How to Sell Your House: Theory and Evidence[†]

Antonio Merlo *University of Pennsylvania*
Françcois Ortalo-Magné *University of Wisconsin – Madison*
John Rust[‡] *University of Maryland*

February, 2008

Abstract

This paper formulates and solves the problem of a homeowner who wants to sell their house for the maximum possible price net of transactions costs (including real estate commissions). The optimal selling strategy consists of an initial list price with subsequent weekly decisions on how much to adjust the list price until the home is sold or withdrawn from the market. The solution also yields a sequence of reservation prices that determine whether the homeowner should accept bids from potential buyers who arrive stochastically over time with an expected arrival rate that is a decreasing function of the list price. This model was developed to provide a theoretical explanation for list price dynamics and bargaining behavior observed for a sample of homeowners in England in a new data set introduced by Merlo and Ortalo-Magné (2004). One of the puzzling features that emerged from their analysis (but which other evidence suggests holds in general, not just England) is that list prices are *sticky*: By and large homeowners appear to be reluctant to change their list price, and are observed to do so only after a significant amount of time has elapsed if they have not received any offers. This finding presents a challenge, since the conventional wisdom is that traditional rational economic theories are unable to explain this extreme price stickiness. Recent research has focused on "behavioral" explanations such as loss aversion in attempt to explain a homeowner's unwillingness to reduce their list price. We are able to explain the price stickiness and most of the other key features observed in the data using a model of rational, forward looking, risk-neutral sellers who seek to maximize the expected proceeds from selling their home net of transactions costs. The model relies on a very small fixed "menu cost" of changing the list price, amounting to less than 6 thousandths of 1% of the estimated house value, or approximately 12 pounds for a home worth 200,000 pounds. A key reason why such a small menu cost produces so much list price stickiness is due to a relatively inelastic relationship between the list price and the expected rate of arrival of potential buyers.

Keywords: housing, bargaining, sticky prices, optimal selling strategy, dynamic programming.
JEL classification: H5

[†] PRELIMINARY DRAFT: DO NOT QUOTE. This research was motivated by "Bargaining over Residential Real Estate: Evidence from England" *Journal of Urban Economics* (2004) by Antonio Merlo and François Ortalo-Magné. We are grateful for research support from the National Science Foundation on NSF collaborative research grant 0635806 "Models of Bargaining and Price Determination of Residential Real Estate, with and without Real Estate Agents". Of course, none of the opinions or conclusions expressed in this paper are endorsed or approved by the NSF.
[‡]**Correspondence address:** Department of Economics, University of Maryland, College Park, MD 20742, phone: (301) 405-3489, email: `jrust@gemini.econ.umd.edu`.

1 Introduction

Buying and selling a home is one of the most important financial decisions most individuals make during their lifetime. Home equity is typically the biggest single component of the overall wealth of most households, and given the highly leveraged situation that most households are in (where mortgage debt is a high fraction of the overall value of the home), mistakes or adverse shocks during the selling process can have very serious consequences for their financial well-being.

Given its importance, we would expect *a priori* that households have strong incentives to behave rationally and strategically when they sell their home. In particular, it seems reasonable to model the household's objective as trying to maximize the expected gains from selling their home net of transactions costs.[1]

Surprisingly, despite its importance, the "home selling problem" has been understudied both theoretically and empirically. In pioneering work Salant (1991) formulated and solved for the optimal selling strategy of a risk neutral seller using dynamic programming. Salant's model involves an initial choice by the household whether to use a real estate agent to help sell their home, versus deciding to save on the high commissions charged by most real estate agencies and follow a "for sale by owner" selling strategy. Under either of these options, the seller must also choose a list price each period the home is up for sale, and whether to accept a bid for the home when one arrives, or to wait and hope that a higher bid will arrive in the near future. Salant showed that the optimal solution generally involves a strictly monotonically declining sequence of list prices, and that it is typically optimal to begin selling the home by owner, but if no acceptable offers have arrived within a specified interval of time, the seller should retain a real estate agent. Under some circumstances, the optimal list price can jump up at the time the seller switches to the real estate agency, but list prices decline thereafter.

To our knowledge, the strong implications of Salant's model — particularly its prediction that list prices should decline monotonically after a home has been listed with a real estate agent — have never been tested empirically. Horowitz (1992) was one of the first attempts to empirically estimate and test a dynamic model of the home seller's problem.[2] Unfortunately, however, Horowitz's work cannot address

[1] It is also possible to model the selling behavior of risk averse sellers via relatively straightforward adjustments to a model of a risk neutral seller. A risk averse seller will set somewhat lower list prices than a risk neutral one, and will accept lower offers in order to reduce the risk of "letting a fish off the hook." However we will show that the broad qualitative features of an optimal selling strategy are the same regardless of the degree of risk aversion.

[2] Horowitz's article is especially important from a methodological point of view, and was awarded the Sir Richard Stone Prize

the issue of whether list prices should decline over time since unlike Salant, Horowitz (who did not cite and thus appeared to be unaware of Salant's work) adopted an infinite horizon stationary search framework that requires the list and reservation prices to be time invariant. Horowitz's model implies that the duration to sale of a house is geometrically distributed, and he estimated his model using data on the transaction price and duration to sale for a sample of 1196 homes sold in Baltimore, Maryland in 1978.

Horowitz concluded that his model "gives predictions of sale prices that are considerably more accurate than those of a standard hedonic price regression." (p. 126). However his other main conclusion, namely that his model "explains why sellers may not be willing to reduce their list prices even after their houses have remained unsold for long periods of time" is unwarranted because time invariance of list and reservation prices are inherent features of Horowitz's stationary search framework, so his model is logically incapable of addressing the issue of whether optimal list prices should decline over time. Further, his data set does not appear to contain any information on changes in the list price between when the home was initially listed and when it was finally sold.

It seems that the question of whether optimal list prices should or should not decline over time can only be addressed in a non-stationary, finite-horizon framework such as Salant's, or else in a stationary infinite-horizon framework that includes state variables such as duration since initial listing, or duration since previous offer, as state variables. However once one includes a state variable such as duration since initial listing, the seller's problem automatically becomes a non-stationary dynamic programming problem that is essentially equivalent to Salant's formulation.

The model presented in this paper is motivated by the empirical findings of Merlo and Ortalo-Magné (2004), (abbreviated as MO below) who introduced a new data set that (to our knowledge) provides the first opportunity to study the house selling decision in considerable detail. MO's study is based on a panel data of detailed transaction histories of 780 homes that were sold via a real estate agency in England between June 1995 and April 1998. The data include all listing price changes and all offers made between initial listing and the final sale agreement. MO characterized a number of key stylized facts pertaining to the sequence of events that occur within individual property transaction histories, and discussed the limitations of existing theories of a home seller's behavior in explaining the data.

From the perspective of this paper, the most striking finding from MO's analysis is that housing list

in 1992 for "for the best paper with substantive econometric application that has been published in the preceding two volumes of the *JAE*".

prices appear to be highly *sticky*. That is, 77% of the house sellers in their data never changed the initial list price between the time the house was initially listed and when it was sold. List prices were changed only once in 21% of the cases, and only twice in the remaining 2% of the cases observed: none of the homeowners made more than 2 changes in their initial listing price over the 10.2 week mean duration between initial listing and the sale of the home. MO conclude that "listing price reductions are fairly infrequent; when they occur they are typically large. Listing price revisions appear to be triggered by a lack of offers. The size of the reduction in the listing price is larger the longer a property has been on the market."

This finding presents a challenge, since the conventional wisdom is that traditional rational, forward looking economic theories are unable to explain this extreme price stickiness. In particular, the findings are inconsistent with Salant's model, which predicts that list prices should decline monotonically over the period the home is on the market. Recent research has focused on "behavioral" explanations for price stickiness. For example, Genesove and Mayer (2001) (abbreviated as GM below) appealed to Kahneman and Tversky's (1991) theory of *loss aversion* to explain the apparent unwillingness of owners of condominiums in Boston to reduce their list price in response to downturns in the housing market.

GM assumed that a seller's previous purchase price serves as the "reference point" required by the model of loss aversion, and use this to explain a pattern of *asymmetric* price stickiness: "In a boom, houses sell quickly at prices close to, and many times above, the sellers' asking prices. In a bust, however, homes tend to sit on the market for long periods of time with asking prices well above expected selling prices, and many sellers eventually withdraw their properties without sale." (p. 1233). GM argued that prices are sticky in the downward direction during a housing bust because "When house prices fall after a boom, as in Boston, many units have a market value below what the current owner paid for them. Owners who are averse to losses will have an incentive to attenuate that loss by deciding upon a reservation price that exceeds the level they would set in the absence of a loss, and so set a higher asking price, spend a longer time on the market, and receive a higher transaction price upon a sale." GM concluded that "The support for nominal loss aversion in the Boston condominium market is quite striking. Sellers whose unit's expected selling price falls below their original purchase price set an asking price that exceeds the asking price of other sellers by between 25 and 35 percent of the percentage difference between the two." (p. 1235).

The loss aversion model of seller behavior, at least as GM describe it[3], is a primarily *backward looking* theory that is inconsistent with the rational forward looking calculations underlying the dynamic programming (DP) models. The DP models assume that homeowners have rational expectations about the amount *prospective buyers* are willing to pay for their home. If the housing market turns bad and it is no longer possible for the homeowner to expect to sell their home at a higher price than they paid for it, a rational seller will regard this as an unfortunate bygone, but will realize that whatever they paid for their house in the *past* may have little bearing on how they should try to sell their house *now*, which requires a realistic assessment of what will happen in the *future.* While many sellers do have the option not to sell their homes if market conditions turn bad, not selling a home or not selling one sufficiently quickly can entail serious losses as well.

Thus, if we take a broader view about what a "loss" is, the loss aversion model — with its prediction that home sellers are backward looking and impose inflexible, unrealistic demands on what they will sell their house for — may actually lead a seller to incur significantly higher *ex post* losses than would have been the case if the seller had been more realistic and forward looking. For example, if a seller sets an unrealistically high price at the start of a market downturn and rejects offers that were close but somewhat less than the price the seller previously paid, the extra delay in selling the home that results from this inflexibility may lead to *lower* transaction prices (or force the seller to withdraw their home from the market) if housing prices have fallen even further in the meantime. The loss aversion model is not a fully specified dynamic behavioral model, and treatments that use it (such as GM) appear to presume that sellers always have a costless option *not* to sell their homes. However this is unlikely to be the case for many sellers: some sellers (such as those facing foreclosure, or who need to sell due to a job move, or a change in family situation such as divorce) are selling under duress, and even others who are under less time pressure may perceive a substantial "hassle cost" of having their home listed, cleaned and ready to show to prospective buyers on short notice. Thus, it is not clear that if we accounted for a broader notion of what "loss" is, then more sophisticated model of loss aversion would necessarily result in the extreme form of downward rigidity that is predicted by the naive versions (i.e. to refuse to sell your house at a lower price

[3] Unlike the work of Salant and Horowitz, the Genesove and Mayer paper only mentions but does not actually develop or solve a formal model of loss aversion. As a result, it is unclear what the specific behavioral implications of loss aversion are in this context. Most applications of loss aversion are in simple static decision contexts. However see Bowman, Minehart and Rabin (1999) for a dynamic application of loss aversion to a consumption/savings problem. This latter work shows that dynamic theories of loss aversion have both backward looking and forward looking elements, and in these more sophisticated versions, agents are modeled as taking into consideration the effect of current information and decisions on future reference points.

than you bought it for). In some sense, the naive version of loss aversion described by GM is an example of how *not* to sell your house.

The primary contribution of this paper is to show that the high degree of list price stickiness is consistent with an optimal selling strategy from a forward looking dynamic programming model with risk-neutral sellers who have rational expectations about the ultimate selling price of their homes. In addition, we show that this model is also consistent with most of the other key features of selling behavior that emerged from MO's empirical analysis of the English housing data. The new model of the seller's decision problem we introduce is specialized to capture specific features of the English housing market. However an important difference in our model relative to Salant's model is that we assume that there is a small fixed *menu cost* involved in changing the list price. As is well known, this type of non-convexity can generate *regions of inaction* where it is optimal for the seller not to change the list price even though the list price inherited from the previous week is not the optimal forward-looking list price that the seller would choose if there was no cost of changing the list price. What we show is that a *very small* menu cost, amounting to less than 6 thousandths of 1% of the estimated house value, or approximately 12 pounds for a home worth 200,000 pounds is sufficient to generate the high degree of price stickiness that we observe in MO's London home transaction data.

A key reason why very small menu cost yield a high degree of list price stickiness in our model is that the estimated relationship between the list price and the expected rate of arrival of potential buyers is relatively inelastic. More precisely, relatively small adjustments in the list price hardly affect the expected sale probability while impacting the expected sale price. When a buyer makes a bid for a home it is generally not equal to the seller's posted list price. In the English housing data, only 15% of all transactions occurred at the list price, and many of these transactions resulted from a bargaining process wherein the first (rejected) offer made by the buyer was significantly below the list price. Thus, both buyers and sellers expect that the list price is not a firm demand, and initial bids are typically lower than the list price and most of the actual price determination process occurs during the ensuing negotiation process. In particular, even though the list price is a piecewise flat function of duration on the market, the seller's reservation price is a continuous and strictly monotonically declining function of duration on the market. Thus, the longer a home has been on the market the lower the expected transaction price will be, and this is largely due to the steady decline in the seller's reservation price rather than any decline in the list price.

An implication of this finding is that houses are generally *overpriced* when they are first listed. In the

English housing data the degree of overpricing is not huge: the initial list is on average 5% higher than the ultimate transaction price for the home. Our model is also able to match the degree of overpricing in the initial list price, and in general our model is consistent with the overall trajectory of list prices, including MO's finding that the magnitude of list price reductions are largest when a home has been on the market for a long time. However our model is also consistent with *underpricing* under different assumptions about arrival rates and buyer behavior. Underpricing can result when the arrival rate of buyers is sufficiently sensitive to the list price, and when multiple buyers can arrive at the same time, resulting in an auction situation and potential "bidding war" that tends to drive the final transaction price to a value far higher than the list price. However even in the absence of auctions, initial bids and final transaction prices in excess of the list price are observed in approximately 4% of all sales in the English housing data. Our model allows for the possibility of such "overbidding" which results from the fact that in England, the seller has no legal obligation to accept a bid that is greater than or equal to the list price. Previous models, including both Salant's and Horowitz's models, do not allow for the possibility that a bid or transaction price would ever exceed the list price.

In this paper, we do not explicitly model the behavior of buyers and the bargaining game that leads to the sale of a house. Typically, when a buyer arrives and makes an initial offer for the home, it is just the first move in a *bargaining subgame* where the buyer and the seller negotiate over the sale price. This negotiation may either lead to a transaction, when the buyer and seller reach an agreement over the terms of the sale, or end with the buyer leaving the *bargaining table* when no agreement can be found. Rather than modeling this situation as a bargaining model with two-sided incomplete information, we capture the key features of this environment by specifying a simplified model of buyers' bidding behavior. In particular, we assume that if a potential buyer arrives, he makes up to n consecutive offers which are drawn from bids distributions that depend, among other things, on the list price and the amount of time the house has been on the market.[4] The seller can either accept or reject each offer, but after any rejection there is a positive probability the buyer "walks" (i.e. decides not to make a further counteroffer and move on and search for other properties). [5]

[4] In our empirical work we assume that $n = 3$, which is consistent with the maximum number of counteroffers observed in the English housing data set.

[5] As is well known, game theoretic models of bargaining with two-sided incomplete information typically admit multiple equilibria — and often a continuum of them. We avoid these problems by treating buyers as *bidding automata* using simple piecewise linear bidding functions with exogenously specified random termination in the bargaining process. Thus, by virtue of treating buyers as simplified automata, we avoid the problems of computing a Bayesian Nash equilibrium for the bargaining

One aspect that we have not attempted to model in this paper is the seller's decision whether to use a real estate agent, something that was a key focus of Salant's analysis. We agree that this is a very interesting and important issue, but it one that we cannot say much about empirically since MO's data set was for a self-selected sample of sellers who had chosen a particular network of real estate agencies. Although it is straightforward to extend our model to include a decision of whether to use a real estate agent (and if so, which one of several competing agents), this decision depends critically on the seller's beliefs of how different real estate agents will affect the arrival rate and selection of potential buyers who view their properties. We need additional data that includes individuals who have chosen different real estate agents and individuals who have abstained from using a real estate agency in order to have a chance of inferring what these beliefs are in a "revealed preference" type of analysis. The final section of our paper does show how our current model can be used to derive "willingness to pay" for the services of a real estate agent, but these calculations depend on hypotheses about the seller's beliefs of what arrival rates would be and what types of potential buyers would be bidding on their home if they were to try to sell their home without the assistance of a real estate agent.

Section 2 provides a brief review of the English housing data analyzed by MO, reviewing the legal environment, the overall housing market, and the way the real estate agency operates in the parts of England where the data were gathered. We refer the reader to MO for a more in depth analysis, but we do attempt to lay out the key features of the data that we attempt to account for in this analysis. Section 3 introduces our model of the seller's decision problem, including the model of buyer arrival and bidding behavior that constitutes the key "belief objects" that must be estimated to empirically implement and test our model. Section 4 presents a simplified model arrival of potential buyers and the "within week bargaining subgame" between the buyer and seller. Section 5 presents estimation results based both on quasi maximum likelihood (QML) and simulated minimum distance (SMD) estimation methods. We show there are substantial problems with the smoothness of the estimation criterion using either of these approaches, which calls into question the validity of standard first order asymptotic theory and the usual methods for computing parameter standard errors and goodness of fit statistics. So instead of focusing on presenting

subgame of the overall selling process. We believe our simplified treatment of buyers is justified by the fact that the English housing data contain very limited information on the buyers: while the data allow us to follow the decisions of sellers, it does not follow individual buyers and record their search and bargaining behavior. Essentially the only information we have about buyers are records of the sequence of bids they make and the identity of the ultimate buyer. We believe that our model may provide a reasonably good approximation to a seller's beliefs in an fluid environment where there is a high degree of heterogeneity in potential bidders and sellers have a great deal of uncertainty about their motivations and outside options.

statistics of dubious validity, we provide a fairly extensive comparison of the predictions of our model to the features we observe in the English housing data. While we have not yet found the "best fitting" parameter estimates or specification of the model (due largely to the non-smoothness of the QML and SMD estimation criteria), we argue that the provisional or trial parameter values and model specification that we present here already provides a very good approximation to a wide range of features that MO documented in their analysis of the English housing data. Section 6 presents a number of hypothetical simulations and calculations using this model. In addition to calculating a seller's willingness to pay for the services of a real estate agency, we also show how risk aversion affects the seller's strategy. We can even use our model to evaluate our own — fully dynamic — model of loss aversion on the part of sellers, although this work is currently in progress. Instead we conduct other calculations with our risk neutral seller model to show how different beliefs on the part of sellers can result in underpricing, and even situations where list prices can increase rather than decrease as a function of time on the market. A final calculation is to show how seller behavior would be changed if seller's were legally obligated to sell to any buyer who is willing to pay the seller's posted list price. Section 7 provides some concluding comments and suggestions for future research.

2 The English Housing Data

In England, most residential properties are marketed under sole agency agreement. This means that a property is listed with a single real estate agency that coordinates all market related activities concerning that property from the time it is listed until it either sells or is withdrawn. Agencies represent the seller only. Listing a property with an agency entails publishing a sheet of property characteristics and a listing price. Although not legally binding, the listing price is generally understood as a price the seller is committed to accept.

The listing price may be revised at any time at the discretion of the seller. The seller does not incur any cost when revising the listing price, except the cost of communicating the decision to the agent. The agent has to adjust the price on the posted property sheet and reprint any property detail sheets in stock, a minimal cost.

Potential buyers search by visiting local real estate agencies and viewing properties. A match between the seller and a potential buyer occurs when the potential buyer makes an offer. Within a match, the general practice is for the seller to either accept or reject offers. In the event the seller rejects an offer, the potential buyer either makes another offer or walks away. If agreement occurs, both parties engage the administrative procedure leading to the exchange of contracts and the completion of the transaction. This procedure typically lasts three to eight weeks. During this period, among other things, the buyer applies for mortgage and has the property surveyed. Each party may cancel the sale agreement up to the exchange of contracts.

For each property it represents, the agency keeps a file containing a detailed description of the property, its listing price, and a record of listing price changes, offers, and terms of the sale agreement, as required by law. The information contained in each individual file is also recorded on the accounting register that is used by each agency to report to the head office. Although all visits of a property by potential buyers are arranged by the listing agency, recording viewings is not required either by the head office or by law. However, individual agencies may require their agents to collect this information for internal management purposes.

The first data set we will use in our research was obtained from the sales records of four real estate agencies in England. These agencies are all part of Halifax Estate Agencies Limited, one of the largest network of real estate agents in England. Three of these agencies operate in the Greater London metropolitan area, one in South Yorkshire. Our sample consists of 780 complete transaction histories of properties listed and sold between June 1995 and April 1998 under sole agency agreement. Each entry in our data was validated by checking the consistency of the records in the accounting register and in the individual files.

Each observation contains the property's characteristics as shown on the information sheet published by the agency at the time of initial listing, the listing price and the date of the listing. If any listing price change occurs, we observe its date and the new price. Each match is described by the date of the first offer by a potential buyer and the sequence of buyer's offers within the match. When a match is successful, we observe the sale agreed price and the date of agreement which terminate the history. In addition, for the properties listed with one of our Greater London agencies (which account for about a fourth of the observations in our sample), we observe the complete history of viewings. Since events are typically recorded by agents within the week of their occurrence, we use the week as our unit of measure of time. Our data spans two geographic areas with different local economic conditions and two different phases of the cycle in the housing market. While the local economy in Greater London has been experiencing a prolonged period of sustained growth, this has not been the case in South Yorkshire. Furthermore, from June 1995 to April 1998, the housing market in the Greater London metropolitan area went from a slow recovery to a boom. While this transition occurred gradually, for ease of exposition we refer to 1995-96 as the recovery and to 1997-98 as the boom.

This data set was the one analyzed by Merlo and Ortalo-Magné (2004), and their main findings can be summarized as follows. First, listing price reductions are fairly infrequent; when they occur they are typically large. Listing price revisions appear to be triggered by a lack of offers. The size of the reduction in the listing price is larger the longer a property has been on the market. Second, the level of a first offer relative to the listing price at the time the offer is made is lower the longer the property has been on the market, the more the property is currently over-priced, and if there has been no revision of the listing price. Negotiations typically entail several offers. About a third of all negotiations are unsuccessful (i.e., they end in a separation rather than a sale). The probability of success of a negotiation decreases with the number of previous unsuccessful negotiations. Third, in the vast majority of cases, a property is sold to the first potential buyer who makes an offer on the property (i.e., within the first negotiation), although not necessarily at the first offer. The vast majority of sellers whose first negotiation is unsuccessful end up selling at a higher price, but a few end up accepting a lower offer. The higher the number of negotiations between initial listing and sale agreement, the higher the sale price.

Figure 2.1 illustrates two typical observations in the data set. We have plotted list prices over the full duration from initial listing until sale as a ratio of the initial listing price. The red dots plot the first offer and the blue squares are the second offers received in a match. The stars plot the final accepted transaction prices. Thus, the seller of property 1046 in the left hand panel of figure 2.1 experienced 3 separate matches. The first occurred in the fourth week that the property was listed, and the seller rejected the first bid by a bidder equal to 95% of the list price. The buyer "walked" after the seller rejected the offer. The next match occurred on the sixth week on he market. The seller once again rejected this second prospective buyer's first bid, which was only 93% of the list price. However this time the bidder did not walk after this first rejection, but responded with a second higher offer equal to 95% of the list price. However when the seller rejected this second higher offer, the second bidder also walked. The third match occurred in the 11th week the home was on the market. The seller accepted this third bidder's opening offer, equal to 98% of the list price. Note that there were no changes in the initial list price during the 11 weeks this property was on the market.

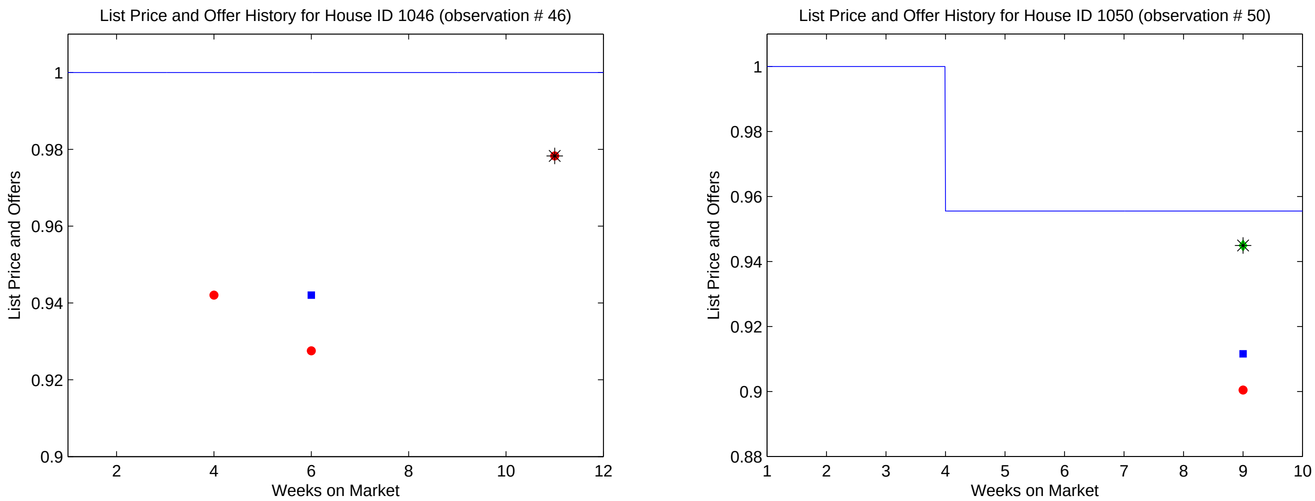

Figure 2.1 Selected Observations from the London Housing Data

The right hand panel plots a case where there was a decrease in the list price by 5% in the fourth week this property was on the market. After this price decrease another 5 weeks elapsed before the first offer was made on this home, equal to 90% of the initial list price. The seller rected this offe and the bidder made a counteroffer equal to 91% of the initial list price. The seller rejected this second offer too, prompting the bidder to make a final offer equal to 94.5% of the initial list price which the seller accepted.

Figure 2.2 plots the number of observations in the data set and the mean and median list prices as a function of the total number of weeks on the market. The left hand panel plots the number of observations (unsold homes reamining to be sold) as a function of duration since initial listing. For example only 54 of the 780 observations remain unsold after 30 weeks on the market, so over 93% of the properties listed by this agency sell within this time frame. If we compute the ratio of first offers received to the number of remaining unsold properties, we get a crude estimate of the offer arrival rate (a more refined model and estimate of this rate and its dependence on the list price will be presented subsequently). There is an 11% arrival rate in the first week a home is listed, meaning that approximately 11% of all properties will receive one or more offers in the first week after the home is listed with the real estate agency. The arrival rate increases to approximately 15% in weeks 2 to 6, then it decreases to approximately 12% in weeks 7 to 12, and then drops to about 10% thereafter, although it is harder to estimate arrival rates for longer durations given the declining number of remaining unsold properties.

The right hand panel of figure 2.2 plots the mean and median list prices of all unsold homes as a

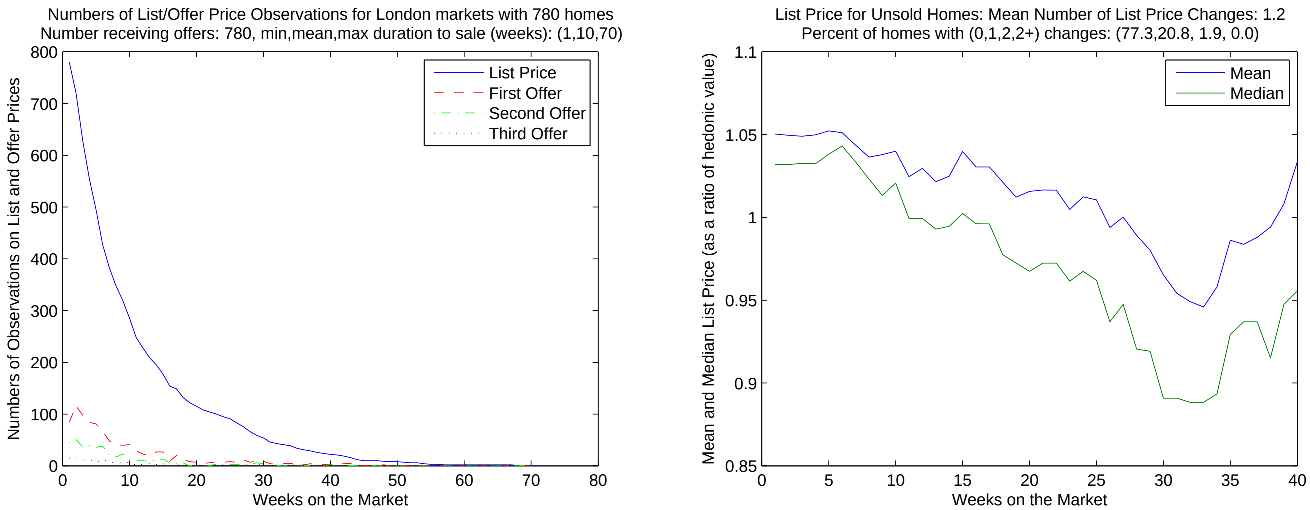

Figure 2.2 Number of Observations and List Prices by Week on Market

function of the duration on the market. We have normalized the list prices by dividing by the predicted sale price from a hedonic price regression using the extensive set of housing characteristics that are available in the data set (e.g. location of home, square meters of floor space, number of baths, bedrooms, and so forth). However the results are approximately the same when we normalize using the *actual* transaction prices instead of the regression predictions: this is a consequence of the fact that the hedonic regression provides a very accurate prediction of actual transaction prices.

We see from the right panel of figure 2.2 that initially houses are listed at an average of a 5% premium above their ultimate selling prices, and there is an obvious downward slope in both the mean and median list prices as a function of duration on the market. However the slope is not very pronounced: even after 25 weeks on the market the list price has only declined by 5%, so that at this point list prices are approximately equal to the *ex ante* expected selling prices. The apparently continuously downward slope in mean and median list prices is misleading in the sense that, as we noted from figure 2.1, individual list price trajectories are piecewise flat with discontinuous jumps on the dates where price reductions occur. Averaging over these piecewise flat list price trajectories creates an illusion that list prices are continuously declining as a function of duration on the market, but we emphasize again that the individual observations do *not* have this property.

Figure 2.3 plots the distribution of sales prices (once again normalized as a ratio to the predicted transaction price) and the distribution of duration to sale. The left hand panel of figure 2.3 plots the distribution

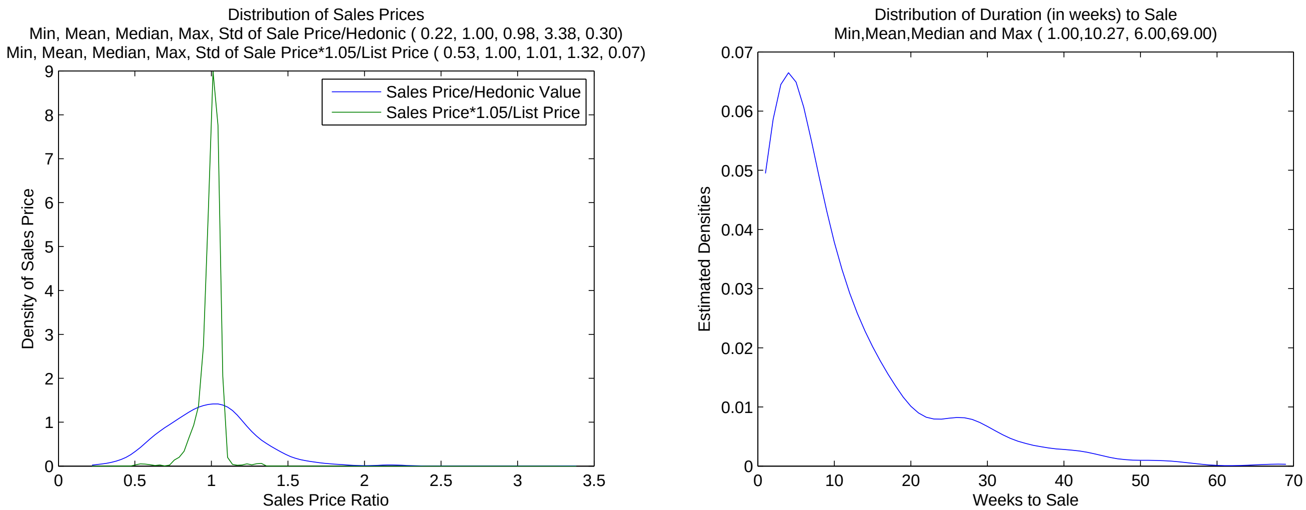

Figure 2.3 Distribution of Sale Prices and Duration to Sale

of sales price ratios. There are two different distributions shown: the blue line is the distribution of ratios of sale price to the hedonic prediction of sales price, and the red line is the distribution of the ratio of sales price to the initial list price, multiplied by 1.05 (this latter factor is the average markup of the initial list price over the ultimate transaction price, as noted above). Both of these distributions have a mean value of 1 (by construction), but clearly the distribution of the adjusted sales price to list price ratio is much more tightly concentrated than the distribution of sales price to hedonic value ratios. Evidently there is significant information about the value of the home that affects the seller's decision of what price to list their home at that is not contained in the *x* variables used to construct the hedonic price predictions. The model we present in section 3 will account for this extra *private information* about the home that we are unable to observe. However even when this extra information is taken into account, there is still a fair amount of variation/uncertainty in what the ultimate sales price will be, even factoring in the information revealed by the initial list price: the sales price can vary from as low of only 53% of the adjusted list price to 32% higher than the adjusted list price.

The right hand panel of figure 2.3 plots the distribution of times to sale. This is a clearly right skewed but unimodal distribution with a mean time to sale of 10.27 weeks and a median time to sale of 6 weeks. As we noted above, over 90% of the properties in our data set were sold within 30 weeks of the date the property was initially listed. Scatterplots relating time to sale to the ratio of the list price to the hedonic value (not shown) do not reveal any clear negative relationship between the degree of "overpricing" (as

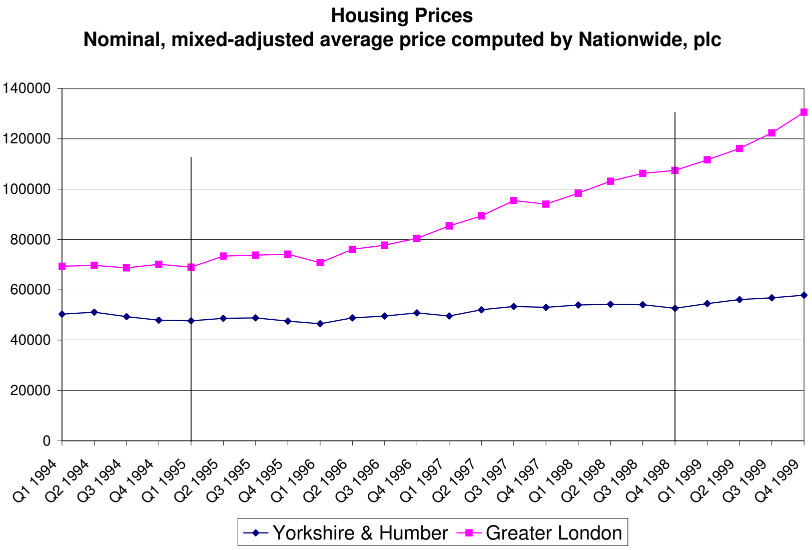

Figure 2.4 Price Indices in the Regions Covered in the English Housing Data

indiciated by high values of this ratio) and longer times to sale. Thus, we do not find any clear evidence at this level supporting the "loss aversion" explanation advocated by Genesove and Mayer. However an alternative explanation is the fact that prices in London were generally rising over the time period of the data (see figure 2.4 above), so an alternative explanation that few of the sellers had experienced any adverse shocks, and thus our sample is not in a regime where the "downward stickiness" prediction of the loss aversion theory is relevant.

We conclude our review of the English housing data by showing figure 2.5, which plots the distributions of the first offer received and the best (highest) offer received as a ratio of the current list price for properties with different durations on the market. The left hand panel of figure 2.5 shows the distributions of first offers. We see that in the first week a home is listed, the mean first offer received is 96% of the list price (which is also the initial list price in this case). However first offers range from a low of only 79% of the list price to a high of 104% of the list price. We see that even accounting for declines in the list price with duration on the market, that first offers made on properties tend to decline the longer the property is on the market. There is a notable leftware shift in the distribution of first offers for offers received on homes that have been on the market for 20 weeks, where the mean first offer is only 91% of the list price in effect for properties that are still unsold after 20 weeks.

The right hand panel of figure 2.5 shows the distribution of the best offers received in a match. In

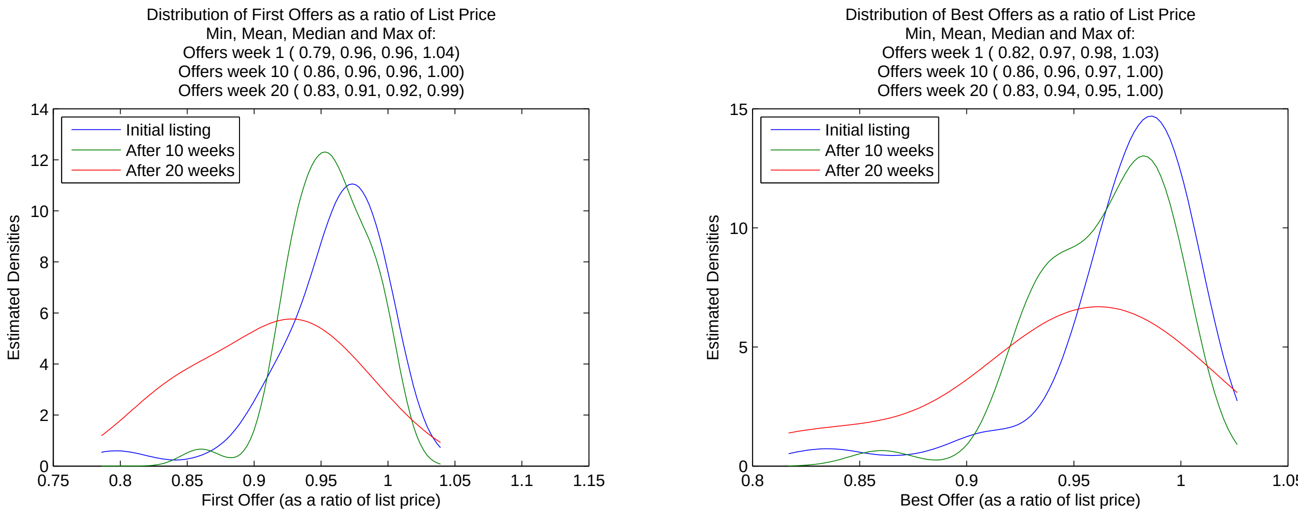

Figure 2.5 Distribution First Offer and Best Offer as a Ratio of List Price

the first few weeks the best offers show only modest improvement over the first offers received (e.g. the best offer is 97% of the list price, whereas the first offer is 96% of the list price). However we see more significant improvement in offers received for homes that were still unsold after 20 weeks: the best offer received is 94% of the current list price, which is 3 percentage points higher than the ratio of the first offer to the list price.

In addition to this data, our analysis will also rely on a new data set for England that we recently collected from a real estate agency in the city of Reading. This agency operates a paperless office where all realtors work cooperatively on all the properties listed. This implies that every real estate agent records every details of any action on every property carefully. The quality of the data is exceptional. We have details about every property that was handled within the agency between January 2000 when they started implementing the paperless office and June 2004 (almost 2500 properties). In particular, unlike the first data set, we have information on all the properties that were listed, regardless of whether they sold or were withdrawn. For every listing, we know the same information as in the first data set. In addition, we know details about all the visits to the property. We know how many hits the property got on a weekly basis on the agency's website. We know if and when the property was advertised through the press and via mailings.

The agency also maintains files for each potential buyer. In particular, we know the date of their first enquiry in the agency, whether they are first-time buyer, buying to move in or to rent the property out, and the price range within which they are looking. We obviously have all the characteristics of the property that are posted by the agency on its website. In addition, we also know the valuation done by the agency's own appraiser. We have also mapped a small subset of the properties with a GIS system to obtain precise information on the area and shape of the parcel of land that supports each property and details about each location. We plan to implement this procedure for all the properties in our sample.

Finally, we have started to collect new data for the market of Madison, Wisconsin and are investigating opportunities to collect data in Chicago. Madison is a particularly interesting case study because it has a very well organized for-sale-by-owner market that is largely dominated by `fsbomadison.com`. The owners of this service have been very cooperative in providing us with data. We are also working closely with various realtors in Madison and their association to assemble data sets that are comparable to the ones we put together for England.

3 The Seller's Problem

This section presents our formulation of a discrete time finite-horizon dynamic programming problem of the seller's optimal strategy for selling a house. We take the decision to sell a house (via a real estate agency) as a given, and consider only the decision of which price to list the house at initially, how to revise this price over time, whether or not to accept offers that are made, and whether to withdraw the house if insufficiently attractive offers are realized. This restriction is motivated by our data, which is a self-selected sample of individuals who chose to list their homes with a real estate agency. We do not have the data on individuals who chose for sale by owner that would be required to estimate a model that endogenizes the choice of real estate agent, but this is definitely an important direction that we hope to extend our model in the future.

Our model differs from the model of Salant (1991) in a number of respects. First, our model has been designed with the specifics of the English real estate market in mind. Our model incorporates a fixed menu cost of changing list prices and models the within week bargaining process between a buyer and the seller as an alternating move bargaining subgame. Consistent with what we observe about real estate transations in many parts of England, the buyer makes the offer and the seller will either accept or reject it. Multiple stages of offers and counter offers interspersed with accept or reject decisions is consistent with what we observe in the English housing data. Our model can also accomodate the possibility of "auctions", i.e. situations where multiple buyers are bidding simultaneously for a home. However since these auction situations are relatively rare in our data, this initial version of the model ignores the possiblity of an auction and presumes that at most one buyer arrives in any given week.

We assume a 2 year horizon, so that if a house is not sold after 2 years, we assume that the house is withdrawn from sale and the seller obtains an exogenously specified "continuation value" representing the use value of owning (or renting) their home over a longer horizon beyond the 2 year decision horizon in this model. This continuation value may or may not equal the seller's belief about the *financial value* of their home. The financial values is the seller's expectation of the ultimate selling price of their home. While it is clear that the ultimate selling price is endogenously determined and partly under control of the seller, we can think of financial value is a realistic initial assessment on the part of the seller of the ultimate outcome of the process. Since the seller's optimal strategy will depend on the financial value, if the financial value is to represent a rational, internally consistent belief on the part of the seller, it will have to satisfy a fixed point condition that guarantees that the seller's financial value is a "self-fulfilling prophecy". Although we do not explicitly enforce this fixed point constraint in our solution of the dynamic programming problem, we verify below (via stochastic simulations) that it does hold for the estimated version of our model. However in future work, it would be possible to extend the dynamic programming problem to explicitly enforce a rationality constraint on the seller's estimate of the financial value of their home. However from our standpoint it is useful to allow for formulations that relax this constraint and thus be able to consider models where sellers do not have fully rational, self-consistent beliefs about the financial value of their homes. Indeed, allowing for inconsistent or "unrealistic" beliefs is an alternative way to explain why some home sellers set unrealistically high listing prices for their homes that would be distinct from the loss aversion approach discussed in the introduction. However as we show below, we do not need to appeal to any type of irrationality or assume sellers have unrealistic beliefs in order to provide an accurate explanation of the English housing data.

Let F_t denote the seller's beliefs about the financial value of their home after t weeks of being listed on the market. We assume that F_t is given by the equation

$$F_t = \exp\{X\beta + \nu_t\} \tag{1}$$

where X are the observed (time-invariant) characteristics of the home (the basis for the traditional hedonic regression prediction of the ultimate sales price discussed in section 2), and ν_t reflects the impact of time-varying variables that can affect the seller's view of their home's financial value. These time varying factors could include the effects of “macro shocks” that affect the entire housing market, regional or neighborhood level shocks such as increases in crime rates, or the impacts of new regional public infrastructure investments such as trains, roads, or subways, etc., as well as idiosyncratic house-specific factors. Ultimately, we would like to adopt a Bayesian learning approach to model the evolution of the financial value in a more formal manner. However for the purposes of this study, we will assume that after consultation with appraisers and the real estate agent, the seller has a firm value of the financial value of their home that does not vary over the course of their selling horizon. In this case $\nu_t = \nu_0$, $t = 1, 2, \ldots$ can be interpreted as time-invariant “random effect” incorporates other factors besides those in the observed X characteristics that affect the seller’s perception of the financial value of their home.

Recall the left panel of figure 2.3 that shows that the adjusted list price is a far more accurate predictor of the ultimate selling price of the home than the hedonic value, $\exp\{X\beta\}$. We assume that ν_0 is a lognormally distributed random variable that is independent of X and reflects factors observed by the seller that affects the seller’s perception of the financial value of their home that is not observed by us, the econometricians. Thus, we can conceptualize ν_0 as reflecting the seller’ *private information* about the financial value of their home that is not already captured in the observable characteristics X. In our estimation of the model we enforce a *rationality constraint* by estimating β via a log-linear regression of the final transaction price on the X characteristics, and assuming that $\exp\{\nu_0\}$ is a lognormally distributed random variable satisfying $E\{\exp(\nu_0)\} = 1$. Again, we can relax this restriction and allow for certain types of less than fully rational buyer behavior. For example if for a certain seller we have $E\{\exp(\nu_0)\} > 1$, we can intrepret this condition as corresponding to an “optimistic seller” who has an upward biased perception of the financial value of their home. However as we noted above, we do not find it necessary to allow such perceptual biases in order to provide a good approximation of the observed outcomes in the English housing data.

Prospective buyers may or may not agree with the seller’s (privately held) belief about the financial value of their home. Thus, we will shortly describe “offer distributions” for the value of offers to buy the home (if made) which will depend on F_0 and also on the current listing price P_t, but which will will not necessarily equal F_0 or P_t but instead relfect the buyer’s own idiosyncratic valuation of the house as well as strategic considerations about the buyer’s optimal search and purchasing strategy. However before we go into these details, we have enough structure already to begin to describe the seller’s decision problem.

Due to the fact that the seller’s optimal selling decision depends critically on the seller’s financial value F_0, which in turn depends on a very high dimensional vector of observed housing characteristics X as well as unobserved characteristics ν_0, straightforward attempts to solve the seller's problem while accounting for all of these variables immediately presents us with a significant "curse of dimensionality". In principle we could treat the estimated hedonic value $\exp\{X_i\hat{\beta}\}$ as a "fixed effect" relevant to property i and solve $N = 780$ individual dynamic programming problems, one for each of the 780 properties in our sample. However the problem is more complicated due to the existence of the unobserved random effect ν_0. This is a one dimensional unobserved random variable and in principle we would solve each of the 780 DP problems over a grid of possible values of ν_0, and thereby approximate the optimal selling strategy explicitly as a function of all possible values of the unobserved random effect ν_0, which would be "integrated out" in the empirical work we present in section 4.

However by imposing *linear homogeneity* assumption, we can solve a single DP problem for the optimal selling strategy where the values and states are defined as *ratios relative to the seller's financial value.* In particular, define the seller's current list price P_t to be the ratio of the actual list price divided by the seller's financial value F_0. Then $P_t = 1.0$ is equivalent to a list price that equals the financial value, and $P_t > 1.0$ corresponds to a list price that exceeds the financial value and so forth. The implicit assumption underlying the linear homogeneity assumption is that, at least within the limited and fairly homoegenous segment of the housing market in our data set, there are no relevant further "price subsegments" that have significantly different arrival rates and buyer behavior depending on whether the houses in these segments are more expensive "high end" homes or not. The homogeneity assumption reflects a reasonable assumption that arrival rates and buyer bidding behavior are driven mostly by the perception of whether a given home is perceived to be a "good deal" as reflected by the ratio of the list price to the financial value. However as we discuss below the actual bid submitted by a buyer will depend on the buyer's private valuation for the home (also expressed as a ratio of the financial value F_0).

Thus the ratio of the list price to the home's financial value can be viewed as a signal to prospective buyers about whether the home is likely to be a "good deal" or not. Arrival rates of matches will be a declining function of this price ratio, and the actual offers submitted by prospective buyers will depend on the list price and the financial value only via this same ratio in a way we will detail below.

To understand why arrival rates of matches and offers submitted by prospective buyers depend on the seller's financial value F_0 as well as the publicly posted list price, note that we model the arrival of *matches* where a match is defined as a buyer who makes an offer on the home. Matches are to be distinguished from *visits* where the real estate shows the home to a prospective buyer. We presume that prior to any match, a prospective buyer has visited the home and observed and verified its characteristics X as well as the unobserved characeristics ν_0, which we assume are common knowledge between the seller and each propective buyer — at least after the buyer has been able to visit the house. We assume that a buyer will have their own private value for the home and will make an offer for the home according to a piecewise linear bidding function that will be described below. We do not derive this bidding function from first principles, i.e. as the endogenous solution to a bargaining game between the seller and the buyer. Instead we treat buyers as *bidding automata* that behave according to fixed, but reasonable, rules to be described below. The piecewise linear bidding strategy was chosen to be simple, yet consistent with the basic facts of bidding behavior observed in the English housing data. But the main point to remember here, is that due to these timing assumptions, it seems reasonable to assume that buyers and sellers share common expectations about the financial value of the home F_0, even though our model is consistent with different buyers having different idiosyncratic valuations for the home. However once we assume that F_0 is common knowledge it is a small additional restriction to assume that buyer arrival rates and bids depend on the list price and the financial value only as a ratio of these two quantities.

Let $S_t(P_t, d_t)$ denote the expected discounted (optimal) value of selling the home at the start of week t, where the current ratio of the list price to the financial value is P_t, and where the duration since the last match is d_t, with $d_t = 0$ indicating a situation where no matches have occured yet. We will get into detail about the timing of decisions and the flow of information shortly, but already we can see that this formulation of the seller's problem has three state variables: 1) the current total time on the market t, 2) the duration since the last match d_t, and 3) the current list price to financial value ratio P_t. The value function $S_t(P_t, d_t)$ provides the value of the home as a ratio of the financial value, so to obtain the actual value and actual list price we simply multiply these values by F_0. Thus $F_0 S_t(P_t, d_t)$ is the present discounted value of the optimal selling strategy, and $F_0 P_t$ is the current list price, both measured in UK pounds. Via this "trick" we can account for substantial heterogeneity in actual list prices and seller valuations by solving just a single DP problem "in ratio form." However an important implication of this assumption is that timing of list price reductions and the percentage size of these reductions implied by the seller's optimal selling strategy are homogeneous of degree 0 in the list price and the financial value.

Our model of the optimal selling decision does not require the seller to sell their home within the 2 year horizon: we assume that the seller has the option to withdraw their home from the market at any time over the selling horizon. Since we do not model the default option of not selling one's house, we do not attempt to go into any detail and derive the form of the value to the seller of withdrawing their home from the market and pursing their next best option (e.g. continuing to live in the house, or renting the home). Instead we will simply invoke a flexible specification of the "continuation value" $W_t(P_t, \tau)$ of withdrawing a home from the market and pursuing the next best opportunity. The parameter τ can be viewed as the seller's "type" and our model can allow for different types of seller who have different continuation values. Fortunately, although our model can all for this possibility, we did not need to appeal to any type of unbobserved heterogeneity in seller types in order for the model to provide a good approximation to the behavior we observe in the English housing data. For this reason we will suppress τ to simplify the notation below.

The Bellman equation for the seller's problem is given in equation (2) below. The seller has 3 main decisions: 1) whether or not to withdraw the property, 2) if the seller opts not to withdraw the property, there is a decision about which list price to set at the beginning of each week the home is on the market, and 3) if a prospective buyer arrives within the week and makes an offer, the seller must determine whether or not to accept the offer, and if the seller rejects the offer and makes a counter offer, whether to accept the counter offer. We assume that the first two decisions are made at the start of each week and that the seller is unable to withdraw their home or change their list price during the remainder of the week. Within the week, if one or more offers arrive, the seller can engage in bargaining with the prospective buyer. The state variables in the model are 1) the listing price set in the previous week, P_t (once again, this is a ratio of the actual list price to the financial value of the home F_0), 2) the duration since the last offer d_t, and 3) the number of weeks since the home was listed, t. Let $S_t(P_t, d_t)$ denote the maximum expected present discounted value of an optimal selling strategy. We have

$$S_t(P_t, d_t) = \max\left[W_t(P_t), \max_P \left[u_t(P, P_t, d_t) + \beta E S_{t+1}(P, d_t)\right]\right] \tag{2}$$

The Bellman equation says that at each week t, the optimal selling strategy involves choosing the larger of 1) the continuation value of (permanently) withdrawing the home from the market, 2) or continuing to sell, choosing an optimal listing price P. The function $ES_{t+1}(P, d_t)$ is the conditional expectation of the week $t+1$ value function S_{t+1} conditional on the current state variables (P_t, d_t). The function $u_t(P, P_t, d_t)$ represents the current week "holding cost" to the seller of having their home on the market. It is the net utility (in money equivalent units) of owning the home less the "hassle costs" of having to show the house to prospective buyers (i.e. having to keep the house clean and tidy, having to vacate the house on short notice when a real estate agent wants to show the house to a prospective buyer, etc.).

We now write a formula for $ES_{t+1}(P,d_t)$ that represents the value of the within week bargaining subgame between the seller and buyer. In order to describe the equation for ES_{t+1}, we need to introduce some additional information to describe the seller's beliefs about the arrival of offers from buyers, the distribution of the size of the offer, and the probability that the buyer will "walk" (i.e. not make a counter offer and search for other houses) if the seller rejects the buyer's offer. Following Merlo and Ortalo-Magne (2004) we assume that the seller's only bargaining decision is to accept or reject offers made by buyers: the seller does not make a price "counter offer" if he/she rejects the buyer's offer. We assume that within a given week there are at most n possible stages of offers and accept/reject decisions betweeen the buyers and the seller. To simplify notiation, we write ES_{t+1} for the case with $n = 3$ within-week bargaining stages and where at most one buyer arrives and makes an offer on the home in any week.

Let $\lambda_t(P,d_t)$ denote the conditional probability that an offer will arrive within a week given that the seller set the list price to be P at the start of the week and the duration since the last offer is d_t. Let O_j be the highest offer received at stage j of the bargaining process. Let $f_j(O_j|O_{j-1},P,d_t)$ denote the seller's beliefs about the counteroffer the buyer would make at stage j given that the buyer did not walk in response to the seller's rejection of the buyer's initial offer. If the seller accepts offer O_j, let $N_t(O_j)$ denote the net sales proceeds (net of real estate commissions, taxes, and other transactions costs) received by the seller. The seller must decide whether to accept the net proceeds $N_t(O_j)$, thereby selling the home and terminating the selling process, or reject the offer and hope that the buyer will submit a more attractive counter offer, or that some better offer will arrive in some future week.

If a seller rejects the offer O_j, there is a probability $\omega_j(O_j,P,d_t)$ that the buyer will "walk" and not make a counter offer as a function of the last rejected offer, O_j, and the current state (P,d_t). With this notation we are ready to write the equation for the within week bargaging problem which determines ES_{t+1} and completes the Bellman equation. We have

$$ES_{t+1}(P,d_t) = \lambda_t(P,d_t)S_{t+1}(P,d_t) + [1-\lambda_t(P,d_t)]\int_{O_1} \max\left[N_t(O_1), ES^1_{t+1}(O_1,P,d_t)\right] f_1(O_1|P,d_t)dO_1. \tag{3}$$

The function $ES^1_{t+1}(O_1,P,d_t)$ is the expectation of the subsequent stages of the within-week bargaining subgame conditional on having received an initial offer of O_1 and conditional on the beginning of the week state vatriables, (P,d_t). We can write a recursion for these within-week expected value functions similar to the overall backward induction equation for Bellman's equation as a "within-period Bellman equation" as follows

$$
\begin{aligned}
ES^1_{t+1}(O_1,P,d_t) = \quad & \omega_1(O_1|P,d_t)S_{t+1}(P,d_t+1) + \\
& [1-\omega_1(O_1|P,d_t)]\int_{O_2} \max\left[N_t(O_2), ES^2_{t+1}(O_2,P,d_t)\right] f_2(O_2|O_1,P,d_t)dO_2. \quad (4)
\end{aligned}
$$

It should now be clear how to define the remaining within week expected value functions, $ES^j_{t+1},\ j = 3,\ldots,n$. Since we have assumed that $n = 3$ in our empirical analysis, we present the final step in the within-week Bellman equation below.

$$
\begin{aligned}
ES^2_{t+1}(O_2,P,d_t) = \quad & \omega_2(O_2|O_1,P,d_t)S_{t+1}(P,d_t+1) + \\
& [1-\omega_2(O_2|O_1,P,d_t)]\int_{O_3} \max\left[N_t(O_3), S_{t+1}(P,d_t+1)\right] f_3(O_3|O_2,P,d_t)dO_3. \quad (5)
\end{aligned}
$$

What equation (5) tells us is that after receiving 2 counteroffers and rejecting the second counteroffer O_2, the seller expects that with probabilty $\omega_2(O_2|O_1,P,d_t)$ the buyer will walk, so that the bargaining ends and the seller's expected value is simply the expectation of next periods' value $S_{t+1}(P,d_t+1)$. However with probability $1-\omega_2(O_2|O_1,P,d_t)$, the buyer will submit a final counteroffer O_3 which is a draw from the conditional density $f(O_3|O_2,P,d_t)$. Once O_3 is revealed to the seller, the seller can either take the offer and receive the net proceeds $N_t(O_3)$, or reject the offer, in which case the bargaining also ends and the seller's expected value is the next week value function, $S_{t+1}(P,d_t+1)$.

4 Models of Bidding by Prospective Buyers

Our initial intention was to develop a highly flexible model of buyer behavior that could be consistent with a wide range of theories of buyer behavior. We attempted to estimate the distibution of the first offer $f_1(O_1|P,d)$ and the conditional densities $f_j(O_j|O_{j-1},P,d)$ representing the improvement in bids when the seller rejects the previous bid and the buyer counteroffers at bidding stages 2 and 3 using non-parametric and semi-parametric estimation methods in a semi-parametric two-step approach to the estimation of our model of seller behavior.

Unfortunately this strategy did not work. Although we were able to estimate the bid densities f_j under fairly weak assumptions, when we used these estimated densities to solve for the optimal selling problem we obtained unreasonable results, including predictions that the seller should set unbounded high list prices.

One important fact about observed bidding behavior is that *there is a positive probability that a prospective buyer will submit a bid equal to the current list price.* In the English housing data, over 15 percent of all accepted offers are equal to the list price and over 10 percent of all *first* offers are equal to the list price. Further, we also observe offers in *excess* of the seller's list price. For example, over 2% of all first offers are above the list price, and nearly 4% of all accepted offers are higher than the list price prevailing when the offer was made.

Thus, any estimation of the offer distributions needs to account for mass points in the distribution, particularly at the list price. We found that we obtained unreasonable implications for the seller model even when we imposed a fair amount of parametric assumptions on the offer distributions, which were intended to help enforce "reasonable" behavioral implications for the seller.

One of these parametric models is a double beta distribution with a mass point at the list price. An example of the double beta density function for bids is presented in the left hand panel of figure 4.1 below. There is a right-skewed component of the bid distribution to the left of the list price mass point, and then a smaller left skewed beta distribution above this mass point. The most important part is the piece below the the list price, which captures the "underbidding" that is the predominant outcome of matches between a buyer and the seller. The right skewed beta component has as its support the interval $[.25, 1]$ where we have assumed that $P = 1$ is the current list price ratio for the house (corresponding to list price equal to the financial value of the home). The lower support .25 represents a bid equal to $1/4$ of the value of the list price of the home.

The distribution plotted in the left hand panel of figure 4.1 is actually a rescaled version of the double beta distribution. The figure does not include the mass point at the normalized list price ratio (equal to $P = 1$) due to problems with plotting density values and the mass point on the same scale. The beta density component to the left of the mass point at 1 has been scaled to have a total mass of .85, representing the probability that a bid will be strictly below the list price. The component of the beta distribution above 1 is scaled to have a total mass of .05, representing a 5% probability of receiving a bid strictly above the list price. The remaining mass is a 10% probability of receiving a bid equal to the list price.

Based on initial empirical work, we judged this double beta model to be a good approximation to the actual distribution of bids we observe in the London housing data. The double beta distribution was specified so that the probabilities of receiving a bid below, equal to, or strictly above the list price was given by a trinomial logit model and the (a, b) parameters of the beta distributions were specified as

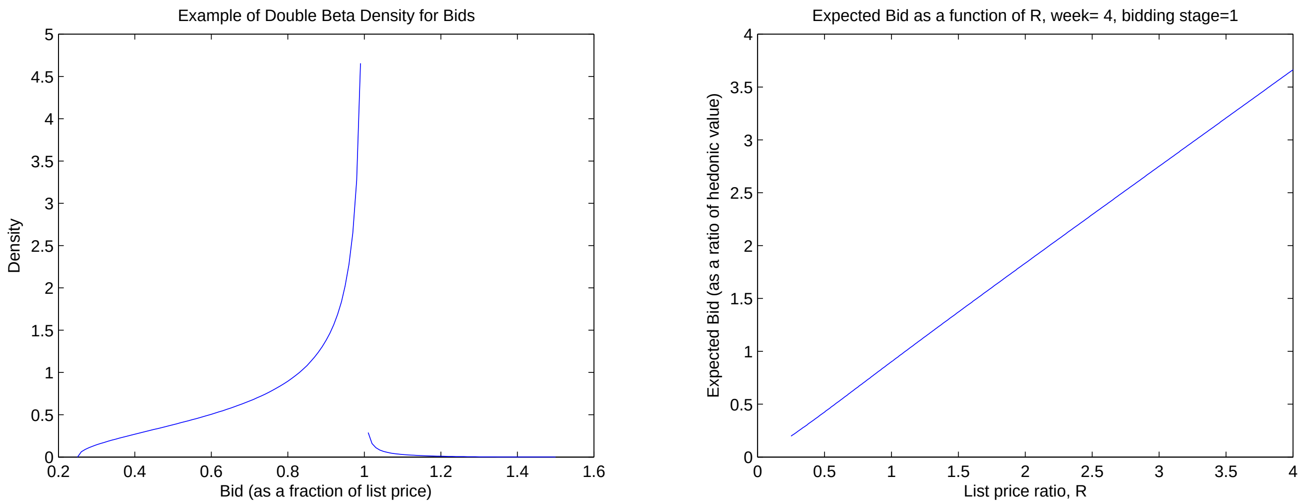

Figure 4.1 Double Beta Distribution of Bids and the implied expected bid function

(exponential) functions of state variables in the model (e.g. number of weeks on the market, the list price, and other variables). Unfortunately, as we see in the right hand panel of figure 4.1, the results of this model have unreasonable implications for sellers' beliefs about the relationship between the list price and the expected bid submitted by buyers. The expected bid function is a monotonically increasing function of the list price. It seems quite unreasonable that a seller should expect to receive to roughly double the expected bid on his house by doubling the list price, but this is exactly what the results from an unrestricted reduced form estimation of the offer distribution implies!

Further, our reduced form estimation results for the arrival rate of matches resulted in a *positive* relationship between list price and arrival rates of buyers, even after controlling for unobserved random effects, as represented by the ν_0 term in the seller's financial value of the home. Combining these two results, it is clear that any seller with such beliefs would find it optimal to set an aribtrarily large list price for their homes, something we rarely observe in practice. So clearly there is some problem with the flexible two step approach to estimating the seller model. The problems we experienced are probably not due to a misspecification of beliefs, since our reduced form model is a highly flexible specification capable of closely approximating the actual distribution of bids (and rates of arrival of matches). We believe the problem is due to the *endogeneity of list prices.* In particular, unobservable characteristics ν_0 that increase the financial value of a home also tend to increase the list price, and also bids made on a home. If we fail to control for these unobservables (as we have in our initial reduced form estimations), it is perfectly conceivable that the endogeneity problems could be strong enough to produce the spurious and implausible monotonic relationship between list price and expected bid values that we see in figure 4.1.

It might be possible to try to use more sophisticated reduced-from econometric methods to overcome the endogeneity problems. However it is clear that the seller's behavior is largely determined by the seller's beliefs about buyers: particularly how the list price affects their rate of arrival and what sort of bids they will make when they do arrive. Thus, there is a huge amount of information that can be brought to bear in estimating these rather slippery objects by adopting a fully structural, simultaneous approach to estimation where we estimate the sellers beliefs along with the other unknown parameters of the seller (e.g. the discount rate, and the parameters affecting hassle costs, and so forth) using nested numerical solution approach. Under this approach we would solve the seller's dynamic programming problem repeatedly for different trial values of the parameters governing the seller's beliefs as well as the other parameters of the model. Trial parameter values that produce "unreasonable" beliefs for the seller (such as shown in figure 4.1) would be discarded by this algorithm since these parameter values imply an optimal selling strategy that is greatly at odds with the behavior we observe in the data.

While it may ultimately be possible to estimate fairly flexible specifications for sellers' beliefs about buyer bids and arrival rates (such as the double beta distribution and even more flexible semiparametric specifications for the offer distributions), we have decided that it would be best to start by providing more structure on the bid distribution. There are two main reasons for this. First, even if we were able to successfully estimate the parameters of the double beta model as structural parameters in a maximum likelihood or simulated minimum distance estimator, there would be the issue of how to interpret these estimated coefficients in terms of an underlying model of bidder behavior. Instead, we felt that more insight could be gained by trying to build some sort of rudimentary model of bidding behavior on the part of buyers. By placing more structure on the offers distributions as we do below, we were able to acheive much more control over the estimation of the model making it much easier to estimate. The semi-reduced form model has fewer free parameters than the more flexibly specified reduced form models of bidding behavior, the parameters are more readily interpretable, and it is easier to see whether the estimated parameters are unreasonable or not, and how to constrain parameters to "reasonable" sections of the parameter space.

The "semi-reduced form model" of buyers' bidding behavior derives the distribution of bids from two underlying "structural" objects: 1) a specification of buyers' bid functions, $b(v,l,F)$, and 2) a specification of the distribution of buyer valuations, $h(v|F,l)$, where v is the buyer's private valuation of the home, F is the financial value of the home, and l is the current list price. In order to maintain the homogeneity restriction, we assume that l and F only enter b and h in a ratio form, i.e. as $p = l/F$. Thus, in the subsequent notation we will write these objects as $b(v,p)$ and $h(v|p)$.

We put "structural" in quotes because a fully structural model of buyer behavior would derive the buyers' bid functions from yet deeper structure: from the solution to their search and bargaining problem. We eventually want to extend the model in this direction, but since the English housing data contain relatively little data on buyers other than the bids they make in matches observed in the data set, it seems sensible to start out with a less complicated and detailed model of their behavior. In particular, since we do not have any data that follows buyers as they search among different homes and allow us to see homes they visit and don't make offers on and homes they visit and do make offers on, it seems that a more complicated buyer search model will have many additional parameters characterizing buyer search costs and opportunity sets and preferences for different locations and types of houses that we could have great difficulty in identifying from our (self-selected) data set of successful matches. This is our justification for failing to pursue a more detailed model of buyer behavior at this point.

The simplest specification for bid functions that we could think of that yields an offer distribution with a mass point at the current list price of the house is the following class of piecewise linear bid functions:

$$b(v,p) = \left\{ \begin{array}{ll} r_1(p)v & \text{if } v \in [\underline{v}, v_1) \\ p & \text{if } v \in [v_1, v_1 + k(p)) \\ r_2(p)v & \text{if } v \in [v_1 + k(p), \overline{v}] \end{array} \right], \tag{6}$$

where $\underline{v}$ and $\overline{v}$ are the lower and upper bounds, respectively, on the support of the distribution of buyer valuations (to be discussed shortly). To ensure continuity of $b(v,l)$ as a function of v, r_1 and r_2 must satisfy the following restrictions

$$\begin{aligned} p &= r_1(p)v_1 \\ p &= r_2(p)(v_1 + k(p)) \end{aligned} \tag{7}$$

This implies that

$$\begin{aligned} v_1 &= \frac{p}{r_1(p)} \\ r_2(p) &= \frac{p}{l/r_1(p) + k(p)} \end{aligned} \tag{8}$$

Thus, the bid functions are fully determined by the two functions $r_1(p)$ and $k(p)$. The first function determines how aggressive the bidder will be in terms of what fraction of the buyer's true valuation the buyer is willing to bid, *for the first bid* (We will consider specifications for 2nd and 3rd bid functions below). The closer $r_1(p)$ is to 1 the more "aggressive" the buyer is in his/her bidding (i.e. the closer they are to truthful bidding). We assume that the buyer interprets the list price l as a signal from the seller about what the seller's reservation value is and as a signal of how reasonable the seller is. If the list price ratio p is substantially bigger than 1, the buyer will interpret this as a sign of an "unreasonable" list price by the seller, and so the buyer will respond by shading their bid to a higher degree. Conversely, a seller that "underprices" their home by setting a list price less than the financial value will result in more aggressive bidding by buyers, i.e. $r_1(p)$ will be closer to 1 when $p < 1$. Thus, we posit that $r_1'(p) < 0$, so that a seller who considers overpricing their home will expect that buyers will shade their first bids to a greater degree.

The bid functions have a flat segment equal to the list price for valuations in the interval $[v_1, v_1 + k(p)]$. As we noted above, this flat section is empirically motivated by the fact that we observe a mass point in bid distributions at the list price. By adjusting the length of this flat segment $k(p)$ we can affect the size of the mass point in the bid distribution and thereby attempt to match observed bid distributions.

We posit that $k'(p) < 0$ for reasons similar to the assumption that $r_1'(p) \leq 0$: a seller who overprices his/her home by setting a list price bigger than 1 will result in a shorter range of valuations over which buyers would be willing to submit a first offer equal to the list price. Conversely, if a seller underprices his/her home by setting a list price less than 1, there should be a wider interval of valuations over which the buyer is willing to submit a first offer equal to the list price. Observe that since the probability of a first offer equal to the list price is the probability that valuations fall into the interval $[v_1, v_1 + k(p)]$, it is not strictly necessary for $k'(p) \leq 0$ in order for the probability of making an offer equal to the list price to be a declining function of l, which is another feature we observe in the English housing data. However initially we will assume that $k'(p) \leq 0$, but we can obviously consider relaxations of this condition later.

The left hand panel of Figure 4.2 plots examples of bid functions for four different values of p. These bid functions were generated from the following specifications for the function $r_1(p)$ and $k(p)$:

$$\begin{aligned} r_1(p) &= .98(1-\theta(p)) + .85\theta(p) \\ k(p) &= .12(1-\theta(p)) + .07\theta(p) \end{aligned} \tag{9}$$

where

$$\theta(p) = \frac{p - \underline{v}}{\overline{v} - \underline{v}}. \tag{10}$$

We see that the bid function for the highest list price, i.e. for a list price of $p = 1.62$ given by the blue dotted line in the left hand panel of figure 4.2, the bid function involves the most shading and it lies uniformly below the bid functions at other list prices. It follows that the list price of $p = 1.62$ is *dominated* in terms of revenue to the seller by lower list prices. However at more moderate list prices, the bid functions generally cross each other and so there is no unambiguous ranking based on strict dominance of the bid functions. For example if we compare the bid function for a list price of $p = 1$ with the bid function with a list price of $p = 1.09$ (the former is the orange dotted line and the latter is the solid red line in the left hand panel of fiugre 4.2), we see that the bid function for the lower list price $p = 1$ is higher for buyers with lower valuations and also for buyers with sufficiently high valuations, but the bid function with $p = 1.09$ (corresponding to a 9% markup over the financial value of the home), is higher for an intermediate range of buyer valuations. Thus the question of which of the two list prices result in higher expected revenues depends on the distribution of buyer valuations: if this distribution has sufficient mass in the intermediate range of buyer valuations where the bid function for the higher list price $p = 1.09$ exceeds the bid function for the lower list price $p = 1$, then the expected bid from setting the higher list price will exceed the expected bid from setting a lower list price. Of course this statement is *conditional* on a buyer arriving and making a bid: we need to factor in the impact of list price on the arrival rate to compute the overall expected revenue corresponding to different list prices.

The right hand panel of figure 4.2 shows how the bid functions change in successive bidding stages. Bid functions for later bidding stages dominate the bid functions for earlier bidding stages, resulting in a montonically increase sequence of bids that is consistent with what we almost always observe in the English housing data. However there are intervals of valuations where the bids lie on the flat segment of the bidding function, so this model can generate a sequence of bids where a previous bid (equal to the list price) is simply resubmitted by the bidder. This is also something we observe in the English housing data.

We complete the description of the semi-reduced form model by describing assumptions about the distribution of buyers' valuations for the home, $h(v|p)$. We assume that $h(v|p)$ is in the Beta family of distributions and thus it is fully specified by two parameters (a,b), as well as its support, $[\underline{v}, \overline{v}]$. We do not place any restriction on the distribution of valuations. In particular, it might be the case that buyers who have relatively higher than average valuations for a given home may choose to make offers: this

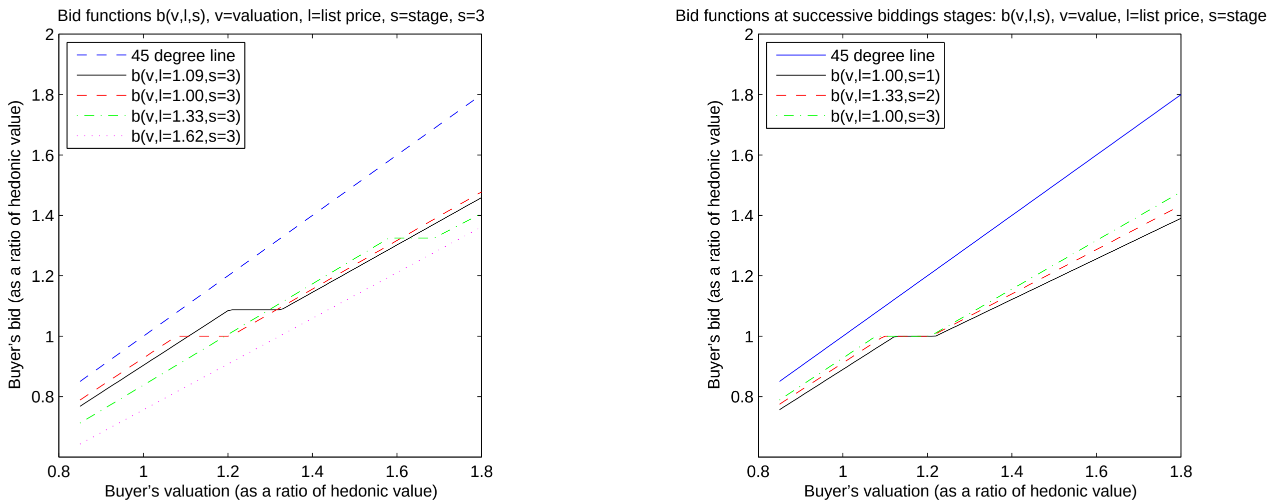

Figure 4.2 Piecewise linear bid functions for different list prices and bidding stages

would argue for a "positively biased" specification where $E\{v|p\} > p$. The direction of the bias might also depend on the list price: overpriced homes that have been on the market for a long time might be more likely to attract "vultures" i.e. buyers with lower than average valuations who are hoping to get a good deal if the seller "caves". We could imagine many other types of stories or scenarios. All of these suggest allowing for a more general model of valuations of the form $f_t(v|p,d)$ where the distribution of valuations of buyers who make an offer on a home with a price ratioof p also depends on the on the duration since the last offer d and the house has been listed, t.

While there is a value (in terms of additional flexiblity in the types of bid distributions that can be generated) by allowing for flexibility in the distribution of buyer valuations, it is clear that if we allow arbitrary amounts of flexibility then we might run into the same sorts of paradoxes that we illustrated for the fully reduced form specification of buyer bidding behavior. In particular if the distribution of buyer valuations shifts upward sufficiently quickly as the list price rises, then it is clearly possible that such a model could result in expected bids that are a monotonically increasing function of p, just as we observed in the double beta specification in figure 4.1. In addition there can be difficult identification problems since higher bids can be increased by either a) fixing a set of piecewise linear bid functions but shifting the distribution of valuation to the right, or b) fixing a distribution of valuations but allowing the piecewise bid functions to rise. For this reason, we have started by fixing the support and (a,b) parameters of the distribution of valuations and focus on estimating the parameters of the piecewise linear bid functions.

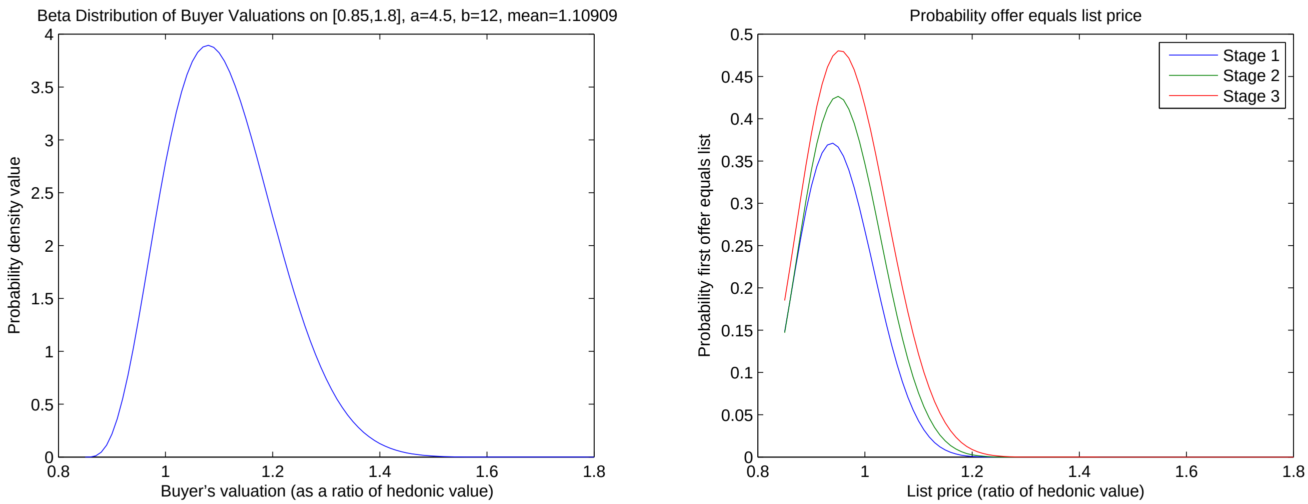

Figure 4.3 Beta distribution of buyer valuations and implied probabilities of bidding the list price

Let $B(u|a,b)$ be a beta distribution on the $[0,1]$ interval with parameers (a,b). We can derive the distribution of bids from this distribution by first rescaling this distriubtion to the $[\underline{v},\overline{v}]$ interval to get the distribution of valuations $H(v)$ given by

$$H(v) = Pr\{\tilde{v} \le v\} = B((v-\underline{v})/(\overline{v}-\underline{v})|a,b)\,. \tag{11}$$

The left hand panel of figure 4.3 plots an example of a beta distribution of valuations on the interval $[\underline{v},\overline{v}] = [.5,3]$ for different values of the (a,b) parameters. These parameters give us the flexiblity to affect both the mode and the tail behavior of the distributions independently of each other. For fixed a, increases in b decrease the expected value $E\{v\}$ and move the mode towards zero *and* thin out the upper tail, whereas for fixed b, increases in a increase the mode, the mean, and thickens the upper tail of $H(v)$ although larger changes are required in a to produce comparably dramatic shifts in $H(v)$ compared with changes in b, at least for $a > 1$.

The right hand panel of Figure 4.3 plots the implied probability that an offer equals the list price, as a function of p at successive stages of the within week bargaining process for buyers whose distribution of valuations is a beta distribution on the support $[.85,1.8]$ with parameters $(a,b) = (4.5,12)$. We see that these implied probabilities are roughly in line with the data for the limited range of list prices that we observe in the English housing data (i.e. a mean first offer that is roughly equal to the financial value, i.e. $E\{b(v,p)\} \simeq 1$, where the mean value of p is approximately equal to 1.05. This implies that $r_1(p) \simeq .95$ when $p \simeq .95$. Actually, for the specification of $r_1(p)$ given above, we have $r_1(1.05) = .9248$.

The implied distribution of bids, $G(x|a,b,l)$, is given by

$$\begin{aligned} G(x|a,b,l) &= Pr\{b(\tilde{v},l) \leq x\} \\ &= Pr\left\{\tilde{v} \leq b^{-1}(x,l)\right\} \\ &= B\left(b^{-1}(x,l) - \underline{v})/(\overline{v} - \underline{v})|a,b\right). \end{aligned} \tag{12}$$

Due to the presence of the flat segment, the usual notion of an inverse of the bid functon does not exist. However if we interpret the inverse of the bid function at the value p as the interval $[v_1, v_1 + k(l)]$, we obtain a distribution of bids that has the observe a mass point in the distribution of bids at the list price. That is, we can write the distribution of bids implied by this specification more explicitly in terms of the functions $r_1(p)$ and $k(p)$ as

$$G(x|a,b,p) = \begin{cases} B((x/r_1(p) - \underline{v})/(\overline{v} - \underline{v})|, a, b) & \text{if } x \in [\underline{v}, p) \\ B((k(p) + l/r_1(p) + k(p) - \underline{v})/(\overline{v} - \underline{v})|a,b) - B((l/r_1(p) - \underline{v})/(\overline{v} - \underline{v})|a,b) & \text{if } x = p \\ B((x(l/r_1(p) + k(p)) - \underline{v})/(\overline{v} - \underline{v})|a,b) & \text{if } x \in (p, \overline{v}] \end{cases} \tag{13}$$

Using this distribution function, we can compute the *expected bid function* $E\{\tilde{b}|p\}$ as

$$\begin{aligned} E\{\tilde{b}|p\} &= \int xG(dx|a,b,p) \\ &= \int_{\underline{v}}^{\overline{v}} b(v,p)H(dv). \end{aligned} \tag{14}$$

Note that expectation depends both on the list price and on the financial value because bids are interpreted as ratios of list price to the financial value of the home.

Figure 4.4 plots the expected bid functions for several different specifications of the distribution of valuations. We see that the expected bid functions are unimodal and are maximized at list prices that are higher than 1, providing an incentive for the seller to "overprice" when the seller sets a list price. Of course this is not the full story, since the seller must also account for the effect of the list price on arrival rates of buyers. The dynamic programming problem takes both factors into account, as well as other dynamic considerations and the fixed menu costs involved in changing the list price.

5 Empirical Results

Not yet written up: to be presented in the seminar

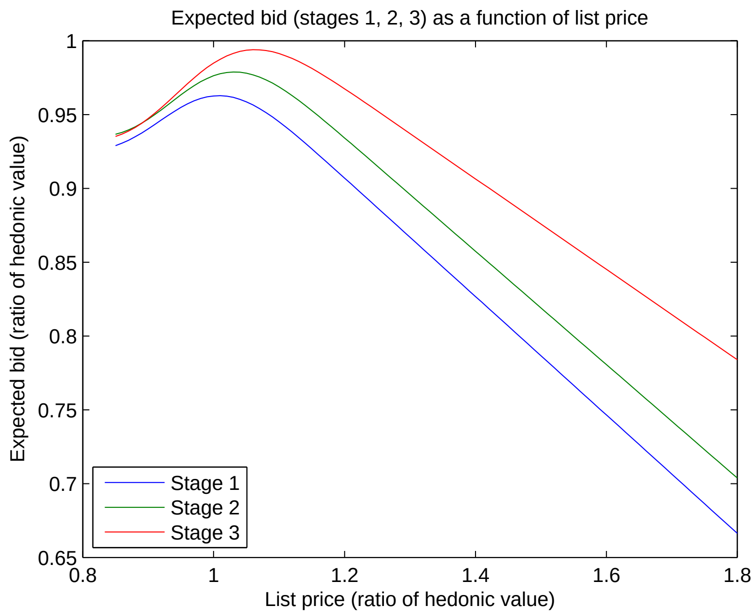

Figure 4.4 Expected bids as a function of the list price and bidding stage

6 Implications of the Model

Not yet written up: to be presented in the seminar

7 Conclusions

Not yet written up: to be presented in the seminar

References

[1] Albrecht, J., A. Anderson, E. Smith and S. Vroman (2006) "Opportunistic Matching in the Housing Market" manuscript, Georgetown University.

[2] Anglin, P.M., R. Rutherford, and T.A. Springer (2003) "The trade-off between the selling price of residential properties and the time-on-the-market" *Journal of Real Estate Finance and Economics* **26** 95–111.

[3] Arnold, M.A. (1999) "Search, bargaining and optimal asking prices" *Real Estate Economics* **27** 453–482.

[4] Bayer, P., R. McMillan, A. Murphy and C. Timmins (2007) "A Dynamic Model of the Housing Market" manuscript, Duke University.

[5] Bowman, D., Minehart, D. and M. Rabin (1999) "Loss aversion in a consumption savings model" *Journal of Economic Behavior and Organization* **38-2** 155–178.

[6] Brown, J.N. and H.S. Rosen (1982) "On the Estimation of Structural Hedonic Price Models" *Econometrica* **50-3** 765–768.

[7] Carillo, P.E. (2006) "An Empirical Two-sided Equilibrium Search Model of the Real Estate Market" manuscript, George Washington University.

[8] Genesove, D. and C.J. Mayer (1997) "Equity and time to sale in the real estate market" *American Economic Review* **87** 255–269.

[9] Genesove, D. and C.J. Mayer (2001) "Loss aversion and seller behavior: Evidence from the housing market" *Quarterly Journal of Economics* **116** 1233–1260.

[10] Glower, M., D.R. Haurin, and P.H. Hendershott (1998) "Selling time and selling price: The influence of seller motivation" *Real Estate Economics* **26** 719–740.

[11] D. Haurin (1988) "The duration of marketing time of residential housing" *Journal of the American Real Estate and Urban Economics Association* **16** 396–410.

[12] Horowitz, J.L. (1992) "The role of the list price in housing markets: Theory and econometric model" *Journal of Applied Econometrics* **7** 115–129.

[13] Kang, H., and M. Gardner, (1989) "Selling Price and Marketing Time in Residential Real Estate Market" *The Journal of Real Estate Research* **4** 21–35.

[14] Kahneman, D. and A. Tversky (1979) "Prospect Theory: An Analysis of Decision under Risk" *Econometrica* **48** 263–291.

[15] J.R. Knight, C.F. Sirmans, and G.K. Turnbull, (1998) "List price information in residential appraisal and underwriting" *Journal of Real Estate Research* **15** 59–76.

[16] Knight, J.R. (2002) "Listing price, time on market, and ultimate selling price: Causes and effects of listing price changes" *Real Estate Economics* **30** 213–237.

[17] Levitt, Steven D. and Chad Syverson (2002) "Market Distortions when Agents are Better Informed: A Theoretical and Empirical Exploration of the Value of Information in Real Estate Transactions" University of Chicago, mimeo.

[18] Miller, N.G. and M.A. Sklarz (1987) "Pricing strategies and residential property selling prices" *The Journal of Real Estate Research* **2** 31–40.

[19] Ortalo-Magne, F. and A. Merlo (2004) "Bargaining over Residential Real Estate: Evidence from England" *Journal of Urban Economics* **56** 192–216.

[20] Springer, T.E. (1996) "Single family housing transactions: Seller motivations, price and marketing time" *Journal of Real Estate Finance and Economics* **13** 237–254.

[21] Salant, S.W. (1991) "For Sale by Owner: When to Use a Broker and How to Price the House" *Journal of Real Estate Finance and Economics* **4** 157–173.

[22] Taylor, C.R. (1999) "Time-on-the-market as a sign of quality" *Review of Economic Studies* **66** 555–578.

[23] Yavaş, A. (1992) "A simple search and bargaining model of real estate markets" *Journal of the American Real Estate and Urban Economics Association* **20** 533–548.

[24] Yavaş, A. and S. Yang (1995) "The strategic role of listing price in marketing real estate: Theory and evidence" *Real Estate Economics* **23** 347–368.

[25] Wheaton, W.C. and G. Nechayev (2006) "Past Housing 'Cycles' and the Current Housing 'Boom': Whats Different This Time?" manuscript, Massachusetts Institute of Technology.

[26] Witte, A.D., H.J. Sumka and H. Erekson (1979) "An Estimate of a Structural Hedonic Price Model of the Housing Market: An Application of Rosen's Theory of Implicit Markets" *Econometrica* **47** 1151–1173.

[27] Zuehlke, T.W. (1987) "Duration dependence in the housing market" *The Review of Economics and Statistics* **69** 701–704.

[28] Zumpano, L., K. Johnson, and R. Anderson (2003) "Internet use and real estate brokerage market intermediation" *Journal of Housing Economics* **12** 134–150.