Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Financial Economics 76 (2005) 627–665

JOURNAL OF Financial ECONOMICS

www.elsevier.com/locate/econbase

Conflicts between principals and agents: evidence from residential brokerage[☆]

R.C. Rutherford[a], T.M. Springer[b], A. Yavas[c,*]

[a]University of Texas at San Antonio, San Antonio, TX 78249, USA
[b]Florida Atlantic University, Jupiter, FL 33458, USA
[c]Smeal College of Business, The Pennsylvania State University, University Park, PA 16802, USA

Received 12 September 2002; received in revised form 28 May 2004; accepted 11 June 2004
Available online 3 March 2005

Abstract

When a homeowner uses an agent to sell his property, he may have less information than his agent and be disadvantaged in price setting and negotiating. This study examines whether the percentage commission structure in real estate brokerage creates agency problems. We investigate whether agents are able to use their information advantage to either sell their own property faster or for a higher price than their clients' properties. The empirical results confirm our theoretical predictions of agency problems, as we find that agent-owned houses sell no faster than client-owned houses, but they do sell at a price premium of approximately 4.5%.

JEL classification: D8; J4; K12; L0

Keywords: Agency theory; Percentage commission contract; Residential brokerage

☆We have benefitted from the comments of Paul Anglin, Richard Arnott, Peter Colwell, Donald Haurin, Stanley Hamilton, Patrick Hendershott, Cemile Yavas, an anonymous reviewer, the participants at the 1998 Research Workshop on Real Estate Brokerage at Ohio State University, the 1999 ASSA Meetings in New York, NY and seminar participants at University of Colorado, Rice University and University of Vermont.
*Corresponding author. Tel.: +1 814 865 0392.
E-mail address: ayavas@psu.edu (A. Yavas).

0304-405X/$ - see front matter
doi:10.1016/j.jfineco.2004.06.006

1. Introduction

Sellers in real estate markets face imperfect information about the market value of their assets and the location of potential buyers. This creates a role for real estate agents who can help sellers both by employing their superior search technology to locate prospective buyers and by using their superior knowledge of the market to suggest an optimal asking price. In return for their services, agents typically receive a percentage (generally 6%) of the transaction price as commission. The significance of the role played by agents is evident: they account for the sale of approximately 81% of single-family dwellings in the housing market (see Federal Trade Commission, 1983). With a sales volume of $1348.2 billion, the estimated total commission income from the sales of existing homes alone exceeded $65.5 billion in 2003 (see www.realtor.org).

The relationship between a home-seller and his agent is a perfect example of a principal-agent problem.[1] Agents have a fiduciary responsibility to represent the interest of their principals (sellers) to the best of their ability, that is, in the same way they would represent themselves. However, whether or not an agent would advance his principal's interest at his own expense is not clear.

The standard agency models in general, and the models of real estate brokerage in particular, argue that the percentage commission system creates an agency problem between an agent and his client because it induces the agent to expend too little effort. A recent study by Williams (1998) questions this result. Williams' (1998) theoretical work shows that a more realistic modification of standard agency models eliminates any agency problems between agents and their clients under the percentage commission system. This is a significant departure from the earlier models and has important policy implications. The significance of the debate is further augmented by the fact that it applies to most classes of professionals, including accountants, consultants, and lawyers. However, there has been no empirical test put to this debate so far. The principal purpose of this study is to offer empirical evidence on whether the percentage commission structure in the brokerage industry creates any agency problems. The paper also offers a theoretical framework that captures the essential ingredients of the agency problem between the agent and his client under the percentage commission structure.

In order to examine whether the agent acts in the best interests of his client, we compare the effort level and the pricing strategy of an agent for two scenarios, the first being that the agent owns the listed property and the second being that the agent is not the owner. We utilize a data set that is sufficiently comprehensive to allow for the identification of sold properties that were owned by listing agents. Thus, we can directly test for the selling price and marketing time differences that would provide evidence of any agency problems. In a world of perfect information with no agency problems, an agent would seek the same price and use the same effort for a contracted seller as he would in selling his own house. That is, equivalent assets will

[1] Note that, although the agent may represent the buyer instead of the seller, the focus in this paper will be on the agency problem between the agent and the seller.

remain in the market for the same length of time and sell for the same price regardless of whether or not they were sold by an owner-agent.

The agency problem in markets for real assets has been the focus of various recent studies. The general conclusion of the earlier models is that although the percentage commission system ensures the interests of the agent to be in the same *direction* as those of the client, it fails to align the *magnitude* of the interests of the agent with those of the client. Geltner et al. (1991), Anglin and Arnott (1991), and Miceli (1991) point out the agency problems with respect to the agent's search effort level and argue that the agent will expend less than the efficient level of effort for the seller. Arnold (1992) studies the pricing aspect of the agency problem and shows that the agent's reservation price for the seller's house will generally be different than the seller's reservation price. The source of the agency problem in these models can be summarized as follows. The seller wants to maximize the selling price while minimizing the time the asset stays on the market. The agent, on the other hand, seeks to maximize his expected commission revenue while minimizing the time on the market. Given that the agent receives a small portion of the transaction price as commission, the agent's goal of maximizing the expected commission may diverge from the seller's goal of maximizing the selling price. Furthermore, given that the targeted selling price will impact the time the asset stays on the market, the agent's desired time on the market may diverge from that of the seller. Since the agent's search efforts are not directly observable and the agent has asymmetric information about the market value of the asset, this divergence can lead to shirking by the agent. As a result, the agent may be motivated to convince the seller to underprice the asset at the listing and/or attempt to induce the seller to accept a suboptimal selling price at the bargaining in order to facilitate a faster sale and obtain a more timely commission. Williams (1998) points out that the existence of such agency problems in earlier models is due to unrealistic assumptions. Specifically, earlier models typically match one agent with one seller and ignore the fact that agents have to search for new clients as well as find buyers for their existing clients. Williams' model allows for multiple sellers per agent and requires agents to spend time selling the assets of current clients and searching for new clients. His conclusion is supported by some of the theoretical studies on optimal contracting which also show that under certain conditions, the optimal contract between a principal and an agent involves a linear compensation scheme (e.g., Diamond, 1998; Hart and Holmstrom, 1987; Holmstrom and Milgrom, 1987; McAfee and McMillan, 1986). Williams was the first to build a formal equilibrium model of brokerage to prove that percentage commission contract can be incentive compatible.

The current paper offers the first direct empirical test of the agency problem in real estate brokerage. Our unique data set allows us to investigate whether real estate agents treat their own properties differently than they do those of their clients.

The next section of the paper presents the predictions of a simple search model of agent behavior that we formally develop in Appendix A. Our theoretical model suggests that an agent will sell his own house for a higher price than a comparable listed house he does not own. For any given listing price, the agent will also expend more effort for his own house. However, since a higher price makes it less likely for a contacted buyer to purchase the house, the time it takes to sell the agent-owned house may be longer or shorter than that of a house not owned by the agent. Section 3 provides an overview of the data. In Section 4 we discuss the estimation of the models. We estimate duration models to examine the effect on time on the market, probit models to generate inverse Mills ratios to correct for sample selection bias and endogeneity, selling price models corrected for sample selection bias, and selling price models corrected for selection bias and endogeneity using data from 306,869 multiple listing service (MLS) listings from several metropolitan counties in Texas during the years 1999 through 2002. Over the study period, 63% of the listings resulted in completed sales. An important feature of this data set is that over 9,911 (3.23%) of the observations involved agent-owned properties.

The empirical results presented in Section 5 show that agent-owned houses sell no faster than client or non-agent-owned houses, but they do sell at a price premium of approximately 4.5% above the price of a comparable house not owned by a real estate agent. This result that the owner-agent is able to obtain a higher selling price is consistent with the prediction of our theoretical model that an agent's interest does not match the interests of the seller. The final section of the paper, Section 6, summarizes the results and offers some concluding remarks.

2. Theoretical predictions

In Appendix A of the paper we present a theoretical model of the agency problem between a real estate agent and his client. We study the problem of a risk-neutral listing agent who chooses both an effort level and a price for a property. We first study the case in which the listing agent is the owner of the property. We then compare this to the case in which the listing agent is not the owner of the property. The purpose of the model is not to design the efficient and incentive compatible contract between the seller and the listing agent. Rather, we take the current commission structure in the industry as given and compare the effort level and price choices of the agent depending on whether he is the owner of the listing or not. The model is representative of the standard models of brokerage with the exception that it captures agency problems with regard to both pricing the client's asset and expending search effort to sell the client's asset. Except for Williams (1998), previous models study either the pricing or the search effort dimension of the problem, but not both.

Our theoretical model incorporates the tradeoff facing property owners in selling their properties; while a higher price is desirable, it decreases the probability of finding a buyer who is willing to pay the price. The theoretical model offers two testable results.

Result 1. The agent will choose a higher price for his own property than he will for his client's property.

This result is due to the fact that the agent retains a much larger portion of a marginal increase in the selling price if he owns the property than if he does not.

Result 2. The marketing time for a property owned by an agent may be shorter or longer than that of a property not owned by the agent.

Although the agent expends a greater effort to search for a buyer for his own property than for a client's property, this does not necessarily result in a shorter marketing time because the agent sets a higher price for his own property, hence decreasing the probability that a contacted buyer purchases the property.

A formal proof of the above two results is provided in Appendix A.

3. Data

The data consist of observations of residential properties (single-family homes) listed between January 1998 and December 2002, collected from a large metropolitan MLS consisting of several counties in Texas. The initial data set has a total of 348,715 single-family residential observations. Due to missing values for some variables included in the models and extreme values of certain variables that suggest data entry problems, the final data set consists of 306,869 observations. The observed dates of sale or withdrawal are between January 1999 and December 31, 2002. Data variable names and definitions are provided in Table 1.

Data collected from the MLS include physical property characteristics (age, square footage, number of fireplaces, pool, large lot), market descriptors (geographic location, calendar information, interest rates), and variables that reflect marketability or possible investor differences between the properties (vacant, tenant-occupied, builder). The data contain numerical MLS-defined geographic variables that we use to control for location in the models. Calendar information includes the listing quarter (Winter, Spring, Summer, and Fall) and the quarter of sale or withdrawal from the market along with the year. These variables are intended to capture contemporaneous changes in market conditions.

Other variables include the list price, selling price, and days on the market. For those properties that are ultimately sold, days on the market is measured as the number of days from the listing date reported in the MLS to the date a sales contract is executed and the property is removed from the market. Similarly, the days on the market for houses that went off the market without a sale is calculated as the number of days between the listing date and the day the property went off the market. Because the MLS does not provide information on properties that are relisted, our calculations of days on the market may be less than the actual time a house remains on the market. For example, if the property is listed with one agent and does not sell within the contract time frame, the seller may select another agent and relist the property. Houses both owned and listed by a real estate agent are identified with a binary variable, *Owner-agent*; a value of "0" denotes a non-agent-owned house.

In Table 2, descriptive statistics are provided for the full sample with a comparison of non-agent-owned and agent-owned houses, depending on whether a real estate agent owned the listed property. The results of differences in means tests comparing the characteristics of the sample of non-agent-owned houses to those involving

Table 1
Definition of variables used in the time-on-the-market hazard models, the selling price regression models, and the probit models

Variable	Description
Selling price	Selling price of the house, expressed as log(sp) in the regression models.
List price	Listed price on the multiple listing service (MLS).
Size	Number of square feet divided by 100.
Age	Year of sale minus year built divided by ten.
Owner-agent	Dummy variable indicating a listing by an agent who is also identified as the owner of the property.
Owner-agent (1–2 listings)	Dummy variable indicating that an owner-agent listed 1–2 properties owned by that agent.
Owner-agent (3–5 listings)	Dummy variable indicating that an owner-agent listed 3–5 properties owned by that agent.
Owner-agent (>5 listings)	Dummy variable indicating that an owner-agent listed more than 5 properties owned by that agent.
Pool	Dummy variable indicating the presence of a pool.
Fireplace	Number of fireplaces.
Large lot	Dummy variable indicating lot size greater than one-half acre.
Tenant	Dummy variable indicating a rental property occupied by a tenant.
Vacant	Dummy variable indicating a vacant property.
Builder	Dummy variable indicating a builder-owned house.
Uniqueness	Haurin's measure of uniqueness of a house.
Market competition	Number of active listing agents in a MLS defined area divided by the total number of listings in the same MLS defined area, multiplied by 100.
Degree of overpricing (DOP)	Residual from a list price equation.
Interest rate change	Change in the level of annual percentage mortgage rates calculated as the difference in the weekly interest rate at the time of listing to the weekly interest rate either when the property sold, went off the market, or the sample period ended.
Open house	Dummy variable indicating the use of an open house in marketing the property.
Tour house	Dummy variable indicating the house was included in a tour of homes.
Listing agent limited experience	Dummy variable indicating the listing agent has less than 3 years of experience.
Listing agent experienced	Dummy variable indicating the listing agent has 5 or more years of experience.
Listing agent broker	Dummy variable indicating the listing agent holds a broker's license.
Internet	Dummy variable indicating the listing was marketed on the internet.
Large firm	Dummy variable indicating the listing agent is associated with a larger firm in the top 20th percentile based on number of agents working for the same sponsoring broker.
Time-on-the-market (DOM)	Calculated number of days from list date until date sold, went off the market, or the end of the sample period, expressed as log(time) in the duration models.
Sold property	Dummy variable indicating the property sold.

houses owned by real estate agents reveal a number of statistically significant differences. The typical agent-owned houses are statistically larger (mean = 2,345 versus 2,247 ft^2), older (20.4 years versus 17.6 years), and on larger lots (17.2%

versus 14.3%), but the actual differences appear to be marginal. With regards to price, on average, agent-owned houses are both listed (list price $227,512 versus $187,989 for all properties listed; $188,573 versus $167,109 for the 193,007 houses that sold) and sold at higher prices (selling price of $184,373 versus $163,368 for the 193,007 houses that sold) than non-agent-owned houses.[2] In addition, agent-owned houses appear to take about the same amount of time to sell as non-agent-owned houses (97 days versus 95 days).

There are several differences relating to marketability and ownership. First, agent-owned houses are more likely to be vacant (44% compared to 27%), occupied by a tenant (3.5% versus 2.8%), and owned by a builder (8.9% versus 8.1%). Each of these variables suggests the possibility of a motivated seller or an investor. A vacant house indicates either that the owner has already moved and thus needs to sell or that an investor is holding the house without a tenant. Whether by choice or by circumstances, the tenant variable clearly indicates ownership by an investor. A circumstantial investor would be the case in which an individual intended to sell the property but failed and thus rented it to mitigate losses. A builder by default is an investor and possibly a motivated seller, given that he has elected to list the property on the MLS and incur the commission costs, rather than market the property directly. Although the ownership or marketability variables could suggest a motivation to discount the price, it is also possible that as an indicator of investor activity, these variables suggest an intention to hold the properties until a profit can be made.[3]

Furthermore, unlike some of the sellers, owner-agents may not have the same timing constraints as non-agent owners in selling their properties. To examine the possibility that an owner-agent is primarily an investor with a real estate license, we also proxy for investor properties by examining the number of properties sold by each owner-agent. We create three variables, *Owner-agent* (1–2 *listings*), *Owner-agent* (3–5 *listings*), and *Owner-agent* (>5 *listings*), to account for situations where multiple houses owned by the same agent were listed during the study period. Of the

[2]As indicated by an anonymous referee, potential property tax incentives may exist to under-report the selling price. We do know that recording fees and commissions are based on the actual selling price and are collected at closing by the title company, thus mitigating the incentive that an individual might have to under-report the selling price. In addition, the listing agent and his broker are responsible for correctly entering all data for the property, thus limiting the opportunity for under-reporting. The owner-agent has the most opportunity for under-reporting since he is the listing agent, but given our results, this does not appear to be the case. Furthermore, if an agent is detected to have misreported a price, he typically will face sanctions from the National Association of Realtors if he is a member, state and local boards of realtors, and the state real estate commission. In one small set of data (119 sales), we are able to obtain the sales price from the Tarrant County Tax Assessors office in which they state that the selling price data was obtained independently from the MLS and then match it with the MLS data. The correlation was 1.00. This indicates that the sales price is most likely reported correctly in the MLS, but of course, the sample is quite small.

[3]Springer (1996) examines in a smaller sample the impact of various factors that indicate a motivated seller on the marketing time and selling prices of single-family houses. The results indicate an approximate 2.5% selling price discount and 19.5% increase in time on the market associated with vacant houses. The selling price discount is consistent with the behavior of a seller that is motivated. The prolonged marketing time is consistent with poor asset marketability.

Table 2
Descriptive statistics for the full sample and subsamples of non-agent-owned and agent-owned houses listed on the MLS in several Texas counties during 1999–2002

Excluding houses with missing observations, the final sample includes 306,869 houses, of which 9,911 are listed and owned by a real estate agent. Univariate test statistics for the difference in characteristics between agent-owned and non-agent-owned houses are presented. The *t*-statistics are calculated to test the null: mean(agent-owned) - mean(non-agent-owned) = 0, with the assumption that the two subsamples are random and independently selected and the sampled population is approximately normal. The nonparametric Wilcoxon statistic is used to test whether the agent-owned and non-agent-owned houses have identical distributions, with the assumption that the two samples are random and independent. Statistics with significance at the 1% level are denoted with a** and the 5% level are denoted with a*.

Summary statistics and univariate tests of key variables	Full sample		Non-agent-owned		Agent-owned houses		*t*-Statistics	Wilcoxon rank sum test
Variable	Mean	Median	Mean	Median	Mean	Median		
Selling price	163,979	133,500	163,368	133,500	184,373	135,000	9.33**	3.76**
List price	189,178	147,000	187,898	146,900	227,512	154,900	17.06**	10.91**
Size	22.500	20.300	22.470	20.290	23.450	20.790	9.19**	4.50**
Age	1.770	1.400	1.765	1.400	2.038	1.600	14.46**	11.35**
Owner-agent	0.032	0.000	—	—	1.000	1.000	—	—
Owner-agent (1–2 listings)	0.021	0.000	—	—	0.638	1.000	—	—
Owner-agent (3–5 listings)	0.005	0.000	—	—	0.168	0.000	—	—
Owner-agent (>5 listings)	0.006	0.000	—	—	0.194	0.000	—	—
Pool	0.163	0.000	0.163	0.000	0.163	0.000	−0.05	−0.05
Fireplace	0.968	1.000	0.967	1.000	0.997	1.000	4.53**	2.17*
Large lot	0.144	0.000	0.143	0.000	0.172	0.000	7.41**	7.94**

Tenant	0.028	0.000	0.028	0.000	0.035	0.000	3.77**	4.18**
Vacant	0.280	0.000	0.274	0.000	0.443	0.000	33.32**	36.77**
Builder	0.081	0.000	0.081	0.000	0.089	0.000	2.86**	2.98**
Uniqueness	0.570	0.497	0.567	0.495	0.654	0.541	19.34**	17.66**
Market competition	39.148	37.879	39.390	37.690	40.380	38.760	8.00**	9.14**
Degree of overpricing (DOP)	0.000	0.002	0.000	0.002	0.000	−0.005	0.00	−0.63**
Interest rate change	−0.079	−0.080	−0.078	−0.080	−0.113	−0.100	−8.89**	−8.58**
Open house	0.029	0.000	0.028	0.000	0.035	0.000	3.48**	3.82**
Tour house	0.067	0.000	0.066	0.000	0.081	0.000	5.43**	5.94**
Listing agent limited experience	0.084	0.000	0.084	0.000	0.091	0.000	2.39*	2.48*
Listing agent experienced	0.803	1.000	0.803	1.000	0.801	1.000	−0.52	−0.52
Listing agent broker	0.288	0.000	0.284	0.000	0.401	0.000	25.17**	27.27**
Internet	0.938	1.000	0.939	1.000	0.891	1.000	−15.35**	−19.69**
Large firm	0.517	1.000	0.522	1.000	0.373	0.000	−30.16**	−29.22**
Time-on-the-market (days)	95.031	76.000	94.970	76.000	96.720	74.000	−2.06**	−4.10**
Sold property	0.629	1.000	0.631	1.000	0.566	1.000	−12.78**	−13.10**
Sample size	306,869		296,958		9,911			
Sample size for the selling price variable above	193,007		187,393		5,614			

9,911 houses listed by owner-agents, the median number of houses listed by an owner-agent is two, with 63.8% of the houses listed by owner-agents that listed one to two owner-agent properties, 16.8% listed by owner-agents that listed three to five owner-agent houses, and 19.4% of the houses listed by owner-agents that listed six or more owner-agent houses. Of the 4,158 listing owner-agents, 56.4% listed one house, 21.2% listed two houses, 16.2% listed three to five houses, and 6.2% of the owner-agents listed six or more houses. These variables allow us to reflect the extent to which an agent may be involved in buying and selling houses as an investment activity. In addition to the physical, marketability, and ownership characteristics, three additional variables are created. The first variable controls for the uniqueness of a house, which has been shown to increase marketing time (Haurin, 1988). This measure is calculated for each house as the sum of the product of the implicit price of each feature and the absolute value of the deviation of that feature, across all features of the house. The second controls for the level of market competition in each MLS-defined location: *market competition* is defined as the number of listing agents active in an area divided by the number of listings in the area multiplied by 100. The last variable is *degree of overpricing* (*DOP*) which measures the percentage by which the list price exceeds the expected list price estimated from a regression model.

One issue that the above variables fail to address is the effort level of the agent. Given owner-agents may be more motivated and expend more effort, we attempt to obtain information that relates to the effort level of the listing agent. We are able to determine whether the listing was included in a tour of homes (8.1% for owner-agents versus 6.6% for non-owner-agents), whether the agent held an open house (3.5% for owner-agents versus 2.8% for non-owner-agents), and whether the agent posted the listing on the internet or withheld it from the internet (89.1% posted for owner-agents versus 93.9% for non-owner agent listings). Also, we obtain data from the Texas Real Estate Commission that indicate the year the agent obtained a license and whether the agent held a real estate broker's license or a real estate agent's license. Over 80% of the listing agents had five or more years of experience with no significant difference between owner-agents and other agents. However, for agents with less than three years of experience, 9.1% were owner-agents compared to 8.4% for non-owner-agent listings, and 40.1% of owner-agents were brokers compared to 28.4% for non-owner-agent listings. These statistics suggest that owner-agents may use open houses and housing tours in an attempt to obtain a better price, possibly indicating additional effort on the agent's part. Clearly, more owner-agents are brokers, but also relatively more owner-agents have less than three years of experience.

4. Methods

The first step in the analysis is to estimate the typical list price for a house described by X under market conditions described by M. The list price (LP) model is

$$E(\log(LP)) = \alpha_X X + \alpha_M M. \tag{1}$$

Because specification testing indicates the presence of heteroskedasticity, the list price model is estimated by generalized least squares (GLS). The residual of the list price model is used to estimate the degree of overpricing, *DOP*, the percentage deviation from a expected list price for a house described by *X* and *M*, where *DOP* is calculated as log(*LP*)−E(log(*LP*); *X*, *M*). Note that *DOP* is expected to influence the time on the market, DOM.

Next, we specify the DOM model with DOM being a function of the characteristics of the house, market conditions, uniqueness, competition, list price residuals, and ownership, where uniqueness, competition, DOP, effort level, and experience are represented by *Z* in the equation below. This equation allows for the assessment of the impact of agent-owned houses on marketing time. Many researchers have used ordinary least squares (OLS) to estimate a DOM model. This method is known to produce unbiased estimates, but is also known to waste information. For example, in a "single risk" model, Lancaster (1990, Chapter 8.8) claims that using a semi-log OLS model to estimate the determinants of DOM is equivalent to throwing away 39% of the data if the true model is exponentially distributed and 43% of the data if a Weibull distribution is more appropriate. In light of this fact, we estimate DOM using a hazard model with a Weibull specification of the baseline hazard function:

$$f(t|X, M, Z,\ \textit{Owner-agent}) = \varphi\lambda(X, M, Z,\ \textit{Owner-agent})^{\varphi} \times t^{\varphi-1}\exp(-(\lambda(X, M, Z,\ \textit{Owner-agent}) * t)^{\varphi}), \tag{2}$$

where φ is a duration dependency parameter, λ is a scaling parameter, t is DOM, and other variables are as previously described. We use a proportional hazards specification to explain the contribution of the independent variables, where

$$\lambda(X, M, Z,\ \textit{Owner-agent}) = \exp(-\beta_X X - \beta_M M + g(\textit{Owner-agent})) \tag{3}$$

for some function $g(.)$.

We modify this likelihood in two ways. First, we assume that unmeasured heterogeneity in the hazard function can be described by a Gamma distribution with mean one and variance θ. Second, the observed DOM is the minimum of two random variables, the time-to-sale and the time-to-withdrawal. Whether a seller is observed selling the house or withdrawing from the market depends on which of these events occurs first. The fact that a seller can withdraw without selling introduces censoring into the duration data which misleadingly shortens the average DOM. The variable *Sold Property* is a binary variable indicating whether a property was sold (*Sold Property* = 1) or withdrawn. For those houses which were withdrawn from the market at time t, the probability that the time-to-sale exceeds t is

$$1 - F(t|X, M, Z,\ \textit{Owner-agent}) = \exp(-(\lambda(X, M, Z,\ \textit{Owner-agent}) * t)^{\varphi}). \tag{4}$$

The maximum likelihood estimates of β, φ, and θ correct for this random and frequent censoring. See Lancaster (1990) for further discussion.

The second model is a typical house price model corrected for sample selection bias following the labor economics literature for wage equations, where one has information on the characteristics of the individual but no wage data for

Table 3
Regression models of house prices corrected for possible sample selection bias based on the probability of sale using Heckman's selection model, based on the full sample during 1999–2002 of 306,869 houses listed on the multiple listing service (MLS) in several Texas counties, of which 193,007 are sold and 113,862 are withdrawn or not sold by the end of the sample period.

The dependent variable is the log of the selling price. All models include quarterly dummy variables (not reported for brevity) to control for potential serial effects and all regressions include dummy variables for MLS specified areas (not reported for brevity) to control for location. The ML estimates of the coefficients are presented in the table, with *t*-statistics reported in parentheses using heteroskedasticity-robust Huebner/White standard errors. Statistics with significance at the 1% level are denoted with a ** and at the 5% level are denoted with a *.

Independent variable	Model 1	Model 2	Model 3	Model 4
Constant	10.98	10.95	10.98	10.95
	(5034.43)**	(2263.51)**	(5034.34)**	(2263.31)**
Size	0.041	0.042	0.041	0.042
	(586.53)**	(596.42)**	(586.68)**	(596.49)**
Age	−0.052	−0.050	−0.052	−0.050
	−(150.15)**	−(144.84)**	−(150.12)**	−(144.81)**
Owner-agent	0.045	0.046		
	(25.18)**	(25.77)**		
Owner-agent (1–2 listings)			0.040	0.040
			(17.65)**	(17.91)**
Owner-agent (3–5 listings)			0.049	0.051
			(10.06)**	(10.59)**
Owner-agent (>5 listings)			0.057	0.058
			(15.85)**	(16.35)**
Pool	0.07	0.08	0.07	0.08
	(74.00)**	(80.52)**	(74.04)**	(80.56)**
Fireplace	0.07	0.07	0.07	0.07
	(75.69)**	(74.09)**	(75.73)**	(74.13)**
Large lot	0.08	0.09	0.08	0.09
	(65.57)**	(69.63)**	(65.57)**	(69.63)**
Tenant	−0.08	−0.07	−0.08	−0.07
	−(40.17)**	−(38.06)**	−(40.17)**	−(38.06)**
Vacant	−0.06	−0.06	−0.07	−0.06
	−(79.73)**	−(77.07)**	−(79.86)**	−(77.20)**
Builder	0.06	0.05	0.06	0.05
	(39.52)**	(35.86)**	(39.45)**	(35.79)**
Uniqueness		−0.09		−0.09
		−(50.40)**		−(50.39)**
Market competition		0.001		0.001
		(9.64)**		(9.68)**
Number of observations	306,869	306,869	306,869	306,869
Number of sold obs.	193,007	193,007	193,007	193,007
Number of unsold obs.	113,862	113,862	113,862	113,862
Log likelihood	−70,656	−69,029	−70,534	−68,904

Table 3 (*continued*)

Independent variable	Model 1	Model 2	Model 3	Model 4
Wald chi^2 (model)	1,800,000	1,930,000	1,810,000	1,940,000
Inverse Mills ratio from a probit model for sold versus not sold to correct for possible sample selection	−0.028 −(16.85)**	−0.021 −(14.06)**	−0.028 −(16.85)**	−0.021 −(14.08)**

unemployed individuals. For our model, we have housing characteristics for all the properties, but no selling price for the 113,862 houses that did not sell during the sample period.

The selling price model is

$$\ln(SP_i) = \beta_0 + \beta_1 \textit{ Owner-agent} + \sum \beta_i X_i + \varepsilon_i, \tag{5}$$

where the vector X_i is similar to the X_i for the time on the market equation, except that the season variables relate to the season of the sale and the overpricing variable is excluded. Table 3 provides the results for the selling price model corrected for possible sample selection bias. The price model includes one additional variable, the inverse Mills ratio (*IMR*). For housing studies and many other applications, the Heckman model (Heckman, 1978) is an appropriate method when one suspects sample selection bias. The Heckman selection model corrects for selectivity bias by adjusting the conditional error terms using the inverse Mills ratio so that the conditional error terms will have zero means. In general, sample selection bias refers to the case in which a dependent variable is only observed for a restricted, non-random sample. In this study, we only observe a selling price if the house is actually sold. In the estimation we take account of this possible selection bias, where in the first stage a probit model (see Table 4, Model 1) is used to predict the probability of a house being sold and in the second stage the *IMR* is included as a regressor in the housing price model. All regression estimates are by maximum likelihood estimation. The dependent variable in the probit model is the variable *Sold Property* and the independent variables are those included in the selling price model plus the variables related to the listing season, excess time on the market, defined as the residual from a log linear time-on-the-market model, and a variable measuring the rate of change in interest rates. While this regression with the *IMR* corrects for possible sample selection bias based on whether or not the property was sold, the issue of endogeneity of the owner-agent variable is unresolved. We attempt to correct for the

Table 4

The results below are for three binary dependent variable models. The first model is a probit model, where the dependent variable equals one for a sold house and zero for an unsold house. From this probit model the inverse Mills ratio is calculated and included in the selling price models in Tables 3 and 5 to correct for possible sample selection bias of sold versus unsold properties. In the second model, the results are for a probit model, where the dependent variable equals one if the owner of the house is also a real estate agent and zero for houses not owned by a real estate agent. The inverse Mills ratio calculated using the results from this probit model is included in selling price models in Table 5 to correct for the possible endogeneity of Owner-agent. In addition, the predicted probability from this model is used in place of Owner-agent in Table 11 of Appendix B, as an alternative means of correcting for possible endogeneity of Owner-agent.

Model three is a linear probability model for Owner-agent, where the binary dependent variable is one if the owner of the house is a real estate agent. The use of a linear probability model and the predicted probability from this model in place of the owner-agent variable has been suggested as a means of correcting for endogeneity and has been implemented in Table 11 of Appendix B. In addition, one other approach suggested in the literature to correct for endogeneity is to include the residual from the linear probability model and the owner-agent variable in a selling price model. This is also implemented as shown in Table 11 of Appendix B.

The dependent variable in all three models is a dummy variable as described above. All models include quarterly dummy variables (not reported for brevity) to control for potential serial effects and all regressions include dummy variables for MLS specified areas (not reported for brevity) to control for location. The estimates of the coefficients are presented in the table, with *t*-statistics reported in parentheses using heteroskedasticity-robust Huebner/White standard errors. Statistics with significance at the 1% level are denoted with a** and at the 5% level are denoted with a*.

Independent variable	Model 1: probit model for sold versus not sold	Model 2: probit model for owner-agent	Model 3: linear probability model for owner-agent
Constant	1.02	−10.37	−0.59
	(24.60)**	−(58.19)**	−(61.60)**
Size	−0.04	−0.69	−0.05
	−(87.63)**	−(73.52)**	−(82.14)**
Age	−0.01	−0.16	−0.01
	−(4.03)**	−(19.22)**	−(28.57)**
Owner-agent	−0.12		
	−(8.76)**		
Large lot	−0.18	−4.00	−0.28
	−(24.10)**	−(54.25)**	−(77.86)**
Tenant	−0.26	−4.68	−0.35
	−(18.07)**	−(49.68)**	−(71.73)**
Vacant	0.10	2.02	0.16
	(16.77)**	(71.46)**	(77.00)**
Builder	0.06	0.89	0.07
	(5.57)**	(23.46)**	(33.32)**
Uniqueness	0.24	2.08	0.24
	(19.16)**	(39.79)**	(64.71)**
Market competition	0.01	0.15	0.01
	(6.68)**	(39.25)**	(55.27)**
Degree of overpricing (DOP)		−0.00	−0.00
		−(2.50)	−(1.59)
Interest rate change	0.06		
	(6.34)**		
Open house	−0.16	−3.05	−0.22
	−(10.65)**	−(37.04)**	−(61.16)**

Table 4 (*continued*)

Independent variable	Model 1: probit model for sold versus not sold	Model 2: probit model for owner-agent	Model 3: linear probability model for owner-agent
Tour house	0.19	3.29	0.26
	(15.69)**	(54.40)**	(72.19)**
Listing agent limited experience	−0.08	−1.79	−0.12
	−(7.54)**	−(33.85)**	−(53.40)**
Listing agent experienced	0.19	3.28	0.25
	(24.60)**	(60.35)**	(76.58)**
Listing agent broker	−0.06	−0.86	−0.07
	−(10.50)**	−(31.56)**	−(45.62)**
Internet		−0.29	−0.02
		−(10.23)**	−(10.21)**
Large firm		−0.22	−0.01
		−(9.39)**	−(9.62)**
Residual from a time-on-the-market	−0.00	−0.02	−0.00
Log linear model	−(26.17)**	−(66.16)**	−(77.21)**
Number of houses sold	193,007	193,007	193,007
Log likelihood	−186,656	−13,314	
Wald chi^2 (model)	28,732	8,297	
Pseudo R^2	0.08	0.48	
F			55.58
Adjusted R^2			0.18

possible endogeneity of *Owner-agent* by estimating a second probit model, where the dependent variable is *Owner-agent* and the independent variables are similar to the probit model for *Sold Property,* with the addition of *Internet* and *Large Firm* (see Table 4, Model 2). The results are discussed below.

5. Results

5.1. Selling price effects corrected for possible sample selection bias

The results for the full sample are provided in Tables 3–6. In Table 3, we present the results for the selling price model corrected for sample selection bias. For each of the models the coefficients for the variables included in the regression are mostly consistent in sign and magnitude with prior research. Housing prices are expected to increase with increasing size, more bathrooms, more fireplaces, and the existence of a pool. Housing prices are expected to decrease with increasing age and for vacant and tenant-occupied houses.

The variable of interest, *Owner-agent*, in Model 1 has a positive and statistically significant coefficient. This result suggests that real estate agents who list their own houses obtain a selling price premium of approximately 4.5% above the price of an equivalent house not owned by a listing agent, after controlling for physical characteristics and including the variables *Tenant*, *Vacant*, and *Builder* to account for the possibility of investor activity.[4]

In the third model shown in Table 3, we include *Tenant*, *Vacant*, and *Builder* as well as the owner-agent variable separated into three groups according to the number of his own houses that an individual agent listed over the study period. The results affirm that owner-agents are able to sell their properties at higher prices. The coefficient on *Owner-agent* (1–2 *listings*) is 4.0%. For *Owner-agent* (3–5 *listings*), the coefficient is 4.9% and for *Owner-agent* (>5 *listings*), the coefficient is 5.7%, suggesting that possible investors hold out for higher prices. In each instance the owner-agent is able to obtain higher selling prices relative to other sales in the market. Thus, after controlling for the probability that some owner-agents are investors, the results further corroborate the evidence for the existence of a principal-agent problem.

In the second and fourth models in Table 3, we include the *Uniqueness* measure and the *Market Competition* variable. Given our evidence that an agency problem exists, an interesting related question is whether the degree of the agency problem varies with the degree of competition among agents. We find that the degree of market competitiveness, defined as the number of active listing agents in a MLS-defined area divided by the total number of listings in the same MLS-defined area multiplied by 100, is associated with a small, but statistically significant, positive impact (0.10%) on the selling price. The coefficient on *Owner-agent* in the selling price model remains significant and does not change after controlling for market competition among agents or uniqueness of the house. This result suggests that the level of competition among agents has a minimal impact on the agency problem between agents and principals as described in this research.

5.2. Selling price results corrected for possible endogeneity of Owner-agent

The prior results are based on selling price equations with a correction for sample selection bias. However, they do not address the question of endogeneity of *Owner-agent*. To control for endogeneity, we estimate another probit model and a linear probability model with *Owner-agent* as the dependent variable. The results for both the probit model and the linear probability model are presented in Table 4. We then estimate two sets of selling price regression models to account for possible endogeneity. In the first model in Table 5, we obtain the inverse Mill's ratio (IMR_2) from the owner-agent probit model and include it in a regression model to account for the possible endogeneity of owner-agent. The coefficient on *Owner-agent* increases to 6.5% as shown in Table 5 versus 4.5% in the model corrected for sample

[4] In each of the selling price models, we find evidence of sample selection bias with the coefficient for the *IMR* being significant at 1% in all models estimated. We also ran a selling price model without correcting for sample selection bias and the coefficient on the owner-agent variables was 4.4% and significant at the 1% level. The correction for sample selection does not appear to affect the magnitude of the coefficient for *Owner-agent*.

Table 5
Regression models of house prices corrected for sample selection bias and possible endogeneity of Owner-agent using the results from the two probit models in Table 4 to estimate the inverse Mills ratios (IMRs). We then include both IMRs in the selling price model. The probit for sold properties and the resulting IMR_1 is based on Model 1 in Table 4. The probit for Owner-agent and the resulting IMR_2, plus the regression model for selling price, are based on the sample of 193,007 houses sold during 1999–2002.

The dependent variable is the log of the selling price. All regressions include quarterly dummy variables (not reported for brevity) to control for potential serial effects and all regressions include dummy variables for MLS specified areas (not reported for brevity) to control for location. The estimates of the coefficients are presented in the table, with *t*-statistics reported in parentheses using heteroskedasticity-robust Huebner/White standard errors. Statistics with significance at the 1% level are denoted with a ** and at the 5% level are denoted with a *.

Independent variable	Model 1	Model 2	Model 3	Model 4
Constant	10.98	10.96	10.98	10.96
	(5069.84)**	(2249.14)**	(5069.78)**	(2249.14)**
Size	0.04	0.04	0.04	0.04
	(426.93)**	(407.21)**	(427.01)**	(407.16)**
Age	−0.05	−0.05	−0.05	−0.05
	−(150.88)**	−(144.48)**	−(150.86)**	−(144.46)**
Owner-agent	0.07	0.07		
	(20.42)**	(20.69)**		
Owner-agent (1–2 listings)			0.06	0.06
			(16.70)**	(16.78)**
Owner-agent (3–5 listings)			0.08	0.08
			(14.79)**	(14.76)**
Owner-agent (>5 listings)			0.06	0.06
			(13.26)**	(14.29)**
Pool	0.07	0.08	0.07	0.08
	(74.05)**	(80.32)**	(74.07)**	(80.34)**
Fireplace	0.07	0.07	0.07	0.07
	(75.73)**	(74.04)**	(75.77)**	(74.08)**
Large lot	0.09	0.10	0.09	0.10
	(70.17)**	(72.43)**	(70.15)**	(72.42)**
Tenant	−0.06	−0.06	−0.06	−0.06
	−(32.38)**	−(30.91)**	−(32.43)**	−(30.95)**
Vacant	−0.07	−0.07	−0.07	−0.07
	−(83.38)**	−(79.53)**	−(83.46)**	−(79.65)**
Builder	0.06	0.05	0.06	0.05
	(38.11)**	(34.19)**	(38.10)**	(34.16)**
Uniqueness		−0.09		−0.09
		−(53.65)**		−(53.63)**
Market competition		0.00		0.00
		(6.88)**		(6.91)**
Number of houses sold	193,007	193,007	193,007	193,007
Adjusted R^2	0.93	0.93	0.93	0.93
F	17,299	18,184	17,025	17,899
Inverse Mills ratio (IMR_1) from a probit model for sold versus not sold to correct for possible Sample selection	−0.13	−0.10	−0.12	−0.10
	−(27.89)**	−(22.82)**	−(27.84)**	−(22.77)**
IMR_2 from a probit model for agent-owned Houses versus client-owned houses to correct for possible endogeneity of Owner-agent	−0.01	−0.01	−0.01	−0.01
	−(5.09)**	−(5.51)**	−(4.46)**	−(4.90)**

selection bias but not endogeneity. Using the correction for endogeneity, we find a significant increase in the values of the owner-agent coefficients. All coefficients for the owner-agent variables increase with a maximum increase for *Owner-agent* (3–5 *listings*) of 3.5% in Model 3. We also estimate models based on the three alternatives suggested previously and find results consistent with models in which the probit model generates the inverse Mills ratio to be included in the selling price equation. In all cases the results indicate that after correcting for endogeneity, owner-agents obtain an even higher price for their houses relative to non-owner-agent houses (see Appendix B).

The results of the various models are robust and suggest that after correction for sample selection bias and endogeneity of *Owner-agent*, owner-agents are able to sell their properties for a higher price than similar non-agent-owned properties. The evidence indicates that once you account for physical characteristics, location, marketing, probability of sale, and possible endogeneity of owner-agent effort, owner-agents are able to obtain higher prices.

5.3. Are owner-agents simply better agents?

In this section we examine the possibility that the estimated price premium for agent-owned listings is a function of owner-agents being superior agents who are able to sell all properties at a higher price without lengthening the marketing time. To test this hypothesis, we first delete all the agent-owned properties and then run a list price model, a time-on-the-market model, and a selling price model corrected for sample selection bias using only the client-owned properties. We find that houses listed for clients by agents who were also owner-agents list at a statistically significant 0.17% price premium compared to houses listed by non-owner agents and sold at a statistically significant 0.25% price premium with a statistically significant reduction of approximately 1.2% on the marketing time. Therefore, the difference in the abilities of the owner-agents compared to other agents may account for part (0.25%) of the selling price differential, but leaves most of the estimated price premium unexplained. The results for the list price and time-on-the-market models, excluding all owner-agent listings, are based on a sample of 296,958 listings with 57% of the sample listed by agents that also were owner-agents during the sample period. The results for the selling price model are based on a sample of 187,393 listings that sold. The full results for the three models are not included in paper, but are available upon request.

The conclusion from the results presented in this section is that a statistically significant difference exists between the selling prices of agent-owned properties and client-owned properties. While there are marginal differences in the magnitude of the *Owner-agent* coefficient for the various model specifications, the coefficient across samples in the selling price equations is consistently significant and positive with a range of 4.5–6.6%. This is a substantial result, even in consideration of the 0.25% estimated premium associated with client-owned houses that were sold by listing agents that were also owner-agents.

5.4. Owner-agents and initial list price

The relation between agent-owned listings and list prices is evaluated by estimating a list price model using an OLS regression. Using the log of the list price as the dependent variable and the independent variables for the time-on-the-market model, the estimated list price models have R^2's of 0.93 (for the full sample, $n = 306,869$ and the sample of sales only, $n = 193,007$). The signs and magnitudes of the coefficients are similar to those of the selling price model for both list price models. For the full sample of listings, the results indicate that agent-owners list their houses at prices 3.5% higher than houses listed by non-agent owners. If we restrict the sample to only the properties that eventually sold, the coefficient on the owner-agent variable in the list price model is 4.1%. This evidence of initial higher pricing suggests that real estate agents that own their properties expect to benefit from setting higher list prices and lends support to the position that agent-owned houses are priced differently than their clients' houses. The list price model results are not included in the paper, but are available from the authors upon request.

It is possible that owner-agents list all their listings at higher prices, both their own and those of their clients. To test this, a list price model is estimated using only the sample of houses listed by the agents that were also owner-agents during the sample period ($n = 179,078$). The magnitude of the *Owner-agent* coefficient for this sample is 3.4% and remains significant at the 1% level. If we estimate the same list price model with only the listings that eventually sold ($n = 114,874$), the coefficient is 3.95%. The combined results from these restricted samples provide strong evidence that owner-agents set the initial price for their own houses higher than the initial price of their clients' houses.

5.5. Time-on-the-market effects

The results for the time-on-the-market models are shown in Table 6. The initial finding is that owner-agents take no more time to sell their properties. Thus, these results provide no evidence that agents are keeping their houses on the market longer in order to obtain the higher prices that the results indicate. In the first model, the coefficient on *Owner-agent* is not significant, suggesting no variation in marketing time when the listing agent is also the owner of the property. The results for the other variables included in the time-on-the-market model are fairly typical. Time on the market is found to increase with house size, overpricing, open houses, whether the agent is inexperienced, whether the agent is a broker, and whether the house is on a large lot, is vacant, is occupied by a tenant, or is builder-owned. In addition, time on the market decreases marginally for older properties, properties with a pool, houses with more fireplaces, increasing interest rates, and listing agent experience.

In the second model in Table 6 we estimate the time-on-the-market model for the complete sample and include the additional variables *Market Competition, Degree of Overpricing,* and *Uniqueness*. The coefficient for *Owner-agent* is again insignificant. In the third model, we separate the owner-agent variable into three groups indicating the number of listings by an individual owner-agent over the study period to further

Table 6
Parametric estimation of accelerated failure time Weibull duration models corrected for censoring based on the full sample during 1999–2002, where 193,007 of 306,869 houses listed eventually sold.

The dependent variable is log of time on the market (DOM). All regressions include quarterly dummy variables (not reported for brevity) to control for potential serial effects and all regressions include dummy variables for MLS specified areas (not reported for brevity) to control for location. The ML estimates of the coefficients are presented in the table, with *t*-statistics reported in parentheses using heteroskedasticity-robust Huebner/White standard errors. Statistics with significance at the 1% level are denoted with a** and at the 5% level are denoted with a*.

Independent variable	Model 1	Model 2	Model 3	Model 4
Constant	1.47	1.43	1.47	1.43
	(475.02)**	(218.93)**	(475.06)**	(218.92)**
Size	0.01	0.01	0.01	0.01
	(119.42)**	(109.16)**	(119.40)**	(109.15)**
Age	0.00	0.00	0.00	0.00
	(2.81)**	(4.13)**	(2.76)**	(4.08)**
Owner-agent	0.00	0.00		
	(0.65)	(0.92)		
Owner-agent (1–2 listings)			0.00	0.00
			(0.52)	(0.81)
Owner-agent (3–5 listings)			0.02	0.03
			(4.21)**	(4.32)**
Owner-agent (>5 listings)			−0.01	−0.01
			−(3.50)**	−(3.51)**
Pool	−0.02	−0.02	−0.02	−0.02
	−(21.54)**	−(19.84)**	−(21.56)**	−(19.86)**
Fireplace	−0.02	−0.02	−0.02	−0.02
	−(21.16)**	−(20.74)**	−(21.18)**	−(20.76)**
Large lot	0.05	0.05	0.05	0.05
	(35.22)**	(36.04)**	(35.20)**	(36.02)**
Tenant	0.05	0.05	0.05	0.05
	(19.22)**	(19.21)**	(19.18)**	(19.17)**
Vacant	0.01	0.01	0.01	0.01
	(13.10)**	(13.27)**	(13.18)**	(13.35)**
Builder	0.04	0.04	0.04	0.04
	(25.58)**	(24.58)**	(25.64)**	(24.63)**
Uniqueness		−0.02		−0.02
		−(7.73)**		−(7.77)**
Market competition		0.00		0.00
		(6.02)**		(6.02)**
Degree of overpricing (DOP)		0.00		0.00
		(37.44)**		(37.44)**
Interest rate change	−0.10	−0.10	−0.10	−0.10
	−(65.99)**	−(66.03)**	−(66.02)**	−(66.06)**
Open house	0.03	0.03	0.03	0.03
	(12.69)**	(11.45)**	(12.67)**	(11.43)**
Tour house	−0.01	−0.01	−0.01	−0.01
	−(2.90)**	−(6.46)**	−(2.90)**	−(6.47)**
Listing agent limited experience	0.01	0.02	0.01	0.02
	(8.55)**	(8.66)**	(8.54)**	(8.65)**
Listing agent experienced	−0.02	−0.02	−0.02	−0.02
	−(16.14)**	−(16.21)**	−(16.13)**	−(16.19)**

Table 6 (*continued*)

Independent variable	Model 1	Model 2	Model 3	Model 4
Listing agent broker	0.01	0.01	0.01	0.01
	(14.38)**	(13.04)**	(14.45)**	(13.12)**
Number of observations	306,869	306,869	306,869	306,869
Number of sold obs.	193,007	193,007	193,007	193,007
Number of unsold obs.	113,862	113,862	113,862	113,862
Log likelihood	−16,306	−15,304	−16,290	−15,286
Wald chi^2 (model)	66,032	67,173	66,080	67,222
P	7.00	7.01	7.00	7.01
AIC	32,869	30,870	32,840	30,839

examine the possibility of owner-agents being investors in the market. Here the results suggest that owner-agents with one or two listings did not keep their properties on the market any longer than listings by non-owner agents. Owner-agents with three to five listings stayed on the market longer by about 2.5%, and owner-agents with more than five listings were able to reduce the time on the market by approximately 1.3%. Though statistically significant, the magnitude of these results suggests a marginal impact. If we use the full sample average of 95 days on the market, a 2.5% increase is only 2.4 days and a 1.3% decrease is only 1.2 days.

The conclusion from these models is that the impact of *Owner-agent* on the time on the market is either statistically insignificant or negligible. Owner-agents do not appear to sell their houses more quickly than other houses, nor do they keep them on the market longer in an attempt to obtain a higher price.

5.6. Restricting the sample

Our objective has been to determine whether owner-agents are able to obtain higher prices relative to their clients. Thus far, the results suggest they set their prices higher and sell their own properties at higher prices relative to the houses of their clients. In this section, we attempt to examine more closely the impact of *Owner-agent* on both time on the market and selling price by creating a restricted sample of agent-owned properties and other properties and then re-estimating the previously described models to correct for sample selection bias and endogeneity based on a restricted sample. The restricted sample is obtained using the following joint constraints. The first constraint restricts the sample to only those properties listed by an agent that is also an owner-agent during the sample period, creating a sample such that agency issues should be more easily identified if they exist. The second constraint restricts the sample to only those properties listed in the same subdivisions as the agent-owned listings creating a matched sample with more similar location and physical characteristics. For the subdivision constraint, the matching is initially done by computer, then followed with a round of hand matching due to significant variation in subdivision names. While the matching process is not as clean as we would have preferred, we believe that our final matched sample is representative of the subdivisions. By restricting the sample, we now compare only listings by agents that were also owner-agents and listings that are in the same subdivision as owner-agent listings. We believe that this will allow us to determine more precisely whether an agency problem exists since we are now comparing only client listings that are listed by the set of owner-agents.

5.7. Results based on the restricted sample

In the restricted matched subset of the data as defined above, we include only listings by agents that are also owner-agents during the sample period and only listings that occur in subdivisions in which an owner-agent has a listing. The objective is to assess coefficient sensitivity and to directly compare owner-agent listings to the other listings by agents that are also owner-agents. In this final set of selling price regressions using the restricted sample, the coefficients for the owner-agent variables are similar to the initial results using the full sample. The coefficient for *Owner-agent* with a correction for sample selection bias is reported in Table 7 and remains 4.5% as in the initial models with all other owner-agent variations within 0.30% of the coefficients in the full sample. When we also correct for endogeneity, the results, found in Table 9, are similar to the results for the full model with a correction for endogeneity, indicating that owner-agents are able to sell their properties for approximately 6% more than the properties of their clients. The coefficients in the restricted sample are all within 0.90% of the coefficients in the full sample. These findings using the restricted sample also provide evidence that reduces the likelihood that selection bias explains the results (Table 8 presents the probit models for the restricted sample).

The owner-agent coefficients for the time-on-the-market model, presented in Table 10, are all significant for the restricted sample. When comparing owner-agent listings to other listings by these same agents in the matching subdivisions, we find that *Owner-agent* is now significant, but the actual time difference is less than one day in the first two models. The last two models, models 3 and 4 with owner-agent separated into groups, have results similar to the full sample, with variations of less than 2.5 days relative to non-agent-owned properties.

5.8. Implications of the findings

The positive and statistically significant coefficient on *Owner-agent* in the selling price models provides empirical evidence that real estate agents sell their own houses for higher prices compared to other houses they list. The evidence of a price premium suggests there is a divergence of interests between principals and agents. It also suggests that agents can in fact have a significant impact on the selling price.

One possible explanation of these results is that an owner-agent devotes more resources to selling his own property, in which case the higher price might simply reflect a normal return to the additional investment in selling. This is a difficult proposition to test because the additional resources by the agent in selling his own

Table 7

Regression models of house prices corrected for possible sample selection bias based on the probability of sale using Heckman's selection model, based on a reduced sample during 1999–2002 of 89,092 houses listed on the multiple listing service (MLS) in several Texas counties, of which 58,096 are sold and 30,996 are withdrawn or not sold by the end of the sample period. To be included in this sample, a listing had to be by an agent who was also an owner-agent and the listings had to occur in a subdivision in which an owner-agent had a listing. This results in a more closely matched sample and provides a comparison of the owner-agent's effort relative to other properties in similar areas that were also listed by agents that were owner-agents.

The dependent variable is log of the selling price. All models include quarterly dummy variables (not reported for brevity) to control for potential serial effects and all regressions include dummy variables for MLS specified areas (not reported for brevity) to control for location. The ML estimates of the coefficients are presented in the table, with *t*-statistics reported in parentheses using heteroskedasticity-robust Huebner/White standard errors. Statistics with significance at the 1% level are denoted with a** and at the 5% level are denoted with a*.

Independent variable	Model 1	Model 2	Model 3	Model 4
Constant	10.97	10.91	10.97	10.91
	(2619.08)**	(1261.74)**	(2618.84)**	(1261.57)**
Size	0.041	0.042	0.041	0.042
	(318.59)**	(326.46)**	(318.78)**	(326.39)**
Age	−0.049	−0.047	−0.049	−0.047
	−(72.36)**	−(69.98)**	−(72.30)**	−(69.92)**
Owner-agent	0.045	0.045		
	(23.35)**	(23.72)**		
Owner-agent (1–2 listings)			0.039	0.038
			(16.51)**	(16.60)**
Owner-agent (3–5 listings)			0.048	0.049
			(9.64)**	(10.04)**
Owner-agent (>5 listings)			0.06	0.061
			(16.10)**	(16.51)**
Pool	0.07	0.08	0.07	0.08
	(41.88)**	(45.15)**	(41.97)**	(45.23)**
Fireplace	0.06	0.06	0.06	0.06
	(39.63)**	(39.50)**	(39.71)**	(39.57)**
Large lot	0.07	0.07	0.07	0.07
	(25.98)**	(28.41)**	(26.01)**	(28.44)**
Tenant	−0.08	−0.07	−0.08	−0.07
	−(23.36)**	−(22.11)**	−(23.36)**	−(22.09)**
Vacant	−0.07	−0.07	−0.07	−0.07
	−(48.67)**	−(46.94)**	−(48.94)**	−(47.24)**
Builder	0.07	0.06	0.07	0.06
	(22.78)**	(21.30)**	(22.61)**	(21.11)**
Uniqueness		−0.08		−0.08
		−(25.81)**		−(25.78)**
Market competition		0.002		0.002
		(8.98)**		(9.04)**
Number of observations	89,092	89,092	89,092	89,092
Number of sold obs.	58,096	58,096	58,096	58,096
Number of unsold obs.	30,996	30,996	30,996	30,996
Log likelihood	−16,747	−16,292	−16,620	−16,163
Wald chi^2 (model)	573,505	620,027	576,865	623,143

Table 7 (*continued*)

Independent variable	Model 1	Model 2	Model 3	Model 4
Inverse Mills ratio from a probit model for sold versus not sold to correct for possible Sample selection	−0.023 −(7.79)**	−0.016 −(6.11)**	−0.023 −(7.81)**	−0.016 −(6.16)**

house are not observed in the data. However, in the selling price models that correct for endogeneity and include additional variables indicative of additional observable effort by the agent (*open house* and *tour house*), the price premium to the owner-agent increases.

Another possible explanation is that owner-agents initially buy higher quality properties and are thus able to obtain higher prices in the market. If true, this could explain the 4.5% premium. Obtaining a quality measure for the full sample is not possible, but we are able to obtain a quality measure for a small sample of 13,153 sales in one county for 2002. We use the indicated quality of the property as determined by the county appraisal board for tax purposes. Eight groups are identified in this sample; poor, fair to average, average, average to above average, good, high, excellent, and excellent to superior quality. Once we control for quality using this sample of 13,153 sales in a regression model that contained all the control variables for the sale price models for the full sample, the estimated coefficient for *Owner-agent* is 4.36%. This result, and the finding that the empirical results do not change when we restrict the sample to properties in similar subdivisions, indicates that quality is not likely to provide a rationale for the premium. Thus, we suggest that the price premium documented in this paper is evidence for the standard argument in the principal-agent theory that the percentage commission structure, or any compensation structure such that the agent receives less than full amount of the surplus, results in positive externalities and fails to align the interests of the principal with those of the agent.

The 4.1% estimated list price premium (see footnote 13) for owner-agents for houses that sold and the marginally larger (4.5%) premium on the selling price are of interest. This provides evidence of higher list prices leading to higher selling prices for houses owned by agents. It also suggest that owner-agents may not discount their properties. We define the percentage discount as $[(1 - \text{selling price}/\text{list price}) * 100]$ for the sold properties and find that the average discount for owner-agents is 1.59% versus 1.80% for client-owned properties. This is similar to the discount obtained when estimating a regression using only the sold properties, where the percentage discount is the dependent variable and we include the independent variables from the selling price model on the right-hand side of the equation. The regression coefficient for *Owner-agent* in this regression model is −0.388, indicating that agent-owned properties have lower discounts of approximately one-third of a percent less than client-owned properties.

The result that owner-agents offer lower concessions/discounts from the list price for their own properties relative to their clients' properties implies that agents obtain

Table 8
The results below are for three binary dependent variable models. The first model is a probit model, where the dependent variable equals one for a sold house and zero for an unsold house. From this probit model the inverse Mills ratio is calculated and included in the selling price models in Table 7 to correct for possible sample selection bias of sold versus unsold properties. In the second model, the results are for a probit model, where the dependent variable equals one if the owner of the house is also a real estate agent and zero for houses not owned by a real estate agent. The inverse Mills ratio calculated using the results from this model is included in the selling price models in Table 5 to correct for the possible endogeneity of Owner-agent. In addition, the predicted probability from this model is used in place of Owner-agent in Table 11 in Appendix B as an alternative means of correcting for possible endogeneity of Owner-agent.

Model three is a linear probability model for Owner-agent, where the binary dependent variable is one if the owner of the house is a real estate agent. The use of a linear probability model and the predicted probability from this model in place of Owner-agent has been suggested as a means of correcting for endogeneity and has been implemented in Table 11 in Appendix B. Also, one other approach suggested in the literature to correct for endogeneity is to include the residual from the linear probability model and Owner-agent variable in a selling price model. This is also implemented, as shown in Table 11 of Appendix B.

The dependent variable in all three models is a dummy variable as described above. All models include quarterly dummy variables (not reported for brevity) to control for potential serial effects and all regressions include dummy variables for MLS specified areas (not reported for brevity) to control for location. The estimates of the coefficients are presented in the table, with *t*-statistics reported in parentheses using heteroskedasticity-robust Huebner/White standard errors. To be included in this sample a listing had to be by an agent that was also an owner-agent and the listings had to occur in a subdivision in which an owner-agent had a listing. This results in a more closely matched sample and provides a comparison of the owner-agent's effort relative to other properties in similar areas that were also listed by agents that were owner-agents. Statistics with significance at the 1% level are denoted with a ** and at the 5% level are denoted with a *.

Independent variable	Model 1: probit model for sold versus not sold	Model 2: probit model for owner-agent	Model 3: linear probability model for owner-agent
Constant	1.27	−7.29	−0.64
	(15.93)**	−(31.05)**	−(36.35)**
Size	−0.04	−1.00	−0.13
	−(47.95)**	−(58.50)**	−(115.45)**
Age	−0.00	−0.10	−0.01
	−(0.76)	−(6.87)**	−(9.47)**
Owner-agent	−0.14		
	−(9.73)**		
Large lot	−0.16	−4.37	−0.56
	−(9.79)**	−(42.43)**	−(89.34)**
Tenant	−0.27	−6.64	−0.84
	−(10.05)**	−(41.18)**	−(89.93)**
Vacant	0.05	1.44	0.19
	(4.20)**	(41.42)**	(71.59)**
Builder	0.13	3.13	0.41
	(6.46)**	(39.17)**	(69.38)**
Uniqueness	0.25	2.93	0.52
	(10.49)**	(32.08)**	(64.53)**
Market competition	0.00	0.10	0.01
	(1.48)	(20.60)**	32.25**
Degree of overpricing (DOP)		−0.00	−0.00
		−(2.26)*	−(2.06)*

Table 8 (*continued*)

Independent variable	Model 1: probit model for sold versus not sold	Model 2: probit model for owner-agent	Model 3: linear probability model for owner-agent
Interest rate change	0.03 (1.54)		
Open house	−0.17 −(5.76)**	−4.56 −(31.59)**	−0.56 −(65.05)**
Tour house	0.22 (9.79)**	5.33 (45.47)**	0.71 (91.29)**
Listing agent limited experience	−0.08 −(3.08)**	−1.91 −(23.54)**	−0.24 −(40.25)**
Listing agent experienced	0.16 (10.22)**	3.66 (48.07)**	0.49 (89.72)**
Listing agent broker	−0.09 −(9.41)**	−2.19 −(42.22)**	−0.29 −(72.74)**
Internet		−0.37 −(8.40)**	−0.05 −(11.25)**
Large firm		−0.29 −(8.08)**	−0.03 −(9.33)**
Residual from a time-on-the-market Log linear model	−0.00 −(11.73)**	−0.02 −(50.68)**	−0.00 −(93.00)**
Number of houses sold	58,096	58,096	58,096
Log likelihood	−53,132	−5,674	
Wald chi^2 (model)	8,175	4,658	
Pseudo R^2	0.08	0.69	
F			136.46
Adjusted R^2			0.51

a higher price for their properties not only by negotiating more fervently for a higher percentage of the list price, but also by listing them at a higher price and exerting the additional effort, but no additional time, to find a buyer who is willing to pay a higher price. Does this mean that a seller can correct the agency problem by increasing his agent's recommended list price by approximately 4.0%? Unfortunately, the answer is no. Although such a strategy may correct the agency problem with respect to the price, it will cause further agency problems with respect to the time on the market. As can be seen from the theoretical model in Appendix A, the agent's choice of effort level, hence time on the market, is a function of the price. Furthermore, for any given price, the listing agent expands a greater effort level if he owns the property than if he does not. Thus, if the seller raises the asking price to the level of the asking price for a comparable agent-owned property, he will experience a longer marketing time relative to an agent-owned property. This is also reflected in the positive coefficient for *Degree of Overpricing* in the time-on-the-market

Table 9

Regression models of house prices corrected for sample selection bias and possible endogeneity of Owner-agent using the results from the two probit models in Table 8 to estimate the inverse Mills ratios (IMRs). We then include both IMRs in the selling price model. The probit for sold properties and the resulting IMR_1 is based on Model 1 in Table 8. The probit for Owner-agent and the resulting IMR_2, plus the regression model for selling price, are based on the sample of 58,096 houses sold during 1999–2002. To be included in this sample, a listing had to be by an agent that was also an owner-agent and the listings had to occur in a subdivision in which an owner-agent had a listing. This results in a more closely matched sample and provides a comparison of the owner-agent's effort relative to other properties in similar areas that were also listed by agents who were owner-agents.

The dependent variable is the log of the selling price. All regressions include quarterly dummy variables (not reported for brevity) to control for potential serial effects and all regressions include dummy variables for MLS specified areas (not reported for brevity) to control for location. The estimates of the coefficients are presented in the table, with *t*-statistics reported in parentheses using heteroskedasticity-robust Huebner/White standard errors. Statistics with significance at the 1% level are denoted with a** and at the 5% level are denoted with a*.

Independent variable	Model 1	Model 2	Model 3	Model 4
Constant	10.97	10.91	10.97	10.91
	(2633.69)**	(1263.12)**	(2633.59)**	(1263.30)**
Size	0.04	0.04	0.04	0.04
	(215.06)**	(202.44)**	(215.64)**	(202.78)**
Age	−0.05	−0.05	−0.05	−0.05
	−(73.41)**	−(70.28)**	−(73.4)**	−(70.25)**
Owner-agent	0.06	0.06		
	(23.97)**	(23.21)**		
Owner-agent (1–2 listings)			0.06	0.05
			(18.86)**	(17.97)**
Owner-agent (3–5 listings)			0.08	0.08
			(14.31)**	(13.52)**
Owner-agent (>5 listings)			0.05	0.06
			(13.56)**	(14.46)**
Pool	0.07	0.08	0.07	0.08
	(41.94)**	(45.00)**	(42.01)**	(45.06)**
Fireplace	0.06	0.06	0.06	0.06
	(39.66)**	(39.42)**	(39.74)**	(39.50)**
Large lot	0.08	0.08	0.08	0.08
	(28.62)**	(30.15)**	(28.62)**	(30.18)**
Tenant	−0.06	−0.06	−0.06	−0.06
	−(18.13)**	−(17.52)**	−(18.21)**	−(17.57)**
Vacant	−0.07	−0.07	−0.07	−0.07
	−(50.23)**	−(47.97)**	−(50.26)**	−(48.09)**
Builder	0.06	0.06	0.06	0.06
	(19.97)**	(18.75)**	(19.99)**	(18.72)**
Uniqueness		−0.09		−0.09
		−(27.38)**		−(27.38)**
Market competition		0.00		0.00
		(8.31)**		(8.36)**
Number of houses sold	58,096	58,096	58,096	58,096
Adjusted R^2	0.94	0.94	0.94	0.94
F	5,643	5,965	5,558	5,876
	−0.13	−0.10	−0.13	−0.10

Table 9 (*continued*)

Independent variable	Model 1	Model 2	Model 3	Model 4
Inverse Mills ratio (IMR_1) from a probit model for sold versus not sold to correct for possible sample selection	−(14.62)**	−(11.09)**	−(14.64)**	−(11.14)**
IMR_2 from a probit model for agent-owned houses versus client-owned houses to correct for possible endogeneity of Owner-agent	−0.01 −(4.36)**	−0.01 −(4.30)**	−0.01 −(3.83)**	−0.01 −(3.79)**

Table 10

Parametric estimation of accelerated failure time Weibull duration models corrected for censoring based on a reduced sample during 1999–2002, where 58,096 of 89,092 houses listed by agents who were also owner-agents eventually sold. To be included in this sample, a listing had to be by an agent who was also an owner-agent and the listings had to occur in a subdivision in which an owner-agent had a listing. This results in a more closely matched sample and provides a comparison of the owner-agent's effort relative to other properties in similar areas that were also listed by agents that were owner-agents.

The dependent variable is the log of time on the market (DOM). All regressions include quarterly dummy variables (not reported for brevity) to control for potential serial effects and all regressions include dummy variables for MLS specified areas (not reported for brevity) to control for location. The ML estimates of the coefficients are presented in the table, with t-statistics reported in parentheses using heteroskedasticity-robust Huebner/White standard errors. Statistics with significance at the 1% level are denoted with a ** and at the 5% level are denoted with a *.

Independent variable	Model 1	Model 2	Model 3	Model 4
Constant	1.45 (238.81)**	1.41 (115.98)**	1.45 (238.91)**	1.41 (115.97)**
Size	0.01 (66.25)**	0.01 (60.33)**	0.01 (66.21)**	0.01 (60.32)**
Age	0.00 (1.28)	0.00 (2.13)*	0.00 (1.18)	0.00 (2.04)*
Owner-agent	0.01 (2.63)**	0.01 (2.82)**		
Owner-agent (1–2 listings)			0.01 (2.27)*	0.01 (2.49)*
Owner-agent (3–5 listings)			0.03 (4.79)**	0.03 (4.90)**
Owner-agent (>5 listings)			−0.01 −(2.44)*	−0.01 −(2.51)*
Pool	−0.02 −(11.31)**	−0.02 −(10.53)**	−0.02 −(11.34)**	−0.02 −(10.55)**
Fireplace	−0.02 −(11.54)**	−0.02 −(11.19)**	−0.02 −(11.56)**	−0.02 −(11.21)**
Large lot	0.04 (13.43)**	0.04 (14.20)**	0.04 (13.35)**	0.04 (14.13)**
Tenant	0.05	0.05	0.05	0.05

Table 10 (*continued*)

Independent variable	Model 1	Model 2	Model 3	Model 4
	(11.51)**	(11.51)**	(11.44)**	(11.44)**
Vacant	0.02	0.02	0.02	0.02
	(9.14)**	(9.33)**	(9.29)**	(9.48)**
Builder	0.04	0.03	0.04	0.03
	(12.06)**	(11.51)**	(12.20)**	(11.65)**
Uniqueness		−0.01		−0.01
		−(3.87)**		−(3.93)**
Market competition		0.00		0.00
		(4.19)**		(4.18)**
Degree of overpricing (DOP)		0.00		0.00
		(18.98)**		(19.00)**
Interest rate change	−0.09	−0.09	−0.09	−0.09
	−(35.79)**	−(35.76)**	−(35.86)**	−(35.83)**
Open house	0.02	0.02	0.02	0.02
	(5.36)**	(4.80)**	(5.32)**	(4.76)**
Tour house	−0.01	−0.01	−0.01	−0.01
	−(1.68)	−(3.59)**	−(1.70)	−(3.62)**
Listing agent limited experience	0.02	0.02	0.02	0.02
	(4.70)**	(4.87)**	(4.67)**	(4.84)**
Listing agent experienced	−0.02	−0.02	−0.02	−0.02
	−(6.74)**	−(6.74)**	−(6.70)**	−(6.69)**
Listing agent broker	0.02	0.02	0.02	0.02
	(10.54)**	(9.94)**	(10.68)**	(10.09)**
Number of observations	89,092	89,092	89,092	89,092
Number of sold obs.	58,096	58,096	58,096	58,096
Number of unsold obs.	30,996	30,996	30,996	30,996
Log likelihood	−2,283	−2,029	−2,266	−2,011
Wald chi^2 (model)	18,825	19,032	18,874	19,084
p	7.04	7.05	7.04	7.05
AIC	4,822	4,320	4,793	4,288

equations; a higher list price typically increases the time on the market and thus decreases the probability of a sale for a given effort level by the agent.

6. Conclusions

This paper addresses the principal-agent relationship in markets for real assets. We examine whether or not agents act in the best interests of their clients by ascertaining whether they sell their clients' assets as favorably as they do their own assets. A simple theoretical model of the agent's problem is developed and empirically evaluated using a comprehensive data set from several metropolitan counties in Texas. The data set enables us to test directly for the presence or absence of an agency problem because a subset of the assets in the data were owned by real estate agents.

The empirical results show that agent-owned houses sell no faster than non-agent-owned houses, but they do sell at a premium of approximately 4.5% above the price of a house not owned by a real estate agent. This result is supported across multiple specifications of the models and implementation of statistical corrections for self selection bias and endogeneity. The empirical evidence from this paper supports the standard agency models in general, indicating that the percentage commission system creates an agency problem between an agent and his client because it induces the agent to expend too little effort for his client.

Appendix A. A simple model of agent effort and price choice

In this appendix, we offer a simple search model of a risk-neutral listing agent who chooses both an effort level and a price for a property. In order to study whether the agent represents the interest of his client in the same way that he would represent himself, we first study the case in which the agent is the owner of the property and then compare this to the case in which the agent is not the owner of the property. The purpose here is not to design the efficient and incentive compatible contract between the seller and the listing agent. Rather, we take the current commission structure in the industry as given and compare the effort level and price choices of the agent depending on whether he is the owner of the listing or not.

To keep the model simple, we assume that the seller of the property sets the asking price suggested by his listing agent. This assumption can be motivated by the fact that agents have superior information about the market and that sellers consult their listing agents in setting their asking prices. Obviously, it is unrealistic to assume that the seller relies completely on the listing agent to set the asking price. This assumption simplifies the model without changing the qualitative results. All we need for the results to hold is that the agent has *some* impact on the seller's choice of the asking price. The degree of this impact will determine the level, but not the direction, of the divergence between the optimal asking price for the seller and the optimal asking price for the agent. To further simplify the analysis, we adopt a simple bargaining model for the negotiation stage of the game; the seller's asking price is treated as a take-it-or-leave-it offer to the buyer and the buyer can either accept or reject the seller's price. Alternative bargaining solutions, such as Rubinstein's (1982) alternating offers bargaining and Nash bargaining, introduce the possibility of a strategic role for the listing price in the negotiation stage, where the listing price can serve as a signal for the seller's reservation price. These are interesting issues and have been examined in Yavas and Yang (1995). As will be seen in the model, the choice of the bargaining solution is not critical for our analysis.

A buyer will accept an asking price only if the asking price is below his reservation price. The density function of buyers' reservation prices is given by $f(.)$ over the interval $[\underline{P}, \bar{P}]$, where f is continuous everywhere.

The listing agent is assumed to be a member of the local multiple listing service (MLS). The MLS requires the listing agent to report his listing to the MLS. Along with the information about the property, the listing agent also indicates the percentage of the price that he will pay as commission if another MLS agent finds the buyer. The MLS then disseminates the listing to all other members of the MLS. There is a large number of agents who are members of the MLS. The member agent that finds the buyer for the listing is referred to as the selling agent.

Following the current practice in the industry, the listing agent receives k percentage of the price as commission from the seller upon the sale of the property. Out of this commission, the listing agent pays k_s, $k_s < k$, percentage of the price to the selling agent. If the listing agent finds the buyer himself, he gets to keep the entire commission. Typically, an agent also shares his commission with his brokerage company in return for the office space and secretarial support provided by the company. An exception is the RE/MAX brokerage firms, where the agent retains 100% of the commission and pays his firm a fixed amount each period for the office space and secretarial support. Such differences are inconsequential for our analysis. We assume that the total commission rate, k, and the selling agent's share, k_s, are determined in the market and taken as exogenous by the players. As indicated earlier, our purpose is to compare the effort level and price choices of a broker for a property that he owns with those of a similar property that he does not own, taking the current commission structure in the industry as given.

We next study the optimal asking price and effort choices of a listing agent for his own property. We then compare these choices to the case in which the listing agent is not the owner of the property.

Listing agent as the owner: The problem of a listing agent when listing his own property is to choose a price, P^{O}, and a search effort level, L^{O}, to maximize his expected profits:

$$\max_{P^{O},L^{O}} \prod\nolimits^{O}(P^{O}, L^{O}) = \psi(L^{O}) \int_{P^{O}}^{\bar{P}} (P^{O} - R) f(p)\, \mathrm{d}p + \Phi \int_{P^{O}}^{\bar{P}} ((1 - k_s)P^{O} - R) f(p)\, \mathrm{d}p - C(L^{O}). \tag{A.1}$$

where $\Psi(L^{O})$ is the probability that the listing agent will contact the buyer himself with $\Psi' > 0$ and $\Psi'' < 0$, R is the listing agent's reservation price, $P^{O} - R$ is the agent's surplus as the owner from the sale of the property, Φ, with $0 < \Phi + \Psi(L^{O}) \leqslant 1\ \forall L^{O}$, is the probability that one of the other members of the MLS contacts the buyer, $(1 - k_s)P^{O} - R$ is the listing agent's surplus as the owner when a selling agent finds the buyer, and $C(L^{O})$, $C' > 0$ and $C'' > 0$, is the search cost function of the listing agent. Given the large number of MLS members and the competition among them to sell the property, we take the probability of a sale by another MLS member as given.

The solution is given by the first-order conditions

$$\psi(L^{O}) \left[\int_{P^{O}}^{\bar{P}} f(p)\, \mathrm{d}p - (P^{O} - R) f(P^{O}) \right] + \Phi \left[\int_{P^{O}}^{\bar{P}} (1 - k_s) f(p)\, \mathrm{d}p - ((1 - k_s)P^{O} - R) f(P^{O}) \right] = 0 \tag{A.2}$$

and

$$\psi'(L^{\mathrm{O}})\int_{P^{\mathrm{O}}}^{\bar{P}}(P^{\mathrm{O}}-R)f(p)\,\mathrm{d}p - C'(L^{\mathrm{O}}) = 0. \tag{A.3}$$

Listing agent not the owner: If the listing agent is not the owner of the property, then his only source of surplus from the listing is the expected commission. He receives all of the commission if he finds the buyer, and shares the commission with the selling agent if the selling agent contacts the buyer. The problem of the listing agent is now to choose a search effort level, L^{N}, and advise the seller to choose a price, P^{N}, to maximize his expected profits

$$\max_{P^{\mathrm{N}},L^{\mathrm{N}}} \prod\nolimits^{\mathrm{N}}(P^{\mathrm{N}},L^{\mathrm{N}}) = \psi(L^{\mathrm{N}})\int_{P^{\mathrm{N}}}^{\bar{P}} kP^{\mathrm{N}}f(p)\,\mathrm{d}p + \Phi\int_{P^{\mathrm{N}}}^{\bar{P}}(k-k_{\mathrm{s}})P^{\mathrm{N}}f(p)\,\mathrm{d}p - C(L^{\mathrm{N}}). \tag{A.4}$$

The first-order conditions are given by

$$[\psi(L^{\mathrm{N}})k+\Phi(k-k_{\mathrm{s}})]\left[\int_{P^{\mathrm{N}}}^{\bar{P}} f(p)\,\mathrm{d}p - P^{\mathrm{N}}f(P^{\mathrm{N}})\right] = 0 \tag{A.5}$$

and

$$\psi'(L^{\mathrm{N}})k\int_{P^{\mathrm{N}}}^{\bar{P}} P^{\mathrm{N}}f(p)\,\mathrm{d}p - C'(L^{\mathrm{N}}) = 0. \tag{A.6}$$

A comparison of the first-order conditions in (A.2)–(A.3) with (A.5)–(A.6) yield the following results:

Lemma A.1. $P^{\mathrm{O}}(L) > P^{\mathrm{N}}(L)\ \forall L$; *for any given L, the listing agent chooses a higher price if he owns the property than if he does not.*

Proof. Eq. (A.5) indicates that P^{N} is independent of L^{N}, and is given by

$$\int_{P^{\mathrm{N}}}^{\bar{P}} f(p)\,\mathrm{d}p = P^{\mathrm{N}}f(P^{\mathrm{N}}). \tag{A.7}$$

Using proof by contradiction, we first show that P^{N} cannot equal P^{O}. Let $P^{\mathrm{O}} = P^{\mathrm{N}}$. Then, substituting (A.7) and $P^{\mathrm{O}} = P^{\mathrm{N}}$ in (A.2) would yield

$$[\psi(L)+\Phi]Rf(P^{\mathrm{O}}) = 0, \tag{A.8}$$

which is a contradiction because expression (A.8) is greater than zero. Therefore, $P^{\mathrm{O}} \neq P^{\mathrm{N}}$.

We next show that $P^{\mathrm{O}}(L) > P^{\mathrm{N}}(L)\ \forall L$. Eq. (A.2) characterizes the P^{O} at which the profit function is *maximized*, thus the slope of the profit function (the left-hand side of Eq. (A.2) will be positive for prices less than the optimal P^{O} and negative for prices greater than the optimal P^{O}. We have just shown that the left-hand side of

(A.2) is positive at $P^{\mathrm{O}} = P^{\mathrm{N}}$. Therefore, we need $P^{\mathrm{O}} > P^{\mathrm{N}}$ in order for (A.2) to equal to zero.

Lemma A.2. $L^{\mathrm{O}}\ (P) > L^{\mathrm{N}}(P)\ \forall P$; *for any given* P, *the listing agent expands a greater effort level if he owns the property than if he does not.*

Proof. Eqs. (A.3) and (A.6) reveal that the comparison of L^{O} and L^{N} for a given P depends on

$$\int_{P}^{\bar{P}} (P - R) f(p)\,\mathrm{d}p \textit{ versus } \int_{P}^{\bar{P}} kPf(p)\,\mathrm{d}p. \tag{A.9}$$

The term on the left is the seller's total expected surplus from the sale of the property. The right-hand term is the total expected commission earned by the listing agent and the selling agent. Since a seller would never list the property with an agent unless his surplus exceeds the total commission he has to pay to agents, the first integral above exceeds the second one for any given P. Given $\Psi(L)$ is a concave function while $C(L)$ is a convex function, the first-order conditions of (A.3) and (A.6) require that $L^{\mathrm{O}} > L^{\mathrm{N}}$ for any given P.

A.1. Equilibrium

We have shown in Lemmas A.1 and A.2 that $P^{\mathrm{O}}(L) > P^{\mathrm{N}}(L)\ \forall L$ and $L^{\mathrm{O}}(\mathrm{P}) > L^{\mathrm{N}}(\mathrm{P})\ \forall P$. This does not necessarily mean however that the equilibrium values of both P and L will be greater if they are chosen by an owner-agent than by a non-owner-agent. The answer will depend on the way P and L interact with each other. We first start with the comparison of the equilibrium price chosen by an owner agent, P^{O^*}, with the equilibrium price chosen by a non-owner agent, P^{N^*}.

Proposition A.1. $P^{\mathrm{O}^*} > P^{\mathrm{N}^*}$.

Proof. The first-order condition for P^{N}, Eq. (A.5), indicates that P^{N} is independent of the agents' search effort choices and is determined by

$$\int_{P^{\mathrm{N}}}^{\bar{P}} f(p)\,\mathrm{d}p = P^{\mathrm{N}} f(P^{\mathrm{N}}). \tag{A.10}$$

Combining this with Lemma A.1 that $P^{\mathrm{O}}(L) > P^{\mathrm{N}}(L)\ \forall L$ yields $P^{\mathrm{O}^*} > P^{\mathrm{N}^*}$.

We next compare the equilibrium effort level chosen by an owner agent, L^{O^*}, with the equilibrium effort level chosen by a non-owner agent, L^{N^*}.

Proposition A.2. L^{O^*} is greater than (smaller than) (equal to) L^{N^*} for

$$\int_{P^{O*}}^{\bar{P}} (P^{O^*} - R)f(p)\,\mathrm{d}p > (<)(=) \int_{P^{N*}}^{\bar{P}} kP^{N^*}f(p)\,\mathrm{d}p. \tag{A.11}$$

Proof. Follows directly from a comparison of (A.3) and (A.6).

The intuition behind this result is simple. The listing agent's effort level depends on his expected commission revenue. He will expend a greater effort level for his own property only if the surplus from the sale of his property exceeds the commission from the sale of his client's property. We know from Proposition A.1 that $P^{O^*} > P^{N^*}$. On the one hand, this generates a higher surplus for the agent from selling his own property. On the other hand, there is a smaller probability that a contacted buyer's reservation price will exceed P^{O^*} than it will P^{N^*}.

Consequently, $P^{O^*} > P^{N^*}$ may yield either of the three possibilities in Proposition A.2 depending on the properties of the function f and the magnitudes of P^{O^*} and P^{N^*}. Proposition A.2 also implies that the probability of a sale, which is the product of the probability of contacting a buyer and the probability that the contacted buyer will have a high enough reservation price, can be smaller or greater for a property owned by an agent than for a similar property not owned by an agent.

Appendix B

An alternative to using the inverse Mills ratio from the second probit model to correct for possible endogeneity is to use the predicted probability from the probit model in place of the inverse Mills ratio and the variable of interest, in this case *Owner-agent* (Vella and Verbeek, 1999). In addition to the inverse Mills ratio and the alternative stated, two other alternatives to correct for possible endogeneity are based on the use of a linear probability model instead of a probit model. In one case, the predicted probability from the linear probability model is used in place of the *Owner-agent* variable, similar to using the predicted probability from a probit model as stated previously. In the second case, the variable of interest, *Owner-agent* along with the residual from the linear probability model, are included in the selling price model. This is similar to including the inverse Mills ratio from a probit and the variable of interest.

The results for this set of models controlling for possible endogeneity using the full sample are presented in Table 11. In the first model we use the predicted probability from a probit instead of the owner-agent variable and the inverse Mills ratio. The coefficient increases to 6.8% for this variable. In the second model, we use the predicted probability from a linear probability model instead of the owner-agent variable. In this case, the coefficient increases to 9.6%. In the third model in this set, we include the owner-agent variable along with the residual from the linear probability model. The coefficient for *Owner-agent* is again 9.6%. In both models,

Table 11

Regression models of house prices corrected for endogeneity using a probit model and a linear probability model to estimate the probability of Owner-agent and the residual. These models have been suggested in the literature as possible alternatives to calculating an inverse Mills ratio from a probit model to correct for endogeneity. All models in this table are based on the sample of 193,007 houses sold during 1999–2002 in several Texas counties.

The dependent variable is the log of the selling price. All regressions include quarterly dummy variables (not reported for brevity) to control for potential serial effects and all regressions include dummy variables for MLS specified areas (not reported for brevity) to control for location. The estimates of the coefficients are presented in the table, with *t*-statistics reported in parentheses using heteroskedasticity-robust Huebner/White standard errors. Statistics with significance at the 1% level are denoted with a** and at the 5% level are denoted with a*.

Independent variable	Model 1 uses the predicted probability from a probit model to correct for endogeneity	Model 2 uses the predicted probability from a linear probability model to correct for endogeneity	Model 3 includes the residual from the linear probability model to correct for endogeneity	Model 4 is a GLS random effects model using the inverse Mills ratio from a probit model on sold properties to correct for sample selection and the inverse Mills ratio from a probit model on Owner-agent to correct for possible endogeneity
Constant	10.96	10.96	10.96	10.88
	(2247.70)**	(2249.93)**	(2250.58)**	(1447.52)**
Size	0.04	0.04	0.04	0.04
	(406.24)**	(404.63)**	(405.26)**	(432.01)**
Age	−0.05	−0.05	−0.05	−0.05
	−(144.28)**	−(144.47)**	−(144.61)**	−(167.63)**
Owner-agent			0.10	0.06
			(19.76)**	(21.63)**
Predicted probability that agent is an owner-agent based on the probit model in Table 4	0.07 (19.20)**			
Predicted probability that agent is an owner-agent based on the linear probability model in Table 4		0.10 (19.73)**		
Residual from the linear probability model in Table 4			−0.05 −(9.95)**	
Pool	0.08	0.08	0.08	0.08
	(80.78)**	(81.30)**	(80.69)**	(81.66)**
Fireplace	0.07	0.07	0.07	0.07
	(74.00)**	(73.75)**	(73.91)**	(85.60)**
Large lot	0.10	0.10	0.10	0.10
	(72.13)**	(72.98)**	(73.14)**	(81.23)**
Tenant	−0.06	−0.06	−0.06	−0.06
	−(31.03)**	−(30.52)**	−(30.42)**	−(30.10)**
Vacant	−0.07	−0.07	−0.07	−0.07

Table 11 (*continued*)

Independent variable	Model 1 uses the predicted probability from a probit model to correct for endogeneity	Model 2 uses the predicted probability from a linear probability model to correct for endogeneity	Model 3 includes the residual from the linear probability model to correct for endogeneity	Model 4 is a GLS random effects model using the inverse Mills ratio from a probit model on sold properties to correct for sample selection and the inverse Mills ratio from a probit model on Owner-agent to correct for possible endogeneity
	−(79.14)**	−(79.55)**	−(79.80)**	−(84.10)**
Builder	0.05	0.05	0.05	0.05
	(34.19)**	(34.33)**	(34.34)**	(37.60)**
Uniqueness	−0.10	−0.10	−0.10	−0.09
	−(53.78)**	−(54.15)**	−(54.21)**	−(53.87)**
Market competition	0.00	0.00	0.00	0.00
	(6.91)**	(6.69)**	(6.63)**	(10.06)**
Number of houses sold	193,007	193,007	193,007	193,007
Adjusted R^2	0.93	0.93	0.93	0.87
F	18,250	18,243	18,185	
Wald chi^2 (model)				1,290,000
Inverse Mills ratio from a probit model for sold versus not sold to correct for possible sample selection bias	−0.10 −(22.32)**	−0.11 −(24.14)**	−0.11 −(24.63)**	−0.10 −(23.45)**
Inverse Mills ratio from a probit model for agent-owned houses versus client-owned houses to correct for possible endogeneity of Owner-agent				−0.01 −(5.54)**

when using the linear probability model we have a much higher coefficient than when we use the inverse Mills ratio from a probit or the predicted probability from the probit. The final model in this set is a generalized least squares random effects model that includes the inverse Mills ratio from a probit to correct for endogeneity. The coefficient is 6.3% and is more in line with the results from the regression models corrected for endogeneity in Model 1 of Table 5 and Table 11 of Appendix B, using the probit models to generate the inverse Mills ratio or the predicted probability.

The results for the restricted sample and the alternative models for endogeneity are presented in Table 12. The owner-agent variables are consistent with the results in Table 9 and suggest that the results in Table 9 are reasonably robust.

Table 12

Regression models of house prices corrected for endogeneity using a probit model and a linear probability model to estimate the probability of Owner-agent and the residual. These models have been suggested in the literature as possible alternatives to calculating an inverse Mills ratio from a probit model to correct for endogeneity. All models in this table are based on the sample of 58,096 houses sold during 1999–2002 in several Texas counties.

The dependent variable is the log of the selling price. All regressions include quarterly dummy variables (not reported for brevity) to control for potential serial effects and all regressions include dummy variables for MLS specified areas (not reported for brevity) to control for location. The estimates of the coefficients are presented in the table, with *t*-statistics reported in parentheses using heteroskedasticity-robust Huebner/White standard errors. Statistics with significance at the 1% level are denoted with a** and at the 5% level are denoted with a*.

Independent variable	Model 1 uses the predicted probability from a probit model to correct for endogeneity	Model 2 uses the predicted probability from a linear probability model to correct for endogeneity	Model 3 includes the residual from the linear probability model to correct for endogeneity	Model 4 is a GLS random effects model using the inverse Mills ratio from a probit model on sold properties to correct for sample selection and the inverse Mills ratio from a probit model on Owner-agent to correct for possible endogeneity
Constant	10.91	10.91	10.91	10.83
	(1263.62)**	(1265.15)**	(1264.12)**	(960.23)**
Size	0.06	0.04	0.04	0.04
	(201.76)**	(199.78)**	(200.48)**	(219.06)**
Age	−0.05	−0.05	−0.05	−0.05
	−(70.08)**	−(70.04)**	−(70.30)**	−(80.95)**
Owner-agent			0.06	0.06
			(20.15)	(23.18)
Predicted probability that agent is an owner-agent based on the probit model in Table 8	0.06			
	(22.83)**			
Predicted probability that agent is an owner-agent based on the linear probability model in Table 8		0.06		
		(19.96)**		
Residual from the linear probability model in Table 8			−0.02	
			−(5.09)**	
Pool	0.08	0.08	0.08	0.07
	(45.51)**	(46.09)**	(45.10)**	(44.91)**
Fireplace	0.06	0.06	0.06	0.06
	(39.27)**	(39.10)**	(39.38)**	(45.73)**
Large lot	0.08	0.08	0.08	0.08
	(29.99)**	(30.13)**	(30.35)**	(33.17)**
Tenant	−0.06	−0.06	−0.06	−0.06
	−(17.57)**	−(17.38)**	−(17.22)**	−(17.48)**
Vacant	−0.07	−0.07	−0.07	−0.07

Table 12 (*continued*)

Independent variable	Model 1 uses the predicted probability from a probit model to correct for endogeneity	Model 2 uses the predicted probability from a linear probability model to correct for endogeneity	Model 3 includes the residual from the linear probability model to correct for endogeneity	Model 4 is a GLS random effects model using the inverse Mills ratio from a probit model on sold properties to correct for sample selection and the inverse Mills ratio from a probit model on Owner-agent to correct for possible endogeneity
	−(47.81)**	−(47.73)**	−(48.00)**	−(50.45)**
Builder	0.06	0.06	0.06	0.06
	(18.83)**	(18.81)**	(18.67)**	(21.52)**
Uniqueness	−0.09	−0.09	−0.09	−0.08
	−(27.32)**	−(27.58)**	−(27.64)**	−(25.25)**
Market competition	0.00	0.00	0.00	0.00
	(8.29)**	(8.28)**	(8.25)**	(10.93)**
Number of houses sold	58,096	58,096	58,096	58,096
Adjusted R^2	0.94	0.94	0.94	0.87
F	5,972	5,952	5,966	
Wald chi^2 (model)				383,415
Inverse Mills ratio from a probit model for sold versus not sold to correct for possible sample selection bias	−0.10 −(10.58)**	−0.10 −(11.11)**	−0.11 −(11.77)**	−0.09 −(11.05)**
Inverse Mills ratio from a probit model for agent-owned houses versus client-owned houses to correct for possible endogeneity of Owner-agent				−0.01 −(3.72)**

References

Anglin, P.M., Arnott, R., 1991. Residential real estate brokerage as a principal–agent problem. Journal of Real Estate Finance and Economics 4, 99–125.

Arnold, M.A., 1992. The principal–agent relationship in real estate brokerage services. Real Estate Economics 20, 89–106.

Diamond, P., 1998. Managerial incentives: on the near linearity of optimal compensation. Journal of Political Economy 106, 931–957.

Federal Trade Commission, 1983. The Residential Real Estate Brokerage Industry, vols. I and II. U.S. Government Printing Office, Los Angeles.

Geltner, D., Kluger, B.D., Miller, N.G., 1991. Optimal price and selling effort from the perspectives of the broker and seller. Real Estate Economics 19, 1–24.

Hart, O., Holmstrom, B., 1987. The theory of contracts. In: Bewley, T. (Ed.), Advances in Economics. Cambridge University Press, New York.

Haurin, D., 1988. The duration of marketing time of residential housing. Real Estate Economics 16, 396–410.

Heckman, J.J., 1978. Sample selection bias as a specification error. Econometrica 47, 153–161.

Holmstrom, B., Milgrom, P., 1987. Aggregation and linearity in the provision of intertemporal incentives. Econometrica 55, 303–328.

Lancaster, T., 1990. The Econometric Analysis of Transition Data. Cambridge University Press, New York.

McAfee, R.P., McMillan, J., 1986. Bidding for contracts: a principal–agent analysis. Rand Journal of Economics 17, 326–338.

Miceli, T.J., 1991. The multiple listing service, commission splits and broker effort. Real Estate Economics 19, 548–566.

Rubinstein, A., 1982. Perfect equilibrium in a bargaining model. Econometrica 50, 97–109.

Springer, T.M., 1996. Single-family housing transactions: seller motivations, price and, marketing time. Journal of Real Estate Finance and Economics 13, 237–254.

Vella, F., Verbeek, M., 1999. Estimating and interpreting models with endogenous treatment effects. Journal of Business and Economic Statistics 17, 473–478.

Williams, J., 1998. Agency and brokerage of real assets in competitive equilibrium. Review of Financial Studies 11, 239–280.

Yavas, A., Yang, X., 1995. The strategic role of listing price in marketing real estate: theory and evidence. Real Estate Economics 23, 347–368.