 1$, respectively. The vector of parameters is thus $\theta = (\alpha_F, \beta_F, \alpha_G, \beta_G, \beta_T, \sigma_p, \epsilon, \underline{v}, \overline{v}, \underline{c}, \overline{c}, \lambda)$.[16] We set $\underline{c} = \underline{v}$ to simplify the computations and $\overline{c} = (1 - \xi_{\text{empirical}})\overline{v}$ and the empirical percentage fee $\xi_{\text{empirical}}$ to 0.06.

Given observations $X = \{X_i\}_{i=1}^N = \{(T_i, P_i, s_i)\}_{i=1}^N$ and the parameter vector θ, the likelihood function for the N observations is $l(X|\theta) := \prod_i l(X_i|\theta)$, where $l(X|\theta)$ is the probability of observing X given θ. The unconditional probability of observing X is denoted by $l(X)$. We are searching for the posterior beliefs about the parameters θ given X, $\pi(\theta|X)$.[17] By Bayes' Law $\pi(\theta|X) = l(X|\theta)\pi(\theta)/l(X)$, where $\pi(\theta)$ is the prior about θ. Assuming a uniform prior $\pi(\theta)$, we obtain the proportionality

---

[16]Note that θ does not include the period length τ, which is the only parameter that we cannot estimate. This is partly because we use a discrete time model, so the distribution of the error in time on market ϵ^T cannot be compared across different τ. Hence comparisons of the likelihood function would not make sense. A remedy would be to use continuous time errors ϵ^T. We chose to run robustness checks with alternative values of τ.

[17]We use lower case π to denote beliefs and, as above, upper case Π to denote expected profits.

$\pi(\theta|X) \propto l(X|\theta)$ as $l(X)$ does not depend on θ.[18]

We are seeking to find Bayesian estimates of the mean and variance of several functions $y(\theta)$, that is, $E_{\theta \sim \pi(\theta|X)}[y(\theta)]$ and $\text{Var}_{\theta \sim \pi(\theta|X)}[y(\theta)]$. These functions $y(\theta)$ are the ratio of linear over optimal profit $\Pi_{\text{linear}}(\theta)/\Pi_{\text{optimal}}(\theta)$, the ratio of proportional over optimal profit $\Pi_{\text{proportional}}(\theta)/\Pi_{\text{optimal}}(\theta)$, the ratio of 6% profit over optimal profit $\Pi_{6\%}(\theta)/\Pi_{\text{optimal}}(\theta)$, the optimal proportional fee $\xi_{\text{proportional}}(\theta)$, and the optimal linear fee with fixed component $\zeta_{\text{linear}}(\theta)$ and slope component $\xi_{\text{linear}}(\theta)$.

Computing the expectations requires computing a 10-dimensional integral. This is numerically challenging as a simple approach would lead to a computational time of several years.[19] We use several numerical improvements to reduce computational time to a few hours (see Section 5 and Appendix B).

While our main interest is the Bayesian posterior beliefs about the different variables, for illustration purposes (for example, to plot the predicted distribution of prices or the relation of price and time on market) we also want a pointwise estimator for θ. For this we take the Maximum A Posteriori Probability estimator $\theta_{\text{MAP}} = \arg \max_\theta \pi(\theta|X)$, that is, the mode of the Bayesian posterior distribution.[20]

3.5 Results

Estimates of parameters In Table 3 we report the Bayesian estimates for the parameter vector θ under the assumption that there are seven matchings per week (that is, $\tau = 1/365$). We report the Bayesian mean (for example $E_{\theta \sim \pi(\theta|X)}[\alpha_F]$ in the first row), the Bayesian standard deviation (for example $\sqrt{\text{Var}_{\theta \sim \pi(\theta|X)}[\alpha_F]}$ in parentheses in the first row), and the computational error (in square brackets). The

---

[18]To be precise, we constrain the uniform prior $\pi(\theta)$ to be the same positive constant wherever the constraints that we impose to avoid numerical problems are satisfied and to be 0 wherever they are not. These constraints are: (a) virtual valuation/cost functions must be increasing (to avoid the need for ironing), (b) $\alpha_F, \beta_F, \alpha_G, \beta_G \geq 1$ (to avoid infinite densities at endpoints of $\hat{f}$ and g), (c) $\overline{v} - \underline{v} \geq 0.2$ (to avoid numerical problems with nearly degenerate distributions that arise when $\overline{v} \approx \underline{v}$).

[19]This estimate is based on extrapolation of the time needed to solve part of the problem. In addition, standard Markov Chain Monte Carlo techniques for multidimensional integration do not work in our setup.

[20]Alternatively, one could use the Bayesian mean $\theta^* = E[\theta]$. However, some of the constraints imposed on θ (for example, increasing virtual valuations) by setting $\pi(\theta|X) = 0$ where constraints are violated, may not be satisfied at θ^*. They hold for θ_{MAP}. Note that for uniform priors, the Maximum A Posteriori Probability estimator coincides with the Maximum Likelihood Estimator.

computational error is as reported by the algorithm in Hahn (2005) (which is summarized in Appendix B). It provides the 99% confidence interval for the results of the (quasi-)Monte Carlo integration used when computing the Bayesian estimates.[21]

Estimated Parameter Values			
Parameters	1993	1994	1995
α_F	2.704(0.356) [0.002]	6.622(0.855) [0.003]	6.697(1.119) [0.006]
β_F	3.696(0.958) [0.003]	8.563(0.344) [0.004]	8.510(0.460) [0.007]
α_G	8.240(0.615) [0.006]	5.381(1.141) [0.003]	5.238(1.709) [0.003]
β_G	6.373(1.998) [0.006]	8.633(2.017) [0.004]	7.893(2.049) [0.005]
β_T	13.96(0.64) [0.01]	11.676(0.657) [0.006]	12.204(0.677) [0.01]
σ_p	0.2060(0.0028) [0.0001]	0.18586(0.00226) [9.e-05]	0.1894(0.0025) [0.0001]
ϵ	0.9683(0.0016) [0.0006]	0.9707(0.0015) [0.0005]	0.9750(0.0019) [0.0008]
$\underline{v}$	0.7605(0.0497) [0.0005]	0.9490(0.0193) [0.0005]	0.9105(0.0316) [0.0007]
$\overline{v}$	1.3492(0.0458) [0.0009]	1.3383(0.0293) [0.0006]	1.351(0.024) [0.001]
λ	0.5743(0.0764) [0.0004]	0.3098(0.1032) [0.0002]	0.5331(0.1218) [0.0003]
# Observations	720	831	782

Table 3: Bayesian estimates of parameter values for 1993 - 1995 for seven matchings per week. Table entries read: Mean (Standard Deviation) [Computational Error].

Figure 2 (a) provides a graphical illustration of parameter estimates for the Maximum A Posteriori Probability estimates of θ for 1993. In particular, it displays the estimated endogenous densities $f(v)$ (dashed line) and $g(c)$ (solid line). Panels (b) to

---

[21]For example, for the first column in the first row, if the correct value of the integral $\int \alpha_F \pi(\theta|X)d\theta$ was above 2.706 or below 2.702, then the null hypothesis that it is a correct calculation would be rejected with a probability of at least 99%. Note the difference to the Bayesian standard deviation of the estimate, which is determined by the (fixed) number of observations. The computational error depends on the number of computational steps. The algorithm increases the number of computational steps until the length of the confidence interval is below the desired level. While it is common not to report the computational error separately, it is useful to do so as it is an indicator of the reliability of the computational method.
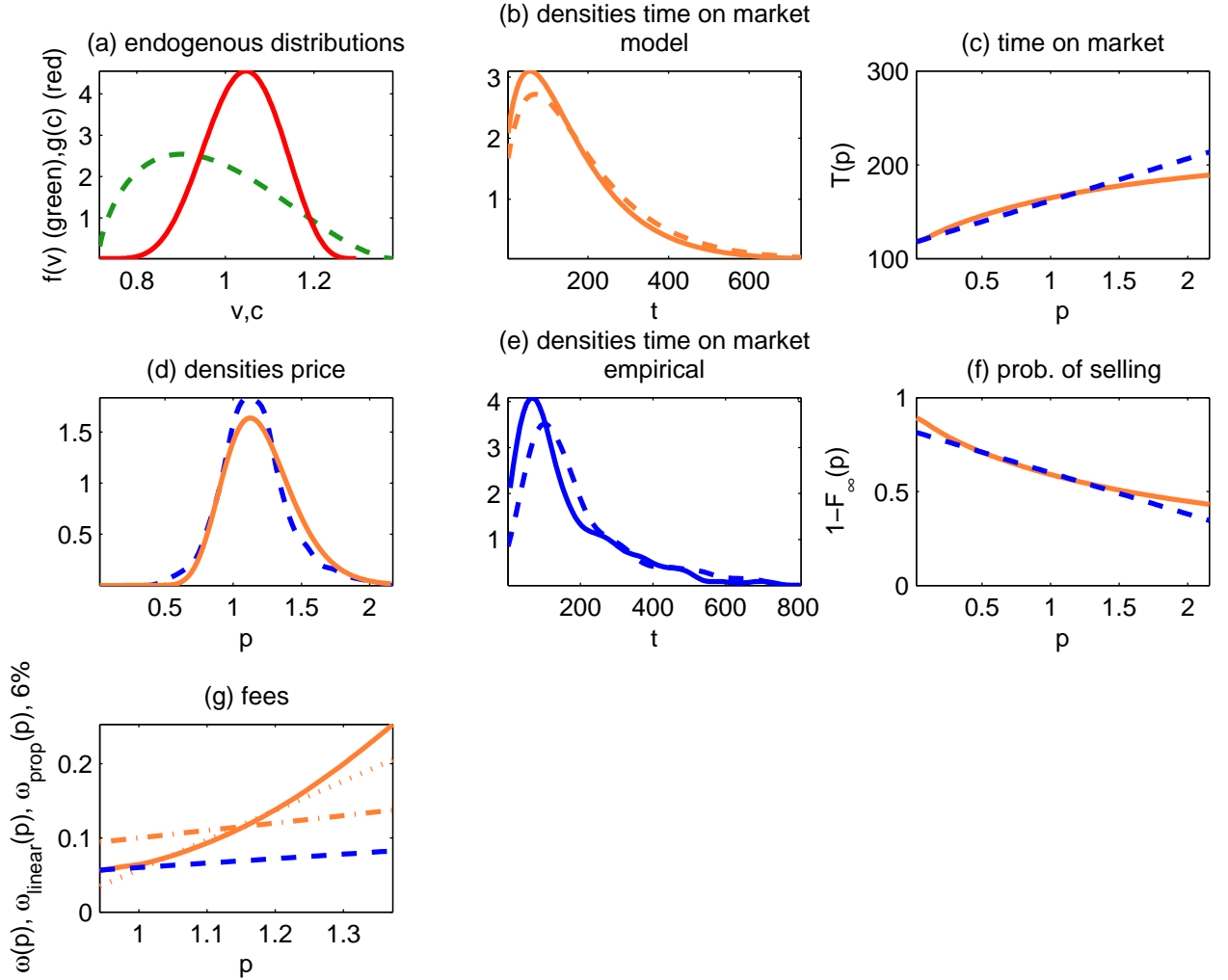
Figure 2: Theory and Empirics (1993, 7 matchings per week). The graphs are based on maximum a posteriori probability estimates of the parameters. (a) Endogenous densities $f(v)$ (dashed) and $g(c)$ (solid), (b) model prediction for densities time on market for sold (solid) and unsold (dashed) houses, (c) time on market as a function of quality-adjusted price, empirical (kernel regression, dashed) and predictions of model (solid), (d) density of quality-adjusted price, empirical (kernel density estimator, dashed) and model prediction (solid), (e) empirical densities time on market (kernel density estimator) for sold (solid) and unsold (dashed) houses, (f) probability of ever selling, empirical (kernel regression, dashed) and model predictions (solid), (g) empirical fees (6%, dashed), counterfactual prediction for best non-linear (solid), linear (dotted), and proportional (dash-dotted) fees.

(f) illustrate how well predictions of the model fit the observed data. Panels (c), (d) and (f) show, respectively, the relationships between time on the market, and quality-adjusted price, density of quality adjusted prices, and the probability of selling. In each case, the empirically observed relationship is displayed as a dashed line while the relationship that our model implies when evaluated at the estimated parameter values is shown as a solid line. The densities of time on the market for sold and unsold houses are displayed, respectively, with solid and dashed lines, in panel (b) as implied by our model and the parameter estimates and in panel (e) for what is observed in the data. In both cases the modes are positive, implying that the densities of time on the market are not exponential and suggesting that time on the market is measured with some error, as one would expect. Panels (b) to (f) in Figure 2 indicate that our model is a fairly good fit to the data. We will return to panel (g) later.

Counterfactual Fees and Profits			
Variables	1993	1994	1995
$\frac{\text{linear profit}}{\text{optimal profit}}$	0.9849(0.0026) [0.0009]	0.9886(0.0032) [0.0008]	0.991(0.003) [0.001]
$\frac{\text{proportional profit}}{\text{optimal profit}}$	0.8909(0.0288) [0.0008]	0.8548(0.0154) [0.0007]	0.872(0.014) [0.001]
$\frac{6\% \text{ profit}}{\text{optimal profit}}$	0.8563(0.0264) [0.0008]	0.8365(0.0371) [0.0007]	0.864(0.017) [0.001]
opt. proportional fee	0.07407(0.01015) [7.e-05]	0.05400(0.00535) [5.e-05]	0.06035(0.00547) [7.e-05]
opt. slope (linear fee)	0.3909(0.0347) [0.0003]	0.5040(0.0168) [0.0004]	0.4922(0.0212) [0.0006]
opt. fixed component (linear fee)	−0.3551(0.0404) [0.0003]	−0.5079(0.0208) [0.0004]	−0.4814(0.0276) [0.0006]
$\frac{6\% \text{ profit}}{\text{linear profit}}$	0.8694(0.0259) [0.0008]	0.8461(0.0360) [0.0007]	0.872(0.016) [0.001]
$\frac{6\% \text{ profit}}{\text{proportional profit}}$	0.9617(0.0305) [0.0009]	0.9784(0.0329) [0.0008]	0.991(0.015) [0.001]
# Observations	720	831	782

Table 4: Bayesian Estimates of Fees and Profits Implied by the Model and the Parameter Estimates (1993 - 1995, seven matchings per week). Table entries read: Mean (Standard Deviation) [Computational Error].

Counterfactual: The opportunity cost of simple fees We use Bayesian estimates of the parameters reported in Table 3 to conduct a counterfactual analysis. Table 4 contains the Bayesian estimates of fees and profits for all three years for which we have data under the assumption that there are seven matchings per week (i.e. $\tau = 1/365$). The results for the alternative assumptions of two and 14 weekly matchings are summarized in Appendix A as are the results for the estimation with a quality index constructed from physical characteristics and for a dataset that does not exclude outliers. According to Table 4 (third row), intermediaries who used a 6% fee achieved between 83% and 86% of the optimal profit, given the parameter estimates. The optimal proportional fee, which is displayed in row four, varies between 5.4% and 7.4% over the three years. Rather remarkably, an intermediary's expected profit under an optimally chosen linear fee falls short of the maximum profit by no more than roughly 1.2% (see the first row in Table 4). Panel (g) in Figure 2 shows the optimal fee $\omega(p)$ as a solid curve, the optimal linear fee $\omega_{linear}(p)$ as a dotted line and the optimal proportional fee $\omega_{prop}(p)$ as a dashed line as implied by our model evaluated at the estimated parameter values. The empirically observed 6% fee is displayed as a dashed line.

4   Discussion

Given that simple percentage fees are easy to implement, they perform fairly well, generating at least 83% of the revenues of the optimal mechanism. However, it is somewhat puzzling that linear fees, which appear similarly simple as percentage fees, achieve 98% of the optimal profit, but are not used by intermediaries. One possible explanation is that the common theoretical assumption that agents' types are private information, but the distributions of types (which correspond to demand and supply functions) are common knowledge, is only an approximation of reality. In our counterfactual analysis, we had the benefit of hindsight: we observe all offered properties and their histories on the market (price offered, sold/unsold, time on market) *ex post*. A real estate agent will typically not know the details of demand and supply in advance and should prefer mechanisms that are robust to "getting things wrong".

We will conduct two simple counterfactual thought experiments that shed some light on robustness. In particular, we will examine robustness with respect to getting the parameters wrong and with respect to getting the quality index wrong.

Parameter uncertainty   For robustness with respect to parameter uncertainty, we will make the following comparison.[22] We use the estimate of the parameters from one year (for example, 1993) to compute the optimal linear fee and compute the profits of a broker that uses this fee structure in another year (for example, 1994). In Table 5, we compare the combinations of different years with profits that can be achieved using a 6% fee. For the example of using the best linear fee from 1993 if the true parameters are those of 1994, one obtains 76.5% of optimal profits, which is less than the 85.2% that a broker achieves with 6% fees. The off-diagonal values in Table 5 show that even when taking three consecutive years into account, using parameter estimates from one year in order to choose linear fees in another year can lead to significant losses in revenue, particularly when evaluated at the minimum.[23] It seems reasonable to assume that over a longer period, linear fees are even less robust to changes in demand and supply. In contrast to linear fees, percentage fees are rather robust. Note that for the sake of computational simplicity, we have calculated profits implied by Bayesian parameter estimates rather than Bayesian estimates of profits.[24]

Uncertainty about quality index   Next, we will consider robustness with respect to the quality index ϑ_i. We have assumed that the quality index for a property is observable by market participant (even if not directly observable by the econometrician) and that it is possible to make contracts contingent on the quality index. In reality, this may only be partially true, either because market participants only have a noisy signal of the quality index or because it is too complicated to write contracts that are contingent on this index. We will report results for a counterfactual analysis

---

[22]We thank Juan Santaló for suggesting this comparison.

[23]For arguments why the minimum may be the appropriate criterion when robustness is a concern, see Bergemann and Morris (2005).

[24]Profits implied by the Bayesian mean of the parameters are slightly different from the Bayesian mean of profits because profits are not a linear function of parameters. Formally, $E[y(\theta)]$ is not (exactly) equal to $y(E[\theta])$. However, it is reasonably close and much simpler to compute.

Robustness with respect to Parameter Uncertainty					
Optimal fee based on estimates from year	ξ	ζ	True parameters from year		
			1993	1994	1995
1993	0.4	-0.36	0.984	0.765	0.913
1994	0.51	-0.51	0.701	0.989	0.941
1995	0.49	-0.48	0.886	0.938	0.991
6%	0.06	0	0.849	0.852	0.873

Table 5: Profits generated by choosing the optimal linear fee (proportional part ξ and fixed part ζ) implied by the estimated parameters of one year in another year. The last row represents profits generated by 6% fees. All profits are provided as a ratio of the optimal profit of the given year. Values are based on Bayesian means of parameter estimates.

in which we assume that the broker chooses the best linear fee assuming that the quality index is some value $\bar{\vartheta}$ (this could be the average quality index) while in truth it is some other value ϑ_i.

So far we have considered the quality-adjusted fee as a function of the quality-adjusted price, which is $\omega(p) = \xi p + \zeta$ for linear fees. The non-quality-adjusted fee in case the quality is ϑ_i is $\omega_i(P_i) = \xi P_i + \vartheta_i \zeta$ with non-quality-adjusted price $P_i = \vartheta_i p$. If the broker adjusts the fee by $\bar{\vartheta}$ rather than the true quality index ϑ_i, the fee is $\overline{\omega}(P_i) = \xi P_i + \bar{\vartheta}\zeta$. Then profits given a price P_i and a quality index ϑ_i are, following (10),

$$
\overline{\Pi}(P_i) = (\xi P_i + \bar{\vartheta}\zeta)\left(-\ln \hat{F}\left(\frac{P_i}{\vartheta_i}\right)\right)\hat{F}\left(\frac{P_i}{\vartheta_i}\right) + \int_{P_i}^{\vartheta_i \bar{v}}(\xi y + \bar{\vartheta}\zeta)\left(-\ln \hat{F}\left(\frac{y}{\vartheta_i}\right)\right)\hat{f}\left(\frac{y}{\vartheta_i}\right)dy
$$
$$
= \vartheta_i\left[\left(\xi p + \zeta \frac{\bar{\vartheta}}{\vartheta_i}\right)(-\ln \hat{F}(p))\hat{F}(p) + \int_p^{\bar{v}}\left(\xi y + \zeta \frac{\bar{\vartheta}}{\vartheta_i}\right)(-\ln \hat{F}(y))\hat{f}(y)dy\right].
$$
(16)

The expression in square brackets in (16) is the quality-adjusted price with a quality-adjusted fee $\omega(p) = \xi p + \zeta \bar{\vartheta}/\vartheta_i$. Taking expectations over the distribution of price P_i, the expression in (16) becomes $\vartheta_i \Pi_{\xi, \zeta\bar{\vartheta}/\vartheta_i}$ with $\Pi_{\xi,\zeta}$ given in (11). We compute profits for the best linear fee (for $\bar{\vartheta}$), $\vartheta_i \Pi_{\xi_{\text{linear}}, \zeta_{\text{linear}}\bar{\vartheta}/\vartheta_i}$ and 6% fees, $\vartheta_i \Pi_{0.06,0}$ for different values of ϑ_i. We report these profits – divided by ϑ_i for easier comparability – in Figure 3. The flat lines represent profits for a 6% fee for the years 1993, 1994, and 1995. The curves represent profits for linear fees. The figure shows that if one has a

very good estimate of ϑ_i, linear fees perform better than a six percent fee. However, if $\overline{\vartheta}$ is sufficiently far from ϑ_i, linear fees perform far worse.
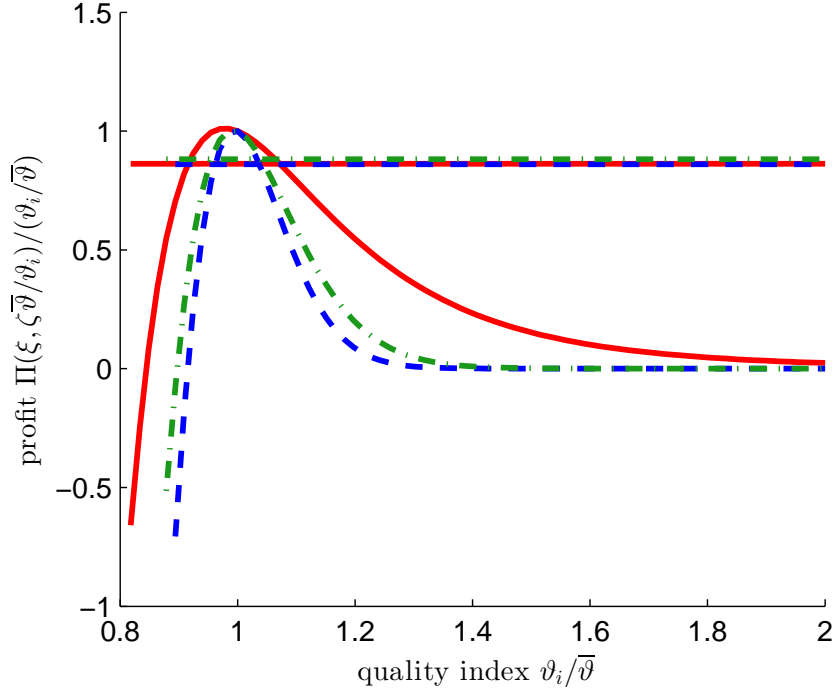

Figure 3: Robustness to choosing fee based on quality index $\overline{\vartheta}$ rather than true quality index ϑ_i. Flat lines are profits for a 6% fee, curves are profits for the best linear fee for a given year. Calculations are provided for the years 1993 (solid line), 1994 (dashed), and 1995 (dash-dotted).

Profit function under percentage fees   A further interesting question concerning percentage fees is how sensitive the profit is with respect to the choice of the percentage. At the optimal percentage fee, the first-order effect of a change of the fee is of course zero. However, the higher-order effects are relevant for larger changes of the fee. Figure 4 shows profits as a function of the percentage fee for the years 1993, 1994, and 1995, indicating a low sensitivity of profits to moderate deviations from the optimal percentage fee.
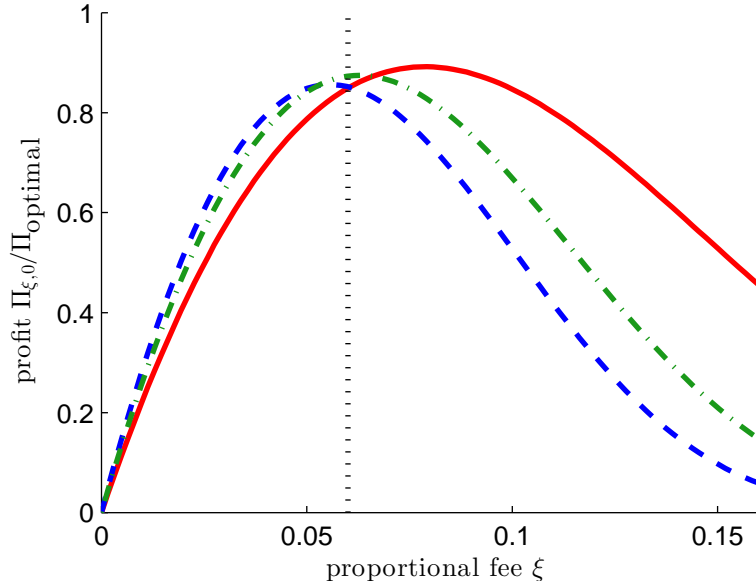
Figure 4: Profit as a function of the percentage fee being charged. Calculations are provided for the years 1993 (solid line), 1994 (dashed), and 1995 (dash-dotted). The vertical line indicates 6% fees.

5 Methodology

The combination of numerical techniques we employ reduces computational time from (estimated) several years to a few hours. The adaptive quasi-Monte Carlo algorithm we use also ensures robustness of the computation and avoids possible pitfalls, such as a Markov Chain Monte Carlo algorithm getting stuck in a local peak. Appendix B provides a more detailed description of the numerical techniques we use.

Our findings also suggest a number of simplifications for future applied research. Our theoretical work has shown that large parts of the problem at hand can be solved in closed-form if G is a mirrored Generalized Pareto distribution and if F is a Genelarized Pareto distribution. To be precise, for $G(c) = [(c - \underline{c})/(\overline{c} - \underline{c})]^\alpha$ and $F(v) = 1 - [(\overline{v} - v)/(\overline{v} - \underline{v})]^\beta$, the virtual cost function Γ and the virtual valuation function Φ are linear. This implies that the optimal fee will be linear and that a seller who faces a linear fee (not necessarily the optimal fee) will set a reserve price that is a linear function of his cost. For example, facing the empirically observed 6% fee a seller with cost c will set the reserve price $\Phi^{-1}(c/0.94)$, which is linear in c because of the linearity of Φ. Given linear virtual cost and virtual valuation functions and

linear fees, the distribution of reserve prices G_p will be a mirrored Generalized Pareto distribution, that is, $G_p(p) = [(p - \underline{p})/(\overline{p} - \underline{p})]^{\alpha}$, where α is the shape parameter of the Pareto distribution G and the boundaries of the support $\underline{p}$ and $\overline{p}$ are given by the parameters of F and G.[25]

Estimating the shape parameter α of a Pareto distribution is a well studied problem, with several well documented frequentist and Bayesian estimation methods (see Coles, 2001) if one is willing to abstract away from the measurement error in the quality-adjusted reserve price. An estimate of α and $\underline{c}$ allows for a closed-form solution for the optimal fee, $\omega(p) = (p - \underline{c})/(\alpha + 1)$. For the current analysis, the simplifying assumption of (mirrored) Generalized Pareto distributions was not admissible because this assumption implies that linear fees are exactly optimal (see Loertscher and Niedermayer, 2012). Therefore, the question how linear fees perform compared to optimal (possibly nonlinear) fees, could not have been answered. However, an analysis that is concerned with the level rather than the linearity of fees would be greatly simplified by this assumption. Given that linear fees achieve more than 98% of the optimal fees, this simplification does not appear to be a severe restriction. If one is further willing to assume that percentage fees, rather than linear fees in general, are optimal, this adds the additional restriction $\underline{c} = 0$. Hence, under this assumption the shape parameter α of the price distribution G_p – which coincides with the elasticity of supply – is sufficient to determine the optimal fee $\omega(p) = p/(\alpha + 1)$.

Even if one does not neglect the measurement error of the quality-adjusted reserve price, Pareto distributions greatly simplify the structural estimation, because of the closed-form solutions for Φ^{-1}, $\hat{F}$, T, and G_{p0}. Further, the functional forms have the effect that the parameters enter the posterior density, loosely speaking, in a more linear way. This should, in turn, also allow for more easily implementable Markov Chain Monte Carlo techniques that are sufficiently fast and robust for the problems at hand.

---

[25]For example for 6% fees, the boundaries $\underline{p} = \Phi^{-1}(\underline{c}/0.94)$ and $\overline{p} = \Phi^{-1}(\overline{c}/0.94)$ are obtained by using the functional form of F, which yields

$$\underline{p} = \frac{\underline{c}\beta(\overline{v} - \underline{v}) + 0.94\overline{v}}{0.94(\beta(\overline{v} - \underline{v}) + 1)} \quad \text{and} \quad \overline{p} = \frac{\overline{c}\beta(\overline{v} - \underline{v}) + 0.94\overline{v}}{0.94(\beta(\overline{v} - \underline{v}) + 1)}.$$

Similarly to Bajari (1997) and Bajari and Hortaçsu (2003) we use a Bayesian estimation method to avoid the difficulties arising when using a maximum likelihood or a general method of moments estimator. The difficulties arise because two regularity conditions are not satisfied in our setup. First, the number of nuisance parameters (here the errors-in-variables for time-on-market and quality adjusted price) increases with the number of observations. If we were to drop the errors-in-variables (there are good reasons not to do so[26]), then another condition would be violated: the boundaries of observable data would depend on the estimated parameters $\underline{v}$, $\overline{v}$, $\underline{c}$, and $\overline{c}$. Second, for Beta distributions there is no simple condition that ensures that the virtual valuation and cost function are increasing (that is, Myerson's regularity condition is satisfied). Not excluding parameters that lead to non-increasing virtual valuations would cause problems both for numerical calculations and for estimation.[27] Therefore, we exclude non-increasing virtual valuations by setting the prior to zero for corresponding parameters. Using an analogous approach for a constrained Maximum Likelihood estimator would lead to discontinuities in the likelihood function, a violation of another regularity condition for maximum likelihood estimation.

There is a strand of literature that views Bayesian estimators as classical (efficient) estimators; see for example Hirano and Porter (2003) and the references therein. Hirano and Porter (2003) note that "in regular parametric models, Bayesian estimators are typically asymptotically equivalent to Maximum Likelihood" estimators and hence efficient and show that for certain non-regular models the Bayesian estimator is efficient, whereas the Maximum Likelihood estimator is not; see also Chernozhukov and Hong (2003).

Donald and Paarsch (1993) and Laffont, Ossard, and Vuong (1995) show that for

---

[26]In our theoretical paper we show that in our setup errors-in-variables are needed to explain that the mode of the distribution of time-on-market is greater than zero and that time-on-market is larger for an unsold house than for a sold house offered at the same quality adjusted price. We observe both in our data.

[27]Distributions that violate Myerson's regularity condition are typically avoided in both theoretical and empirical work. For computation, a violation of the condition would greatly increase the complexity of the computational procedure, since one would have to use ironing techniques (see Myerson (1981)). For estimation, this would mean that choices of agents would be discontinuous function of their types. In particular, this would mean that some reserve prices (and also transaction prices) would never be chosen in equilibrium, which is a problem for identification.

standard first- and second-price auctions maximum likelihood and general method of moment estimators are biased. They derive methods to correct the bias. Guerre, Perrigne, and Vuong (2000) derive a non-parametric estimator for first-price auctions. Unfortunately, one cannot apply the techniques developed in these papers in our setup for a variety of reasons, which include our errors-in-variables, the facts that we do not take a stance on whether bargaining is a first- or second-price auction, and that we use time-on-market for estimation rather than the distribution of bids.

6 Conclusions

In this paper we have used a structural model based on Loertscher and Niedermayer (2012) to estimate demand and supply parameters for a data set constructed by Genesove and Mayer (2001), which covers the Boston condominium market in the 1990s. Using these parameter estimates for a counterfactual analysis, we have found that for the parametrization with seven matchings per week the empirically observed 6% fee achieves 83% or more of the maximum profit that can be achieved with an optimal Bayesian mechanism. With an optimally structured linear fee, brokers achieve at least 98% of the maximum profit. Of course, the optimal mechanism and the optimal linear fee both vary with the parameter estimates, which exhibit some variation in the data set. Even well-informed brokers will face non-trivial uncertainty about the relevant parameters values and hence about the optimal mechanism, be it linear or unconstrained. In contrast, 6% fees are obviously independent of these specific parameter values and the finer details of the design problem such as the objective quality of a property and supply and demand parameters. Thus, they resemble robust mechanisms in the sense of Wilson (1987), which makes their good empirical performance even more remarkable.

Appendix

A Robustness Checks

In this appendix we report robustness checks. Tables 6 and 7 summarize the result-
ing fees and profits when alternative rematching frequencies are chosen (two and 14
rematchings per week rather than seven).

Table 8 performs an additional robustness check based on an alternative quality
index constructed by Genesove and Mayer (2001), which relies on physical charac-
teristics rather than the previous transaction price. Given the vector Y_i of physical
characteristics of property i (such as the number of bedrooms and whether there is air
conditioning), Genesove and Mayer (2001) ran a regression with the observed listing
price $\hat{P}_i$ as the dependent variable:

$$\ln \hat{P}_i = Y_i\beta + \epsilon_i.$$

For our purposes, we interpret $\exp(Y_i\beta) = \vartheta_i$ as the quality index and the statistical
error $\exp(\epsilon_i) = p_i$ as the quality adjusted price.[28]

We also report estimation results if we do not exclude outliers based on time on
market, but only exclude property with a price less than 1/10 of or more than 10
times the original sales price adjusted by changes of the real estate price index, in
Table 9.

Tables 6, 7, 8, and 9 suggest that our results are robust to a change of the matching
frequency, quality index, and outliers to be excluded.

Data before 1993 and after 1995 cannot be used for a reliable estimation for
several reasons. First, our theory is about cross-sectional variations and relies on the
simplifying assumption that the environment is (nearly) stationary. Since the years
before 1993 and after 1995 show considerable change of real estate prices over time,
a model based on stationarity cannot be used for estimation. Second, data at the
beginning and the end of the periods we observe suffers from truncation issues. The
reason for this is that property that was listed before 1990 or delisted after 1997

---

[28]Note that the quality adjusted prices based on physical characteristics are a biased estimate,
since the residuals $\hat{\epsilon}_i$ are correlated.

does not show up in the dataset. While extending our theory to a non-stationary environment is possible in principle, it would add a significant amount of complexity. For the sake of completeness, we also report estimates for the years before 1993 and after 1995 in Table 10. While interpreting estimates of a stationary model based on non-stationary data is difficult, Table 10 should at least give a rough idea of what a more complex model that takes non-stationarity and truncation into account might predict. Linear fees perform similarly well, while there is variation in the performance of proportional fees. The optimal proportional fee ranges between 4% and 15% and a 6% fee achieves about 70% of optimal profits.

Counterfactual Fees and Profits			
Variables	1993	1994	1995
$\frac{\text{linear profit}}{\text{optimal profit}}$	0.9827(0.0022) [0.0007]	0.9826(0.0014) [0.0008]	0.979(0.002) [0.001]
$\frac{\text{proportional profit}}{\text{optimal profit}}$	0.9039(0.0184) [0.0007]	0.8866(0.0083) [0.0007]	0.825(0.013) [0.001]
$\frac{\text{6\% profit}}{\text{optimal profit}}$	0.8899(0.0154) [0.0007]	0.8843(0.0113) [0.0007]	0.788(0.021) [0.001]
opt. proportional fee	0.06812(0.00597) [5.e-05]	0.05839(0.00281) [5.e-05]	0.04775(0.00207) [6.e-05]
opt. slope (linear fee)	0.3595(0.0249) [0.0003]	0.3998(0.0123) [0.0003]	0.4832(0.0117) [0.0006]
opt. fixed component (linear fee)	−0.3288(0.0292) [0.0002]	−0.3865(0.0166) [0.0003]	−0.4864(0.0134) [0.0006]
$\frac{\text{6\% profit}}{\text{linear profit}}$	0.9055(0.0143) [0.0007]	0.9000(0.0108) [0.0007]	0.805(0.021) [0.001]
$\frac{\text{6\% profit}}{\text{proportional profit}}$	0.9846(0.0130) [0.0007]	0.9975(0.0069) [0.0008]	0.955(0.018) [0.001]
# Observations	720	831	782

Table 6: Bayesian Estimates of Fees and Profits Implied by the Model and the Parameter Estimates (1993 - 1995, two matchings per week).

Counterfactual Fees and Profits			
Variables	1993	1994	1995
$\frac{\text{linear profit}}{\text{optimal profit}}$	0.9872(0.0025) [0.0009]	0.9766(0.0026) [0.0009]	0.97(0.001) [0.01]
$\frac{\text{proportional profit}}{\text{optimal profit}}$	0.9024(0.0125) [0.0008]	0.7799(0.0187) [0.0007]	0.759(0.006) [0.009]
$\frac{\text{6\% profit}}{\text{optimal profit}}$	0.7130(0.0273) [0.0006]	0.7735(0.0229) [0.0009]	0.752(0.011) [0.009]
opt. proportional fee	0.11198(0.00545) [0.0001]	0.06178(0.00456) [6.e-05]	0.0542(0.0030) [0.0004]
opt. slope (linear fee)	0.4512(0.0198) [0.0004]	0.5180(0.0121) [0.0005]	0.547(0.009) [0.007]
opt. fixed component (linear fee)	−0.3908(0.0192) [0.0003]	−0.5010(0.0132) [0.0005]	−0.547(0.012) [0.007]
$\frac{\text{6\% profit}}{\text{linear profit}}$	0.7222(0.0274) [0.0006]	0.7920(0.0220) [0.0009]	0.777(0.011) [0.01]
$\frac{\text{6\% profit}}{\text{proportional profit}}$	0.7901(0.0273) [0.0007]	0.9918(0.0168) [0.0009]	0.99(0.01) [0.01]
# Observations	720	831	782

Table 7: Bayesian Estimates of Fees and Profits Implied by the Model and the Parameter Estimates (1993 - 1995, 14 matchings per week).

Counterfactual Fees and Profits			
Variables	1993	1994	1995
$\frac{\text{linear profit}}{\text{optimal profit}}$	0.988(0.002) [0.001]	0.9771(0.0032) [0.0006]	0.994(0.002) [0.001]
$\frac{\text{proportional profit}}{\text{optimal profit}}$	0.905(0.016) [0.001]	0.7845(0.0158) [0.0005]	0.922(0.019) [0.001]
$\frac{\text{6\% profit}}{\text{optimal profit}}$	0.830(0.019) [0.001]	0.7044(0.0323) [0.0004]	0.918(0.023) [0.001]
opt. proportional fee	0.08625(0.00603) [0.0001]	0.04377(0.00253) [3.e-05]	0.05678(0.00255) [7.e-05]
opt. slope (linear fee)	0.4100(0.0212) [0.0005]	0.5428(0.0166) [0.0003]	0.4598(0.0235) [0.0006]
opt. fixed component (linear fee)	−0.3806(0.0271) [0.0005]	−0.5753(0.0191) [0.0003]	−0.4703(0.0263) [0.0006]
$\frac{\text{6\% profit}}{\text{linear profit}}$	0.840(0.020) [0.001]	0.7208(0.0315) [0.0004]	0.923(0.022) [0.001]
$\frac{\text{6\% profit}}{\text{proportional profit}}$	0.917(0.027) [0.001]	0.8977(0.0316) [0.0005]	0.995(0.007) [0.001]
# Observations	740	854	799

Table 8: Bayesian Estimates of Fees and Profits Implied by the Model and the Parameter Estimates (1993 - 1995, seven matchings per week) when physical characteristics are used to control for house quality.

Counterfactual Fees and Profits			
Variables	1993	1994	1995
$\frac{\text{linear profit}}{\text{optimal profit}}$	0.981(0.003) [0.009]	0.982(0.018) [0.002]	0.987(0.001) [0.008]
$\frac{\text{proportional profit}}{\text{optimal profit}}$	0.888(0.013) [0.008]	0.848(0.027) [0.001]	0.890(0.007) [0.007]
$\frac{6\% \text{ profit}}{\text{optimal profit}}$	0.837(0.016) [0.008]	0.799(0.046) [0.001]	0.886(0.012) [0.007]
opt. proportional fee	0.0811(0.0055) [0.0007]	0.04852(0.00299) [8.e-05]	0.0569(0.0032) [0.0004]
opt. slope (linear fee)	0.395(0.015) [0.004]	0.4441(0.0580) [0.0007]	0.428(0.012) [0.003]
opt. fixed component (linear fee)	$-0.387(0.021)$ [0.004]	$-0.4730(0.0683)$ [0.0008]	$-0.436(0.018)$ [0.003]
$\frac{6\% \text{ profit}}{\text{linear profit}}$	0.853(0.016) [0.008]	0.814(0.044) [0.001]	0.897(0.011) [0.008]
$\frac{6\% \text{ profit}}{\text{proportional profit}}$	0.942(0.020) [0.009]	0.942(0.032) [0.002]	0.995(0.009) [0.008]
# Observations	761	869	807

Table 9: Bayesian Estimates of Fees and Profits Implied by the Model and the Parameter Estimates (1993 - 1995, seven matchings per week, no exclusion of outliers based on time on market, only excluding outliers with a price less than 1/10 of or more than 10 times the adjusted original sales price). Table entries read: Mean (Standard Deviation) [Computational Error].

	Counterfactual Fees and Profits				
Variables	1990	1991	1992	1996	1997
$\frac{\text{linear profit}}{\text{optimal profit}}$	0.940 (0.049) [0.01]	0.98 (0.004) [0.01]	0.99 (0.001) [0.01]	0.989 (0.003) [0.003]	0.975 (0.002) [0.006]
$\frac{\text{proportional profit}}{\text{optimal profit}}$	0.717 (0.052) [0.008]	0.886 (0.022) [0.009]	0.95 (0.009) [0.01]	0.819 (0.019) [0.003]	0.641 (0.050) [0.004]
$\frac{6\% \text{ profit}}{\text{optimal profit}}$	0.701 (0.057) [0.007]	0.729 (0.039) [0.007]	0.642 (0.029) [0.007]	0.728 (0.059) [0.002]	0.641 (0.050) [0.004]
opt. proportional fee	0.0548 (0.0027) [0.0006]	0.1076 (0.0127) [0.001]	0.144 (0.008) [0.002]	0.0449 (0.0035) [0.0001]	0.0598 (0.0002) [0.0004]
opt. slope (linear fee)	0.468 (0.094) [0.005]	0.414 (0.023) [0.004]	0.356 (0.018) [0.004]	0.602 (0.024) [0.002]	0.734 (0.029) [0.005]
opt. fixed component (linear fee)	−0.540 (0.120) [0.005]	−0.406 (0.039) [0.004]	−0.269 (0.019) [0.003]	−0.614 (0.028) [0.002]	−0.700 (0.024) [0.005]
$\frac{6\% \text{ profit}}{\text{linear profit}}$	0.746 (0.058) [0.008]	0.743 (0.041) [0.007]	0.648 (0.029) [0.007]	0.736 (0.058) [0.002]	0.657 (0.051) [0.004]
$\frac{6\% \text{ profit}}{\text{proportional profit}}$	0.98 (0.02) [0.01]	0.824 (0.057) [0.008]	0.678 (0.029) [0.008]	0.888 (0.057) [0.003]	1.00 (0.0004) [0.007]
# Observations	519	656	666	688	531

Table 10: Bayesian Estimates of Fees and Profits Implied by the Model and the Parameter Estimates (1990 - 1997 without 1993 - 1995, seven matchings per week). Note that the estimates are not reliable because of non-stationarity and truncation of the data.

B Numerical Methods

In this appendix, we briefly describe the main numerical methods we use. The Matlab and Fortran-MEX files used are available upon request from the authors.

The computation of the likelihood function can be sped up by rewriting the expressions. By using the true values t and p as summation/integration variables rather than the error terms ϵ^T and ϵ^P, (14) can be transformed to

$$\sum_{k=1}^{T_i/\tau} \int_{\underline{p}}^{\overline{p}} h_{tps}(k\tau, p, s_i) h_t(T_i - k\tau) h_p(P_i - p) dp, \tag{17}$$

where the summation variable $k\tau$ represents the true time on market t.

Note the difference in terms of the numerical evaluation of (14) and (17) with (13). For the former, one must evaluate a three-dimensional function $h_{tps}(\cdot, \cdot, \cdot)$ in four loops (summation and integration, looping over all observations and repeatedly evaluating the likelihood function in a Monte Carlo simulation). For the latter one has two one-dimensional functions $F(p)$ and $g_{p0}(p)$ combined with simple addition and multiplication operations. One can greatly increase the computational speed by approximating $F(p)$ and $g_{p0}(p)$ with Chebyshev polynomials and then using the closed-form solutions of the integrals of the polynomials (i.e. reusing precomputed values for multiple Gauss-Chebyshev quadratures).

The Bayesian estimation further requires us to integrate over the 10-dimensional posterior density function $\pi(\theta|X)$. We do this by using the Divonne algorithm, which is an extension of the CERNLIB D.151 algorithm (see Hahn, 2005; Friedman and Wright, 1981, for a detailed description). (Note that the usual numerical techniques, such as Markov Chain Monte Carlo sampling/integration, fail to provide a useful accuracy level. Further, tests indicate that the integrands are not smooth enough for sparse-grid quadrature.)[29] The basic intuition for how the algorithm works can be illustrated by example of the simple one-dimensional case. The algorithm relies on the Koksma-Hlawka theorem, which states that the error of

---

[29]One reason may be that some of the variables are strongly correlated (in terms of posterior distribution). Another is that the posterior density has multiple peaks (for example, not allowing distributions with non-increasing virtual valuations and approximation errors when computing the integrands). A Markov Chain Monte Carlo method (without additional refinements) has difficulties "walking" over narrow ridges and might additionally get stuck in local maxima.

a finite sum approximation of an integral of a function $y(x)$ is bounded above by the product of the integrated function's total variation and the discrepancy of the point set at which the function is evaluated. Formally, $\left| \frac{1}{N} \sum_{i=1}^{N} y(x_i) - \int_0^1 y(u)\, du \right| \leq V(y)\, D_N^*(x_1, \ldots, x_N)$, where the total variation is $V(y) = \int_0^1 |y'(x)| dx$ if y is differentiable and $V(y) = \max_{x \in [0,1]} y(x) - \min_{x \in [0,1]} y(x)$ if y is monotone and where $D_N^* = \sup_{t \in [0,1]} \left| \frac{|\{x_1, x_2, \ldots, x_N\}| \cap [0,t)|}{N} - t \right|$ is the discrepancy and measures how "concentrated" the points $\{x_i\}$ are. Discrepancy is maximal if all points coincide (if $x_i = \bar{x}\ \forall i$ then $D_N^* = \max\{\bar{x}, 1 - \bar{x}\}$) and minimal if all points are equidistant ($D_N^* = 1/N$). For uniformly distributed (pseudo-)random numbers expected discrepancy is proportional to $1/\sqrt{N}$, for quasi-random numbers[30] worst-case discrepancy is proportional to $(\ln N)/N$. The algorithm reduces the approximation error in two ways. First, it uses quasi-random rather than pseudo-random numbers. The difference can be quite substantial. For example, to reduce the approximation error by a factor of 10, one needs to increase N by a factor of 100 using a (pseudo-)Monte Carlo method, whereas with a quasi-Monte Carlo method a factor close to 10 will be sufficient. Second, the algorithm reduces total variation (in individual subregions) by iteratively dividing the interval of integration into subintervals (i.e. iterative partition refinement) in the following way. At the beginning of each iteration there are M subintervals $[a_j, b_j]$ with $\cup_j [a_j, b_j] = [0, 1]$. The subintervals are disjoint except at the boundaries. The interval k with the largest spread $(b_j - a_j)(\max_{x \in [a_j, b_j]} y(x) - \min_{x \in [a_j, b_j]} y(x))$ is chosen for subdivision. A cut c in this subregion is chosen such that $y(c) \approx \frac{1}{2}(\max_{x \in [a_k, b_k]} y(x) - \min_{x \in [a_k, b_k]} y(x))$ and region $[a_k, b_k]$ is divided into $[a_k, c]$ and $[c, b_k]$. The next iteration is conducted with the obtained $M + 1$ subregions. Iterations are repeated until the estimated integration error (estimated by the sum of spreads) is sufficiently small. After this partitioning phase, the subintervals are fixed. In the following integration phase, the same number of points is sampled from each subinterval $[a_j, b_j]$. For the multi-dimensional case, both

---

[30] Pseudo-random numbers are deterministically computed numbers that "behave like" random numbers for most practical purposes. Quasi-random number sequences are a notion from number theory. Such sequences are constructed to have a low discrepancy, at the price of not satisfying some properties of random numbers (in particular they fail to be i.i.d.). See Judd (1998, p.309) for an introduction to quasi-random numbers.

the algorithm and the definitions of total variation and discrepancy are more complex (the definitions are given in Judd (1998, p.309)). For the details of the workings of the algorithm we refer interested readers to the above-mentioned papers.[31]

The Divonne algorithm works well if subdivisions along the axes help to reduce variation. Correlated variables (which occurred in our problem) cause a problem as the optimal subdividing cuts are diagonal to the axes. To avoid the resulting problems and speed up computation, we do a first, low-precision pseudo-Monte Carlo computation of the mean $\hat{\mu}_\theta$ and covariance matrix $\hat{\Sigma}_\theta$ of the distribution of θ. This allows us to construct a transformed variable $\hat{\theta} = (\theta - \hat{\mu}_\theta)L^{-1}$ where L is the Cholesky decomposition with $LL' = \hat{\Sigma}_\theta$. $\hat{\theta}$ has a mean of approximately 0 and the covariance is approximately the identity matrix. As the distributions can be relatively well approximated by a multivariate Gaussian in some directions, but not in others (in particular fat tails and multiple local maxima), we chose a middle ground for further transformation. We transform $\hat{\theta}$ to $\tilde{\theta}$ by the inverse of the (fat tail) student's t distribution H_{student}. This has the advantage over a Gaussian transformation of avoiding numerical division by a number close to zero when computing $\int_{-\infty}^{\infty} y(\theta)d\theta = \int_0^1 y(H_{\text{student}}^{-1}(\tilde{\theta})L + \mu_\theta) \det(L)^{-1}(\prod_i h_{\text{student}}(\hat{\theta}_i))^{-1}d\tilde{\theta}$ at the tails, at the same time making the integrand relatively flat in approximately Gaussian directions. The Divonne algorithm can compute the integral quite efficiently using the transformed variable $\tilde{\theta}$.

The numerical evaluation of the optimal profit Π_{optimal} and the profit with optimal linear fees Π_{linear} can be computed much faster using revenue equivalence results based on mechanism design. In particular, the dominant strategy implementation of the optimal mechanism gives the same expected revenue as the implementation using optimal fees, but its revenue is much easier to compute numerically. The dominant strategy implementation that results in Π_{optimal} is the following: let the seller report his cost c and the buyer(s) report their valuations $\{v_i\}_{i=1}^k$. If $\max_i \Phi(v_i) \geq \Gamma(c)$, then trade occurs, otherwise it does not. The seller receives a payment which equals the

---

[31]For d dimensions, the theoretical worst-case discrepancy is $(\ln N)^d/N$. In practice, the average case discrepancy is typically much smaller. Further, the subdivision of intervals is replaced by the subdivision of hyperrectangles, which adds the additional complexity of having to choose the right dimension to cut.

highest cost he could have reported while still selling the good, that is $\max_i \Phi(v_i)$. The winning buyer i pays the lowest valuation he could have reported while still receiving the good, that is $\max\{\max_{j \neq i}\{v_j\}, \Gamma(c)\}$. It can be shown by a Vickrey-type of argument that it is a dominant strategy to report one's type truthfully and that the expected revenue is the same as charging a (possibly non-linear) fee $\omega_{\text{optimal}}(p)$. This transformation is reflected in (8), which can be evaluated faster than the expression that would result from plugging $\omega_{\text{optimal}}(\cdot)$ into (9).

Further, we obtain a fast approximation of the incomplete Beta function $B_x(\alpha, \beta) = \int_0^x t^{\alpha-1}(1-t)^{\beta-1}dt$ by using splines. First, we approximate $g(x) = (1-w)\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} + w$ by a monotone piecewise cubic Hermite spline, where w is the weight of the uniform distribution that is mixed with the beta distribution in order to avoid division by zero when computing the virtual cost Γ. Note that g can have a narrow peak at its modal value $x^* = (\alpha - 1)/(\alpha + \beta - 2)$ for α, β large. To deal with this, we take n points to the left of x^* and n points to the right and construct the interpolation data $\{(x_i, y_i)\}$ with $y_i = g(x_i)$.

Second, we take a monotone piecewise cubic Hermite spline approximation (Fritsch and Carlson, 1980) based on data $\{(x_i, y_i)\}$. A *monotone* spline ensures that, for example, the approximating function cannot be negative with non-negative data $\{y_i\}$. This gives us the interpolating polynomial function $\hat{g}(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i$ for $x \in [x_i, x_{i+1})$. $\hat{G}(x)$ can be obtained from the closed-form solution of the integral of the polynomial. $\hat{g}'(x)$ can be obtained from the relation $\hat{g}'(x) = \hat{g}(x)((\alpha - 1)/x + (\beta - 1)/(1 - x))$ (rather than the numerically imprecise derivative of the approximating polynomial). F, f, and f' can be approximated in a similar fashion.

References

ANDERSON, S., AND R. RENAULT (1999): "Pricing, Product Diversity, and Search Costs: A Bertrand-Chamberlin-Diamond Model," *RAND Journal of Economics*, 30(4), 719–735.

BAJARI, P. (1997): "The First Price Sealed Bid Auction with Asymmetric Bidders: Theory and Applications," *University of Minnesota, unpublished Ph. D. dissertation.*

BAJARI, P., AND A. HORTAÇSU (2003): "The winner's curse, reserve prices, and endogenous entry: empirical insights from eBay auctions," *RAND Journal of Economics*, 34(2), 329–355.

BERGEMANN, D., AND S. MORRIS (2005): "Robust Mechanism Design.," *Econometrica*, 73(6), 1771 – 1813.

BURDETT, K., AND D. MORTENSEN (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2), 257–273.

CHERNOZHUKOV, V., AND H. HONG (2003): "An MCMC approach to classical estimation," *Journal of Econometrics*, 115(2), 293–346.

CHU, L. Y., AND D. SAPPINGTON (2007): "Simple Cost-Sharing Contracts," *American Economic Review*, 97(1), 419–428.

CHU, S., P. LESLIE, AND A. SORENSEN (2011): "Bundle-size Pricing as an Approximation to Mixed Bundling," *American Economic Review*, 101(1), 263–303.

COLES, S. (2001): *An introduction to statistical modeling of extreme values*. Springer Verlag.

DONALD, S., AND H. PAARSCH (1993): "Piecewise pseudo-maximum likelihood estimation in empirical models of auctions," *International Economic Review*, pp. 121–148.

FRIEDMAN, J., AND M. WRIGHT (1981): "A nested partitioning procedure for numerical multiple integration," *ACM Transactions on Mathematical Software (TOMS)*, 7(1), 76–92.

FRITSCH, F., AND R. CARLSON (1980): "Monotone piecewise cubic interpolation," *SIAM Journal on Numerical Analysis*, 17(2), 238 – 246.

GENESOVE, D., AND L. HAN (2012): "Search and matching in the housing market," *Journal of Urban Economics*, 72(1), 31–45.

GENESOVE, D., AND C. J. MAYER (1997): "Equity and Time to Sale in the Real Estate Market," *American Economic Review*, 87(3), 255–269.

——— (2001): "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, 116(4), 1233–1260.

GLOWER, M., D. HAURIN, AND P. HENDERSHOTT (1998): "Selling Time and Selling Price: The Influence of Seller Motivation.," *Real Estate Economics*, 26(4), 719–721.

GUERRE, E., I. PERRIGNE, AND Q. VUONG (2000): "Optimal Nonparametric Estimation of First-price Auctions," *Econometrica*, 68(3), 525–574.

HAHN, T. (2005): "Cuba – a library for multidimensional numerical integration," *Computer Physics Communications*, 168(2), 78 – 95.

HENDEL, I., A. NEVO, AND F. ORTALO-MAGNÉ (2009): "The Relative Performance of Real Estate Marketing Platforms: MLS versus FSBOMadison.com," *American Economic Review*, 99(5), 1878 – 98.

HIRANO, K., AND J. PORTER (2003): "Asymptotic Efficiency in Parametric Structural Models with Parameter-Dependent Support," *Econometrica*, 71(5), 1307–1338.

HOLMSTROM, B., AND P. MILGROM (1987): "Aggregation and Linearity in the Provision of Intertemporal Incentives.," *Econometrica*, 55(2), 303 – 328.

HSIEH, C.-T., AND E. MORETTI (2003): "Can Free Entry Be Inefficient? Fixed Commissions and Social Waste in the Real Estate Industry," *Journal of Political Economy*, 111(5), 1076–1122.

JUDD, K. L. (1998): *Numerical Methods in Economics*. MIT Press, Cambridge MA.

KANG, H. B., AND M. J. GARDNER (1989): "Selling Price and Marketing Time in the Residential Real Estate Market," *Journal of Real Estate Research*, 4(1), 21–35.

KRISHNA, V. (2002): *Auction Theory*. Elsevier Science, Academic Press.

LAFFONT, J., H. OSSARD, AND Q. VUONG (1995): "Econometrics of first-price auctions," *Econometrica*, pp. 953–980.

LOERTSCHER, S., AND A. NIEDERMAYER (2012): "Fee Setting Intermediaries: On Real Estate Agents, Stock Brokers, and Auction Houses," *Working Paper*.

MCAFEE, P. (2002): "Coarse Matching," *Econometrica*, 70(5), 2025–34.

MERLO, A., AND F. ORTALO-MAGNÉ (2004): "Bargaining over residential real estate: evidence from England," *Journal of Urban Economics*, 56(2), 192–216.

MYERSON, R. (1981): "Optimal Auction Design," *Mathematical Operations Research*, 6(1), 58–73.

QUIGLEY, J. M. (1999): "Real Estate Prices and Economic Cycles," *International Real Estate Review*, 2(1), 1–20.

ROGERSON, W. (2003): "Simple Menus of Contracts in Cost-Based Procurement and Regulation," *American Economic Review*, 93(3), 91926.

SHILLER, B., AND J. WALDFOGEL (2011): "Music for a Song: An Empirical look at Uniform Pricing and its Alternatives," *Journal of Industrial Economics*, 59(4), 630–660.

SHNEYEROV, A. (2006): "An empirical study of auction revenue rankings: the case of municipal bonds," *RAND Journal of Economics*, 37(4), 1005–1022.

SHY, O., AND Z. WANG (2011): "Why Do Payment Card Networks Charge Proportional Fees?," *American Economic Review*, 101(4), 1575–1590.

WHEATON, W. (1990): "Vacancy, search, and prices in a housing market matching model," *Journal of Political Economy*, pp. 1270–1292.

WILSON, R. (1987): "Bidding," *The New Palgrave of Economics*, 1, 238 – 242.

WOLINSKY, A. (1986): "True monopolistic competition as a result of imperfect information," *Quarterly Journal of Economics*, 101(3), 493–511.