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

Simon Loertscher† Andras Niedermayer‡

December 1, 2011

Abstract

We use a structural model to estimate how much profit brokers make who employ simple percentage fee mechanisms relative to the profit they would make with 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 proportional fee mechanisms with fees ranging from 4% to 10% achieve 80% or more of the maximum profit. With the empirically observed 6% fees intermediaries achieve at least 75% of the maximum profit. With an optimally structured linear fee, they achieve 98% or more of the maximum profit. Moreover, we find that the theoretical model is consistent with a variety of stylized facts observed in real estate brokerage and with the joint distribution of prices and time on the market for houses observed in the data set.

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] At least on its surface,

*We want to thank Aviv Nevo and Harry J. Paarsch for comments early on in this project and Philipp Schmidt-Dengler for comments on the paper in its present form. We also want to thank David Genesove and Chris Mayer for providing us with their data set and Christian Michel for excellent research assistance. Financial support through a research grant by the Faculty of Business and Economics at the University of Melbourne is also gratefully acknowledged.

†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 (2009), Rogerson (2003), Shy and Wang (2011) and Hsieh and Moretti (2003).

this simplicity contrasts with the prescriptions and predictions of Bayesian mechanism design that 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, 2011), 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 rather than optimal mechanisms in the case of real estate brokerage. Using a structural model and data generously provided to us by Genesove and Mayer (2001), we show in our counterfactual analysis that intermediaries who employ the best proportional fee mechanisms with fees ranging from 4% to 10% achieve 80% or more of the maximum profit. With the empirically observed 6% fees intermediaries achieve at least 75% of the maximum profit. With an optimally structured linear fee, they achieve 98% or more of the maximum profit. Though there is obviously no clear cut metric to categorize these numbers, the fact that seemingly very little profit is lost through the use of very simple mechanisms suggests – to us at least – that concerns of practicality and economic principles may be rather well aligned.

That simple mechanisms can achieve a large percentage of the optimal surplus or profit has been shown in a variety of contexts. McAfee (2002) and Rogerson (2003) provide theoretical analysis of assortative matching and incentives in procurement, respectively. Recent empirical work includes Shiller and Waldfogel (2009) for uniform pricing by iTunes and Chu, Leslie, and Sorensen (2011) for bundling of theater tickers. To the very best of our knowledge, the present paper is the first to quantify the performance of simple fee setting mechanisms in the brokerage industry. By providing empirically based measures how well simple mechanisms fare that real estate brokers employ our paper also contributes to the literature on real estate and real estate brokerage, which has witnessed an upsurge of interest over the last decade; see e.g. Genesove and Mayer (2001), Hsieh and Moretti (2003), Rutherford, Springer, and Yavas (2005), Levitt and Syverson (2008), Hendel, Nevo, and Ortalo-Magné (2009), Genesove and Han (2010) and Jia and Pathak (2011).

The remainder of this paper is structured as follows. Section 2 introduced the theory. The empirical analysis and results are described in Section 3. Section 4 concludes. The appendix (not for publication) contains additional results and a description of the numerical methods.

2 Theory

The theoretical model is based on the dynamic random matching model analyzed in Loertscher and Niedermayer (2011, LN hereafter).[2] LN have the following stage game for a buyer viewing a real estate property that a seller offers for sale through a broker. The buyer's reservation price – the maximal amount he is willing to pay – v is his private information and drawn from the distribution F, which is common knowledge. The seller's opportunity cost of selling – the minimal 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 have to 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 have some (limited) assumptions about the equilibrium concept once we do the counterfactual analysis.

Theoretical Predictions As Hsieh and Moretti (2003) note, the fees charged by brokers are almost always very close to 6%. While the buyer and the seller often bargain over the transaction price, the transaction price is typically close to the listing price set by the seller. As an approximation, we model the behavior of market participants the following way. The seller sets a price p that maximizes his gains from trade times the

[2] For more details, see also the discussion paper version (Loertscher and Niedermayer, 2008).

probability of trade $(94\%p - c)(1 - F(p))$, the buyer accepts iff $p \leq v$, and in case of trade the broker gets 6% of the price p. The exposition becomes simpler by defining some additional functions. Define the virtual valuation function $\Phi(v) := v - (1 - F(v))/f(v)$ and the virtual cost function as $\Gamma(c) := c + G(c)/g(c)$. Further, denote the (price) elasticity of demand as $\eta_d(v) := vf(v)/(1 - F(v))$ and the (price) elasticity of supply as $\eta_s(c) := cg(c)/G(c)$.[3] Observe that $\Phi(v) = v(1 + 1/\eta_d(v))$ and $\Gamma(c) = c(1 + 1/\eta_s(c))$. The first-order condition for the maximization problem of seller with cost c reveals that he optimally sets the price $p = \Phi^{-1}(c/0.94)$.

If there is a trade, the buyer and the seller leave the market. If there is no trade, they stay in the market until the next rematching, which happens after time τ has passed. With exogenous probability $1 - \epsilon$ a seller or a buyer drop out of the market without having traded. The drop out probability is a simple device to take into account a variety of sources of impatience: traders may have deadlines until when they want to move, the buyer may have to rent temporarily at a high price if he does not buy, etc. The probability that a seller who sets the price p does not leave the market between two subsequent rematchings is $\epsilon F(p)$. Hence, the time on market t of a property has a geometric distribution with the distribution function $1 - (\epsilon F(p))^{t/\tau}$ and mean

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

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

$$1 - F_\infty(p) = \frac{1 - F(p)}{1 - \epsilon F(p)}. \tag{2}$$

When taking the model in Loertscher and Niedermayer (2011) to the data, we make three modifications of the baseline model. First, there is some heterogeneity in preferences: a buyer and a seller are a match with probability λ. Second, the price is a quality adjusted

[3] To see that these are indeed the price elasticities, interpret the probability with which at price p a buyer is willing to buy and a seller is willing to sell as quantity demanded and quantities supplied, respectively, which are denoted $q^d(p) := 1 - F(p)$ and $q^s(p) := G(p)$. 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.

price, which is observed with an error. Third, time on market is observed with an error. We will reiterate on this later on.

Theory for Counterfactual Analysis Now we can consider the model for the counterfactual analysis. We need one additional assumption for this: when the broker proposes the exchange mechanism for the buyer and seller, he faces a competitive threat of the buyer and seller going to a different broker in the future if they do not like his offer. We model it the following way. The broker can only extract rents from the current trade, but not from future trades. Hence, his 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, say, 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 incentive compatibility and individual rationality constraints of participants. Loertscher and Niedermayer (2011) show that no mechanism exists that generates higher profits than a fee setting mechanism defined as follows. The broker announces a fee function $\omega(p)$, then the seller sets a price p, the buyer either reject the offer or accepts it. In case of acceptance, the buyer pays the gross price p to the seller, who in turn pays the fee $\omega(p)$ to the broker. The optimal fee function is

$$\omega(p) = p - E_v[\Gamma^{-1}(\Phi(v))|v \geq p], \tag{3}$$

which induces a seller with cost c to set the price

$$P(c) = \Phi^{-1}(\Gamma(c)), \tag{4}$$

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

Constant elasticities of supply are helpful to get 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 distributions with linear virtual cost functions[4] $G(c) = [(c - \underline{c})/(\overline{c} - \underline{c})]^\alpha$ the optimal fee function is linear, $\omega(p) = p/(1+\alpha) - \underline{c}/(1+\alpha)$ (see Prop. 4 in LN). Note that for linear virtual cost distributions, the fees are determined solely by the seller's elasticity of supply and independent of the buyer's distribution.

LN identify two forces that lower the elasticity of supply and make it "more constant": one arising in a static setup (i.e. when $\epsilon = 0$) and one that comes from dynamics. To see where the static effect comes from, 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 gets 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. It also makes them "more linear" in the following sense and way. 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 (Prop. 2 in LN).

The dynamic effect is due to high cost sellers setting higher prices and having a lower probability of sales. Hence higher cost sellers are overrepresented in the market compared to the entrant population. This transformation from inflow to stock distribution increases the density of sellers with higher costs. This 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 Prop. 6 and 7 and 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.

In the current paper, we return to the ratio of linear to optimal profits and bring the theoretical model to the data. The optimal profit is obtained by setting the fee $\omega(p)$

[4]These are equivalent to distributions with linear inverse hazard rates $G(c)/g(c)$.

given by (3), which results in price $P(c)$ given by (4), probability of sale $1 - F(P(c))$, and expected profits

$$\Pi_{\text{optimal}} = E_c[\omega(P(c))(1 - F(P(c)))].$$

The best linear fee is obtained if we restrict the fee function to be linear, $\omega_{\xi,\zeta}(p) = \xi p + \zeta$, and maximize with respect to the slope ξ and the intercept ζ of the fee:

$$\Pi_{\text{linear}} = \max_{\xi,\zeta} E[\omega_{\xi,\zeta}(P_{\xi,\zeta}(c))(1 - F(P_{\xi,\zeta}(c)))],$$

where $P_{\xi,\zeta}(c) = \Phi^{-1}((c + \zeta)/(1 - \xi))$ is the price set by a seller facing fee $\omega_{\xi,\zeta}$.[5] Besides the performance of linear fees, we will also estimate the profits generated by the best proportional fee function $\Pi_{\text{proportional}} = \max_{\xi} E[\omega_{\xi,0}(P_{\xi,0}(c))(1 - F(P_{\xi,0}(c)))]$ and by the six percent fee that is a reasonable approximation of the fee typically chosen by brokers, $\Pi_{6\%} = E[\omega_{0.06,0}(P_{0.06,0}(c))(1 - F(P_{0.06,0}(c)))]$.

3 Empirical Analysis

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 date of 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 of a property, which Genesove and Mayer (2001) (and we) use to account for unobserved property heterogeneity in constructing a quality adjusted price as discussed above. The data set includes the listing of property from January 6, 1990 to December 28, 1997 and delisting (either due to sale or withdrawal) from May 10, 1990 to March 16, 1998. We only use data from January 1, 1993 to December 31, 1996 to avoid truncation problems. This gives us 3167 observations. We exclude data with a quality adjusted price larger than 2 and less than 1/2 and properties that were on the market for more than 2 years. This applies to 4.8% of observations.

Table 1 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

[5] $P_{\xi,\zeta}(c)$ is the solution of the first-order condition of the sellers maximization problem $\max_p((1 - \xi)p - \zeta)(1 - F(p)))$.

Ortalo-Magné (2004, Table 1) for two regions in the U.K. Observe also that unsold houses stayed longer on the market than houses that did sell. If a house is delisted and relisted within 4 weeks, it is considered to be the same transaction. For further details about the data, see Genesove and Mayer (2001).

Variables	All Houses	Sold Houses	Unsold Houses
Observations	3014	1987	1027
Listing Price	\$223,337 (\$179,860)	\$234,702 (\$183,341)	\$201,349 (\$170,880)
Quality Adjusted Listing Price	1.140 (.225)	1.128 (.224)	1.163 (.226)
$100\frac{\text{Listing Price}}{\text{Transaction Price}}$		91%	
Time on Market	143 (131) Days	125 (120) Days	178 (144) Days

Table 1: Descriptive Statistics: Sample Means (Standard Deviations)

Empirical Model We use the quality index constructed by Genesove and Mayer (2001) that uses previous transaction prices and the movement of the real estate price index P_t^{index}. Formally, at time t the 'objective' quality ϑ_i of house i that had been traded previously at time s for the price $\hat{P}_{is}$ is

$$\vartheta_i = \frac{P_t^{\text{index}}}{P_s^{\text{index}}} \hat{P}_{is}.$$

The quality adjusted price p_i of property i is with listing price $\hat{P}_i$ is then given as $p_i = \hat{P}_i / \vartheta_i$.[6]

While having a complex empirical model taking into account many effects besides those of our model would be interesting, looking how well our baseline model with minimal modifications can explain observations is equally interesting and informative. We add three modifications to our baseline model. First, the true quality adjusted price p_i is observed with an error ϵ_i^P, hence the observed quality adjusted price is $P_i = p_i + \epsilon_i^P$. One can imagine ϵ^P coming from an imperfect measurement of the quality index. Second, the true time on market t_i is also observed with an error (ϵ_i^T), hence $T_i = t_i + \epsilon_i^T$.

[6]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 s and of structural changes in demand from time s to time t.

ϵ^T can come from the broker starting to show the property only some time after it is listed, a property being delisted with some delay after the buyer and the seller agreed on a deal (because of negotiations taking time, the broker delisting the property with some delay after negotiations ended, and the like). In the data set we use (described above), most properties are listed and delisted on a Sunday, Thus, we have essentially weekly data and delay happens at least until the end of the week. Third, we return to the assumption of some simple heterogeneity of houses. We assume that a buyer likes a houses with probability λ, the overall probability of being willing to buy the house at the quality adjusted price p is the probability of liking it λ times the probability of finding the price acceptable $1 - F(p)$. We take this into account by replacing F with F_λ where F_λ is implicitly defined by $1 - F_\lambda(p) = \lambda(1 - F(p))$.

Denote the joint density of prices and times on market as predicted by our baseline model as $h_{tps}(p, t, s)$, the density of the error term for the time on market ϵ^T as $h_t(\epsilon^T)$, and the density of the error term for the quality adjusted price ϵ^P as $h_p(\epsilon^P)$. Let the measured time on market T_i and the measured quality adjusted price P_i be given as $T_i = t_i + \epsilon^T$ and $P_i = p_i + \epsilon^P$, where t_i and p_i represent the true values. Let s_i be 1 if the house was sold and 0 otherwise.

We first describe the prediction of the baseline model without error terms. Denote with $P_I(p) = 0.94\Phi(p)$ the empirical inverse price function (i.e. a seller with cost c will set the price $\Phi^{-1}(c/.94)$). Since 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 get the *entrant* distribution of prices, we have to divide by the time on market and get the *entrant* density of prices $\sigma g(P_I(p))P_I'(p)/T(p) =: g_{p0}(p)$, where $T(p)$ is the average time on market and σ is a constant making sure density adds up to 1.

Since rematching happens every τ periods, and a house drops out of the market with probability $1 - \epsilon$ and is sold with probability $1 - F_\lambda(p)$, the probability that a house is still in the market after t periods is $(\epsilon F_\lambda(p))^{t/\tau}$. Hence the joint distribution of p, t, and

s has the density

$$h_{tps}(t,p,s) = \begin{cases} (1-F_\lambda(p))(\epsilon F_\lambda(p))^t g_{p0}(p) & \text{if } s=1, \\ (1-\epsilon)F_\lambda(p)(\epsilon F_\lambda(p))^t g_{p0}(p) & \text{if } s=0. \end{cases} \tag{5}$$

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

$$l(X_i|\theta) = \sum_{k=1}^{T_i/\tau} \int_{-\infty}^{\infty} h_{tps}(T_i - k\tau, P_i - \epsilon^P, s_i) h_t(k\tau) h_p(\epsilon^P) d\epsilon^P \tag{6}$$

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

Identification If quality adjusted price and time on market were observable without the errors ϵ^P and ϵ^T, it would be rather obvious that our model is non-parametrically identifiable given observations of quality adjusted price, time on market and whether a house was sold. Rearranging (1) and (2), the expressions for the 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 get

$$\begin{aligned} 1 - F(p) &= \frac{1 - F_\infty(p)}{T(p)/\tau}, \\ \epsilon &= \frac{T(p_2) - T(p_1)}{T(p_2)(1 - F_\infty(p_1)) - T(p_1)(1 - F_\infty(p_1))}, \\ \tau &= \frac{T(p_1)F_\infty(p_2) - T(p_2)F_\infty(p_1)}{F_\infty(p_2) - F_\infty(p_1)}, \end{aligned}$$

where p_1 and p_2 are two arbitrary prices (or – with some modification of the equations – price segment).[7] This makes F, ϵ, and τ non-parametrically identifiable. Note that the distinction between F and F_λ is irrelevant both for identification and for the best response fees we calculate.[8] Assuming that a seller with cost c chooses the profit maximizing price

[7]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.

[8]The distribution F_λ can be seen as a distribution with a mass point with probability $(1-\lambda)$ at 0 and a distribution F with probability λ. The distinction between a distribution F with heterogeneity factor λ on the one hand and a distribution F_λ and homogeneity on the other hand boils down to how

$\Phi^{-1}(c/(1-0.06)) = \arg\max_p((1-0.06)p-c)(1-F(p))$ given 6% fees, the distribution of sellers' costs G is non-parametrically identifiable by the distribution of prices G_p: $G(c) = G_p(\Phi^{-1}(c/(1-0.06)))$. Having F and G our theory provides the unique best-response fee setting mechanism of the broker.

With measurement errors of quality adjusted price ϵ^P and time on market ϵ^T, the argument is more involved, but identification is still possible. Denote the observed quality adjusted price as $P = p + \epsilon^P$, and redefine the observed probability of ever selling $1 - F_\infty(P)$ and the time on market of sold and unsold houses $T^s(P)$ and $T^u(P)$ as depending on the observed quality adjusted price P rather than the true (quality adjusted) price p. Further, let $\hat{T}^s(P) := T^s(P) + E[\epsilon^T]$ and $\hat{T}^u(P) := T^u(P) + E[\epsilon^T]$ be the observed average times on market. The probability of ever selling given P and ϵ_P is $\text{Prob}(s = 1|P, \epsilon^P) = (1-F(P-\epsilon^P))/(1-\epsilon F(P-\epsilon^P))$ and the probability of never selling $\text{Prob}(s = 0|P, \epsilon^P) = (1-\epsilon)F(P-\epsilon^P)/(1-\epsilon F(P-\epsilon^P))$. Given the unconditional density $h_p(\epsilon^P)$, the conditional densities are $h_p(\epsilon^P|P, s = 1) \propto h_p(\epsilon^P)/\text{Prob}(s = 1|P, \epsilon^P)$ and $h_p(\epsilon^P|P, s = 0) \propto h_p(\epsilon^P)/\text{Prob}(s = 0|P, \epsilon^P)$ by Bayes' Law. This gives us

$$
\begin{aligned}
1 - F_\infty(P) &= E_{\epsilon^P \sim H_p}\left[\frac{1 - F(P-\epsilon^P)}{1-\epsilon F(P-\epsilon^P)}\right], \qquad (7)\\
\hat{T}^s(P) &= E_{\epsilon^P \sim H_p(\cdot|P,s=1)}\left[\frac{\tau}{1-\epsilon F(P-\epsilon^P)}\right] + E[\epsilon^T],\\
\hat{T}^u(P) &= E_{\epsilon^P \sim H_p(\cdot|P,s=0)}\left[\frac{\tau}{1-\epsilon F(P-\epsilon^P)}\right] + E[\epsilon^T]
\end{aligned}
$$

Note that (7) and $\hat{T}^u(P) - \hat{T}^s(P)$ do not require any knowledge about the distributions of ϵ^T and c and identify F and H_p for given ϵ and τ. The density of time on market t conditional on a particular price P, $E_{\epsilon^P \sim H_p, \epsilon^T \sim H_t}[(\epsilon F(P-\epsilon^P))^{t/\tau - \epsilon^T} | \epsilon^T \leq t/\tau]$ identifies H_t. This uses only one price P. The different densities of t for two additional prices P pin down ϵ and τ. Now only the seller's distribution G remains to be identified. It is non-parametrically identifiable the same way as without measurement errors ϵ^P and ϵ^T:

buyers behave at future transactions. For the former, a buyer who had a low valuation today (because the property was a bad fit with probability $1-\lambda$) might have a high valuation for the next property he looks at. For the latter, the buyers willingness will be low for the next transaction as well. Since we have no data on buyer behavior, the two cases cannot be distinguished non-parametrically. (The distinction does not matter for the choice of the optimal fee either.) However, they are parametrically distinguishable if one assumes that the distribution F has no atoms.

$G(c) = G_p(\Phi^{-1}(c/(1-0.06)))$.

Estimation Procedure While our model is in principle non-parametrically identifiable, we use a parametrization to get a Bayesian estimate of the parameters. The main reason is that our main hypothesis is – loosely speaking – that if one were to pick distributions F and G randomly, then "most of the time" (or "on average") linear fees would perform close to optimally. Drawing F and G from the Bayesian posteriors distribution is a natural choice and also gives a clearer meaning to "on average".

For our estimation we make the following functional form assumptions and take the following parametrization. We assume that $F(v)$ 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 $\theta = (\alpha_F, \beta_F, \alpha_G, \beta_G, \beta_T, \sigma_p, \epsilon, \underline{v}, \overline{v}, \underline{c}, \overline{c}, \lambda)$.[9] 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 $l(X)$. We are looking for the posterior beliefs about the parameters θ given X, $\pi(\theta|X)$.[10] By Bayes' Law $\pi(\theta|X) = l(X|\theta)\pi(\theta)/l(X)$, where $\pi(\theta)$ is the prior about θ.

[9]Notice that θ does not include the period length τ, which is the only parameter that we cannot estimate. This is partly due to the fact that we use a discrete time model, so that the distribution of the error in time on market ϵ^T cannot be compared across different τ and 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.

[10]We use lower case π to denote beliefs and keep using upper case Π to denote expected profits.

Assuming a uniform prior $\pi(\theta)$, we get the proportionality $\pi(\theta|X) \propto l(X|\theta)$ since $l(X)$ does not depend on θ.[11]

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

Computing the expectations requires computing a 10-dimensional integral. This turns out to be numerically quite challenging, since a simple approach would lead to a computational time of several years (estimate based on extrapolation of time needed to solve part of the problem). We use several numerical improvements to reduce computational time to a few hours (see Appendix B).

While our main interest is a Bayesian estimate of the different variables, for illustrational purposes (to plot an example of the predicted distribution of prices, relation price/time on market, etc.) we also want a pointwise estimator for θ. For this we take the (constrained) Maximum Likelihood Estimator $\theta_{\mathrm{MLE}} = \arg\max_\theta l(X|\theta)$.[12]

Results Table 2 contains the results for all four years for which we have data under the assumption that there are 7 matchings per week (i.e. $\tau = 1/365$). The results for the alternative assumptions of 2 and 14 weekly matchings are summarized in Appendix A, which also provides the parameter estimates for all years for the specification with 7 matchings per week. According to Table 2 (third row), intermediaries who used a 6% fee achieved between 75% and 88% of the optimal profit, given the parameter estimates. The optimal proportional fee, which is displayed in row 4, varies between 4% and 10% over

[11] 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 f and g), (c) $\overline{v} - \underline{v} \geq 0.2$ (to avoid numerical problems with nearly degenerate distributions that arise when $\overline{v} \approx \underline{v}$).

[12] Alternatively, one could also use the Bayesian mean $\theta^* = E[\theta]$. However, some of the constraints imposed on θ (e.g. increasing virtual valuations) by setting $\pi(\theta|X) = 0$ where constraints are violated, may not be satisfied at θ^*. They do hold for θ_{MLE}.

Variables	1993	1994	1995	1996
$\frac{\text{linear profit}}{\text{optimal profit}}$	0.979(0.032) [0.003]	0.993(0.059) [0.01]	0.978(0.058) [0.003]	0.983(0.001) [0.009]
$\frac{\text{proportional profit}}{\text{optimal profit}}$	0.883(0.019) [0.002]	0.883(0.018) [0.009]	0.814(0.013) [0.002]	0.840(0.008) [0.008]
$\frac{6\%\text{ profit}}{\text{optimal profit}}$	0.749(0.035) [0.002]	0.882(0.02) [0.009]	0.812(0.014) [0.002]	0.758(0.0397) [0.007]
opt. proportional fee	0.103(0.01) [0.0003]	0.06(0.002) [0.001]	0.06(0.003) [0.0002]	0.044(0.003) [0.0004]
opt. slope (linear fee)	0.407(0.02) [0.001]	0.424(0.024) [0.004]	0.478(0.016) [0.001]	0.501(0.016) [0.004]
opt. fixed component (linear fee)	−0.337(0.027) [0.001]	−0.396(0.023) [0.004]	−0.46(0.018) [0.001]	−0.494(0.02) [0.004]
$\frac{6\%\text{ profit}}{\text{linear profit}}$	0.764(0.036) [0.002]	0.889(0.022) [0.009]	0.831(0.013) [0.002]	0.771(0.041) [0.007]
$\frac{6\%\text{ profit}}{\text{proportional profit}}$	0.848(0.046) [0.002]	0.999(0.008) [0.010]	0.998(0.004) [0.003]	0.902(0.042) [0.008]

Table 2: Fees and Profits Implied by the Model and the Parameter Estimates (1993 to 1996, 7 matchings per week). Table entries read: Mean (Standard Deviation) [Computational Error].

the four 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 2% (see the first row in Table 2).[13] In Table 3 we report the Bayesian estimates for the parameter vector θ.

Figure 1 provides a graphical illustration using the Maximum Likelihood Estimates of θ for the year 1993. Panel (a) displays the estimated endogenous densities $f(v)$ in green and $g(c)$ in red. 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 dashed line while the relationship that our model implies when evaluated at the estimated parameter values is shown as solid line. The densities of time on the market

[13] A possible explanation why proportional fees rather than linear fees are used by real estate brokers is that the fixed component of the linear fee varies with the quality ϑ of the house. Though we assume that ϑ is observable for buyers, sellers and brokers, it may be impossible or prohibitively costly to write contracts that depend on it.

Parameter	1993	1994	1995	1996
α_F	3.672(0.476) [0.004]	2.900(0.648) [0.033]	2.774(0.305) [0.004]	7.357(0.483) [0.061]
β_F	5.473(0.438) [0.006]	3.010(0.491) [0.033]	6.203(1.274) [0.010]	7.917(0.512) [0.065]
α_G	9.786(1.317) [0.012]	4.196(0.700) [0.046]	7.227(1.989) [0.010]	9.534(0.501) [0.077]
β_G	9.343(0.751) [0.011]	3.615(0.29) [0.037]	12.054(3.467) [0.018]	10.458(0.883) [0.084]
β_T	15.967(0.622) [0.02]	11.214(1.159) [0.124]	10.301(0.831) [0.018]	9.808(0.508) [0.081]
σ_p	0.231(0.004) [0.000]	0.212(0.003) [0.002]	0.216(0.003) [0.0003]	0.225(0.003) [0.001]
ϵ	0.961(0.002) [0.001]	0.971(0.003) [0.010]	0.978(0.002) [0.002]	0.971(0.002) [0.007]
$\underline{v}$	0.659(0.027) [0.000]	0.872(0.014) [0.009]	0.886(0.024) [0.001]	0.909(0.016) [0.007]
$\overline{v}$	1.480(0.099) [0.002]	1.267(0.023) [0.013]	1.383(0.051) [0.002]	1.239(0.020) [0.010]
λ	0.556(0.083) [0.001]	0.300(0.049) [0.003]	0.715(0.070) [0.001]	0.755(0.074) [0.005]

Table 3: Estimated Parameter Values for 1993 to 1996 for 7 matchings per week.

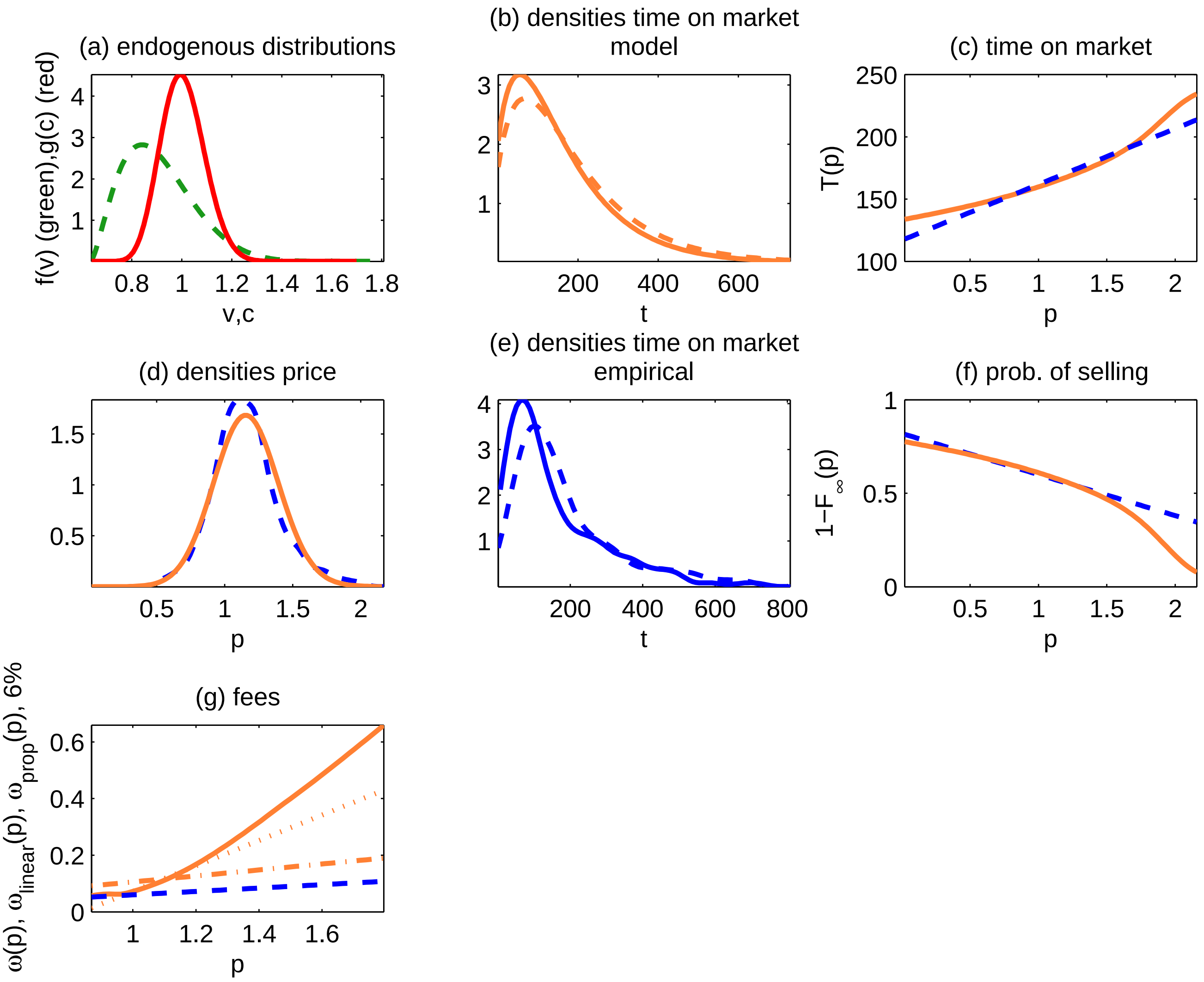

Figure 1: Theory and Empirics (1993, 7 matchings per week). The graphs are based on (constrained) maximum likelihood estimates.

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. Panel (g) shows the optimal fee $\omega(p)$ as a solid curve, the optimal linear fee $\omega_{linear}(p)$ as dotted line and the optimal proportional fee $\omega_{prop}(p)$ as dashed lined as implied by our model evaluated at the estimated parameter values. The empirically observed 6% fee is displayed as a dashed line. Figure 1 indicates that our model is a fairly good fit to the data. However, it should also be mentioned that the correlation between time on the market and quality adjusted price is weak and is negative for certain years and matching frequencies.

4 Conclusions

In this paper we have used a structural model based on Loertscher and Niedermayer (2011) to estimate demand and supply parameters for a data set that is due to Genesove and Mayer (2001) and covers the Boston condominium market in the 1990s. Using these parameter estimates for a counterfactual analysis, we have found that for our preferred parametersization with seven matchings per week the empirically observed 6% fee achieves 75% 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 quite some variation in the data set. Even very savvy brokers will therefore 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. Thus, they resemble robust mechanisms in the sense of Wilson (1987). In light of this, the performance of the 6% fee is even more remarkable.

Appendix

A Robustness (not for publication)

In this appendix we report robustness checks. Tables 4 and 5 summarize the resulting fees and profits when alternative rematching frequencies are chosen (2 and 14 rematchings per week rather than 7).

Table 6 performs an additional robustness check based on an alternative quality index constructed by Genesove and Mayer (2001) that relies on physical characteristics rather than the previous transaction price. Given the vector Y_i of physical characteristics of property i (such as the number of bedrooms, whether there is air conditioning, etc.), Genesove and Mayer (2001) run a regression with the observed listing price as dependent variable:

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

For our purposes, we interpret $Y_i\beta = \vartheta_i$ as the quality index and the residual $\epsilon_i = p_i$ as the quality adjusted price.[14] Tables 4, 5, and 6 suggest that our results are robust to a change of the matching frequency and of the quality index.

B Numerical Methods (not for publication)

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.

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, (6) can be transformed to

$$\sum_{k=1}^{T_i/\tau} \int_{\underline{p}}^{\overline{p}} h_{tps}(k\tau, p, s_i) h_t(T_i - k\tau) h_p(P_i - p) dp, \tag{8}$$

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

Note the difference in terms of numerical evaluation of (6) on the one hand and (8) with (5) on the other hand. For the former, one has to evaluate a three dimensional

[14] Note that the quality adjusted prices based on physical characteristics are a biased estimate, since the residuals ϵ_i are correlated.

Variables (2 matchings/week)	1993	1994	1995	1996
$\frac{\text{linear profit}}{\text{optimal profit}}$	0.965(0.008) [0.0006]	0.952(0.012) [0.010]	0.976(0.002) [0.001]	0.993(0.002) [0.006]
$\frac{\text{proportional profit}}{\text{optimal profit}}$	0.848(0.024) [0.0005]	0.772(0.026) [0.009]	0.850(0.013) [0.001]	0.924(0.012) [0.005]
$\frac{\text{6\% profit}}{\text{optimal profit}}$	0.759(0.025) [0.0004]	0.745(0.047) [0.008]	0.839(0.015) [0.001]	0.746(0.054) [0.004]
opt. proportional fee	0.093(0.005) [0.00006]	0.0740(0.011) [0.0009]	0.053(0.002) [0.00006]	0.040(0.002) [0.0002]
opt. slope (linear fee)	0.412(0.020) [0.0002]	0.472(0.019) [0.005]	0.418(0.016) [0.0005]	0.426(0.040) [0.002]
opt. fixed component (linear fee)	−0.3526(0.023) [0.0002]	−0.436(0.013) [0.005]	−0.403(0.018) [0.0004]	−0.420(0.045) [0.002]
$\frac{\text{6\% profit}}{\text{linear profit}}$	0.787(0.022) [0.0005]	0.781(0.041) [0.009]	0.859(0.014) [0.001]	0.751(0.054) [0.004]
$\frac{\text{6\% profit}}{\text{proportional profit}}$	0.895(0.023) [0.0005]	0.9643(0.033) [0.011]	0.987(0.008) [0.001]	0.807(0.050) [0.004]

Table 4: Fees and Profits Implied by the Model and the Parameter Estimates (1993 to 1996, 2 matchings per week).

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 quadra-

Variables (14 matchings/week)	1993	1994	1995	1996
$\frac{\text{linear profit}}{\text{optimal profit}}$	0.964(0.002) [0.0009]	0.980(0.004) [0.004]	0.980(0.008) [0.002]	0.985(0.004) [0.0006]
$\frac{\text{proportional profit}}{\text{optimal profit}}$	0.801(0.010) [0.0008]	0.850(0.017) [0.004]	0.794(0.097) [0.002]	0.757(0.019) [0.0005]
$\frac{\text{6\% profit}}{\text{optimal profit}}$	0.451(0.030) [0.0004]	0.730(0.043) [0.003]	0.768(0.103) [0.002]	0.747(0.020) [0.0005]
opt. proportional fee	0.182(0.013) [0.0002]	0.098(0.012) [0.0004]	0.070(0.012) [0.0002]	0.065(0.006) [0.00004]
opt. slope (linear fee)	0.547(0.016) [0.0005]	0.494(0.016) [0.002]	0.512(0.084) [0.001]	0.616(0.031) [0.0004]
opt. fixed component (linear fee)	−0.415(0.010) [0.0004]	−0.443(0.016) [0.002]	−0.479(0.078) [0.001]	−0.594(0.037) [0.0004]
$\frac{\text{6\% profit}}{\text{linear profit}}$	0.468(0.031) [0.0004]	0.745(0.044) [0.003]	0.783(0.102) [0.002]	0.758(0.020) [0.0005]
$\frac{\text{6\% profit}}{\text{proportional profit}}$	0.563(0.037) [0.0005]	0.859(0.056) [0.004]	0.967(0.036) [0.002]	0.986(0.016) [0.0006]

Table 5: Fees and Profits Implied by the Model and the Parameter Estimates (1993 to 1996, 14 matchings per week).

ture.)[15] 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 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

[15] One of the reasons may be that some of the variables are strongly correlated (in terms of posterior distribution). Another is that the posterior density has multiple peaks (due e.g. 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.

Variables (phys.characteristics)	1993	1994	1995	1996
$\frac{\text{linear profit}}{\text{optimal profit}}$	0.976(0.006) [0.008]	0.990(0.002) [0.009]	0.990(0.002) [0.006]	0.980(0.002) [0.009]
$\frac{\text{proportional profit}}{\text{optimal profit}}$	0.835(0.094) [0.007]	0.898(0.020) [0.008]	0.891(0.015) [0.005]	0.829(0.010) [0.008]
$\frac{\text{6\% profit}}{\text{optimal profit}}$	0.613(0.074) [0.005]	0.892(0.029) [0.008]	0.851(0.021) [0.005]	0.785(0.027) [0.007]
opt. proportional fee	0.129(0.008) [0.001]	0.057(0.004) [0.0005]	0.0490(0.002) [0.0003]	0.048(0.003) [0.0004]
opt. slope (linear fee)	0.489(0.072) [0.004]	0.421(0.035) [0.004]	0.465(0.020) [0.003]	0.494(0.014) [0.005]
opt. fixed component (linear fee)	−0.416(0.057) [0.003]	−0.416(0.043) [0.004]	−0.478(0.023) [0.003]	−0.488(0.017) [0.005]
$\frac{\text{6\% profit}}{\text{linear profit}}$	0.628(0.076) [0.005]	0.901(0.028) [0.008]	0.859(0.020) [0.005]	0.800(0.028) [0.008]
$\frac{\text{6\% profit}}{\text{proportional profit}}$	0.733(0.031) [0.006]	0.993(0.014) [0.009]	0.956(0.018) [0.006]	0.947(0.03) [0.010]

Table 6: Fees and Profits Implied by the Model and the Parameter Estimates (1993 to 1996, 7 matchings per week) when physical characteristics are used to control for house quality.

is proportional to $1/\sqrt{N}$, for quasi-random numbers[16] 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 10, one needs to increase N by a factor 100 with 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 subdividing the interval of integration into subintervals (i.e. iterative partition refinement) the following way. At the beginning of each iteration there are M subintervals $[a_j, b_j]$ with $\cup_j [a_j, b_j] = [0, 1]$ and the subintervals being 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 done 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 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 the interested reader to the above-mentioned papers.[17]

The Divonne algorithm works well if subdivisions along the axes help reduce variation. Correlated variables (which happen to occur in our problem) cause a problem, since the optimal subdividing cuts are diagonal to the axes. To avoid the resulting problems and to 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

[16] 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 e.g. Judd (p.309 1998) for an introduction to quasi-random numbers.

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

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 mean approximately 0 and covariance approximately the identity matrix. Since 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 choose a middle ground for the 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 $\int_{-\infty}^{\infty} y(\theta)d\theta = \int_0^1 y(H^{-1}_{\text{student}}(\tilde{\theta})L + \mu_\theta)\det(L)^{-1}(\prod_i h_{\text{student}}(\hat{\theta}_i))^{-1}d\tilde{\theta}$ at the tails, at the same time making the integrand relatively flat in the 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} can be computed much faster by using the revenue equivalence result that $\Pi_{\text{optimal}} = E_c[\omega(P(c))(1 - F(P(c)))]$ is equal to $\int_{\Phi^{-1}(\underline{c})}^{\overline{v}} \int_{\underline{c}}^{\Gamma^{-1}(\overline{v})} \max\{\Phi(v) - \Gamma(c), 0\}dcdv$. Linear (and proportional) profits can be evaluated much faster using a change of variables from costs c to prices $p = P_{\xi,\zeta}(c) = \Phi^{-1}((c+\zeta)/(1-\xi))$ and from the intersect of the fee function ζ with the lowest price of trade $\underline{p} = \Phi^{-1}((\underline{c}+\zeta)/(1-\xi))$, which is equivalent to $\zeta = (1-\xi)\Phi(\underline{p}) - \underline{c}$:

$$\begin{aligned}\Pi_{\xi,\zeta} &= E_c[(\xi P_{\xi,\zeta}(c) + \zeta)(1 - F(P_{\xi,\zeta}(c)))] \\ &= \int_{\underline{p}}^{\overline{v}} (\xi p + (1-\xi)\Phi(\underline{p}) - \underline{c})(1 - F(p))g((1-\xi)(\Phi(p) - \Phi(\underline{p})) + \underline{c})(1-\xi)\Phi'(p)dp.\end{aligned}$$

This avoids the need to compute the inverse of Φ and allows us to maximize over the simple rectangular boundary $(\xi, \underline{p}) \in [0,1] \times [\underline{v}, \overline{v}]$.

The maximization problem $\max_{\xi,\underline{p}} \Pi_{\xi,\zeta}$ can be further sped up by having an initial guess $\xi_0, \underline{p}_0$ (or ξ_0, ζ_0) close to the optimum. Prop. 2 suggests that $p - \Gamma^{-1}(p)$ is a good approximation of the fee function $\omega(p)$. We choose the initial guess to minimize $\int_{\underline{p}}^{\overline{v}}[(p - \Gamma^{-1}(p)) - (\xi p + \zeta)]^2 dp = \int_{\underline{c}}^{\Gamma^{-1}(\overline{v})}[(1-\xi)\Gamma(c) - c - \zeta]^2\Gamma'(c)dc$. Using a Chebyshev polynomial approximation of the integrand, the expression becomes $\begin{bmatrix} \zeta \\ 1-\xi \end{bmatrix}^T A \begin{bmatrix} \zeta \\ 1-\xi \end{bmatrix} + 2b^T \begin{bmatrix} \zeta \\ 1-\xi \end{bmatrix}$, where $A = \begin{bmatrix} S(1) & -S(\Gamma(c_i)) \\ -S(\Gamma(c_i)) & S(\Gamma(c_i)^2) \end{bmatrix}$, $b = \begin{bmatrix} S(c_i) \\ S(c_i\Gamma(c_i)) \end{bmatrix}$, $S(x_i) := \sum_{i=1}^n w_i\Gamma'(c_i)x_i$, and w_i and c_i are the weights and nodes of the Chebyshev quadrature rule. Solving the first order conditions of this quadratic minimization problem yields $\begin{bmatrix} \zeta_0 \\ 1-\xi_0 \end{bmatrix} = -A^{-1}b$.

Further, we get 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. In order 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 makes sure e.g. that 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

Chu, S., P. Leslie, and A. Sorensen (2011): "Bundle-size Pricing as an Approximation to Mixed Bundling," *American Economic Review*, 101(1), 263–303.

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 (2010): "Search and Matching in the Housing Market," *CEPR Discussion Papers*, (7777).

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

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.

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.

Jia, P., and P. Pathak (2011): "The Costs of Free Entry: Evidence from Real Estate Agents in Greater Boston," *Working Paper, MIT.*

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

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

Loertscher, S., and A. Niedermayer (2008): "Fee Setting Intermediaries: On Real Estate Agents, Stock Brokers, and Auction Houses," *Northwestern University Discussion Paper Nr. 1472.*

——— (2011): "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.

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

Rutherford, R., T. Springer, and A. Yavas (2005): "Conflicts between Principals and Agents: Evidence from Residential Brokerage," *Journal of Financial Economics*, 76, 627–65.

Shiller, B., and J. Waldfogel (2009): "Music for a Song: An Empirical Look at Uniform Song Pricing and its Alternatives," *Mimeo, Wharton.*

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

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